Path Integral Methods and Applications*

Richard MacKenzie†
Laboratoire René-J.-A.-Lévesque
Université de Montréal
Montréal, QC H3C 3J7 Canada

UdeM-GPP-TH-00-71

Abstract

These lectures are intended as an introduction to the technique of path integrals and their applications in physics. The audience is mainly first-year graduate students, and it is assumed that the reader has a good foundation in quantum mechanics. No prior exposure to path integrals is assumed, however.

The path integral is a formulation of quantum mechanics equivalent to the standard formulations, offering a new way of looking at the subject which is, arguably, more intuitive than the usual approaches. Applications of path integrals are as vast as those of quantum mechanics itself, including the quantum mechanics of a single particle, statistical mechanics, condensed matter physics and quantum field theory.

After an introduction including a very brief historical overview of the subject, we derive a path integral expression for the propagator in quantum mechanics, including the free particle and harmonic oscillator as examples. We then discuss a variety of applications, including path integrals in multiply-connected spaces, Euclidean path integrals and statistical mechanics, perturbation theory in quantum mechanics and in quantum field theory, and instantons via path integrals.

For the most part, the emphasis is on explicit calculations in the familiar setting of quantum mechanics, with some discussion (often brief and schematic) of how these ideas can be applied to more complicated situations such as field theory.

*Lectures given at Rencontres du Vietnam: VIth Vietnam School of Physics, Vung Tau, Vietnam, 27 December 1999 - 8 January 2000.

†rbmack@lps.umontreal.ca

1 Introduction

1.1 Historical remarks

We are all familiar with the standard formulations of quantum mechanics, developed more or less concurrently by Schroedinger, Heisenberg and others in the 1920s, and shown to be equivalent to one another soon thereafter.

In 1933, Dirac made the observation that the action plays a central role in classical mechanics (he considered the Lagrangian formulation of classical mechanics to be more fundamental than the Hamiltonian one), but that it seemed to have no important role in quantum mechanics as it was known at the time. He speculated on how this situation might be rectified, and he arrived at the conclusion that (in more modern language) the propagator in quantum mechanics "corresponds to" $\exp iS/\hbar$, where S is the classical action evaluated along the classical path.

In 1948, Feynman developed Dirac's suggestion, and succeeded in deriving a third formulation of quantum mechanics, based on the fact that the propagator can be written as a sum over all possible paths (not just the classical one) between the initial and final points. Each path contributes $\exp iS/\hbar$ to the propagator. So while Dirac considered only the classical path, Feynman showed that all paths contribute: in a sense, the quantum particle takes all paths, and the amplitudes for each path add according to the usual quantum mechanical rule for combining amplitudes. Feynman's original paper,[1] which essentially laid the foundation of the subject (and which was rejected by Physical Review!), is an all-time classic, and is highly recommended. (Dirac's original article is not bad, either.)

1.2 Motivation

What do we learn from path integrals? As far as I am aware, path integrals give us no dramatic new results in the quantum mechanics of a single particle. Indeed, most if not all calculations in quantum mechaincs which can be done by path integrals can be done with considerably greater ease using the standard formulations of quantum mechanics. (It is probably for this reason that path integrals are often left out of undergraduate-level quantum mechanics courses.) So why the fuss?

As I will mention shortly, path integrals turn out to be considerably more useful in more complicated situations, such as field theory. But even if this were not the case, I believe that path integrals would be a very worthwhile contribution to our understanding of quantum mechanics. Firstly, they provide a physically extremely appealing and intuitive way of viewing quantum mechanics: anyone who can understand Young's double slit experiment in optics should be able to understand the underlying ideas behind path integrals. Secondly, the classical limit of quantum mechanics can be understood in a particularly clean way via path integrals.

It is in quantum field theory, both relativistic and nonrelativistic, that path integrals (functional integrals is a more accurate term) play a much more important role, for several

[1] References are not cited in the text, but a short list of books and articles which I have found interesting and useful is given at the end of this article.

reasons. They provide a relatively easy road to quantization and to expressions for Green's functions, which are closely related to amplitudes for physical processes such as scattering and decays of particles. The path integral treatment of gauge field theories (non-abelian ones, in particular) is very elegant: gauge fixing and ghosts appear quite effortlessly. Also, there are a whole host of nonperturbative phenomena such as solitons and instantons that are most easily viewed via path integrals. Furthermore, the close relation between statistical mechanics and quantum mechanics, or statistical field theory and quantum field theory, is plainly visible via path integrals.

In these lectures, I will not have time to go into great detail into the many useful applications of path integrals in quantum field theory. Rather than attempting to discuss a wide variety of applications in field theory and condensed matter physics, and in so doing having to skimp on the ABCs of the subject, I have chosen to spend perhaps more time and effort than absolutely necessary showing path integrals in action (pardon the pun) in quantum mechanics. The main emphasis will be on quantum mechanical problems which are not necessarily interesting and useful in and of themselves, but whose principal value is that they resemble the calculation of similar objects in the more complex setting of quantum field theory, where explicit calculations would be much harder. Thus I hope to illustrate the main points, and some technical complications and hangups which arise, in relatively familiar situations that should be regarded as toy models analogous to some interesting contexts in field theory.

1.3 Outline

The outline of the lectures is as follows. In the next section I will begin with an introduction to path integrals in quantum mechanics, including some explicit examples such as the free particle and the harmonic oscillator. In Section 3, I will give a "derivation" of classical mechanics from quantum mechanics. In Section 4, I will discuss some applications of path integrals that are perhaps not so well-known, but nonetheless very amusing, namely, the case where the configuration space is not simply connected. (In spite of the fancy terminology, no prior knowledge of high-powered mathematics such as topology is assumed.) Specifically, I will apply the method to the Aharonov-Bohm effect, quantum statistics and anyons, and monopoles and charge quantization, where path integrals provide a beautifully intuitive approach. In Section 5, I will explain how one can approach statistical mechanics via path integrals. Next, I will discuss perturbation theory in quantum mechanics, where the technique used is (to put it mildly) rather cumbersome, but nonetheless illustrative for applications in the remaining sections. In Section 7, I will discuss Green's functions (vacuum expectation values of time-ordered products) in quantum mechanics (where, to my knowledge, they are not particularly useful), and will construct the generating functional for these objects. This groundwork will be put to good use in the following section, where the generating functional for Green's functions in field theory (which *are* useful!) will be elucidated. In Section 9, I will discuss instantons in quantum mechanics, and will at least pay lip service to important applications in field theory. I will finish with a summary and a list of embarrassing omissions.

I will conclude with a few apologies. First, an educated reader might get the impression that the outline given above contains for the most part standard material. S/he is likely correct: the only original content to these lectures is the errors.[2]

Second, I have made no great effort to give complete references (I know my limitations); at the end of this article I have listed some papers and books from which I have learned the subject. Some are books or articles wholly devoted to path integrals; the majority are books for which path integrals form only a small (but interesting!) part. The list is hopelessly incomplete; in particular, virtually any quantum field theory book from the last decade or so has a discussion of path integrals in it.

Third, the subject of path integrals can be a rather delicate one for the mathematical purist. I am not one, and I have neither the interest nor the expertise to go into detail about whether or not the path integral exists, in a strict sense. My approach is rather pragmatic: it works, so let's use it!

[2] Even this joke is borrowed from somewhere, though I can't think of where.

2 Path Integrals in Quantum Mechanics

2.1 General discussion

Consider a particle moving in one dimension, the Hamiltonian being of the usual form:

$$H = \frac{p^2}{2m} + V(q).$$

The fundamental question in the path integral (PI) formulation of quantum mechanics is: If the particle is at a position q at time $t = 0$, what is the probability amplitude that it will be at some other position q' at a later time $t = T$?

It is easy to get a formal expression for this amplitude in the usual Schroedinger formulation of quantum mechanics. Let us introduce the eigenstates of the position operator $\hat{q}$, which form a complete, orthonormal set:

$$\hat{q}\,|q\rangle = q\,|q\rangle\,, \qquad \langle q'|\,q\rangle = \delta(q'-q), \qquad \int dq\,|q\rangle\,\langle q| = 1.$$

(When there is the possibility of an ambiguity, operators will be written with a "hat"; otherwise the hat will be dropped.) Then the initial state is $|\psi(0)\rangle = |q\rangle$. Letting the state evolve in time and projecting on the state $|q'\rangle$, we get for the amplitude A,

$$A = \langle q'|\,\psi(T)\rangle \equiv K(q',T;q,0) = \langle q'|\,e^{-iHT}\,|q\rangle\,. \tag{1}$$

(Except where noted otherwise, $\hbar$ will be set to 1.) This object, for obvious reasons, is known as the propagator from the initial spacetime point $(q,0)$ to the final point (q',T). Clearly, the propagator is independent of the origin of time: $K(q',T+t;q,t) = K(q',T;q,0)$.

We will derive an expression for this amplitude in the form of a summation (integral, really) over all possible paths between the initial and final points. In so doing, we derive the PI from quantum mechanics. Historically, Feynman came up with the PI differently, and showed its equivalence to the usual formulations of quantum mechanics.

Let us separate the time evolution in the above amplitude into two smaller time evolutions, writing $e^{-iHT} = e^{-iH(T-t_1)}e^{-iHt_1}$. The amplitude becomes

$$A = \langle q'|\,e^{-iH(T-t_1)}e^{-iHt_1}\,|q\rangle\,.$$

Inserting a factor 1 in the form of a sum over the position eigenstates gives

$$\begin{aligned} A &= \langle q'|\,e^{-iH(T-t_1)}\underbrace{\int dq_1\,|q_1\rangle\,\langle q_1|}_{=1}\,e^{-iHt_1}\,|q\rangle \\ &= \int dq_1\,K(q',T;q_1,t_1)K(q_1,t_1;q,0). \end{aligned} \tag{2}$$

This formula is none other than an expression of the quantum mechanical rule for combining amplitudes: if a process can occur a number of ways, the amplitudes for each of these ways add. A particle, in propagating from q to q', must be *somewhere* at an intermediate time t_1; labelling that intermediate position q_1, we compute the amplitude for propagation via the point q_1 [this is the product of the two propagators in (2)] and integrate over all possible intermediate positions. This result is reminiscent of Young's double slit experiment, where the amplitudes for passing through each of the two slits combine and interfere. We will look at the double-slit experiment in more detail when we discuss the Aharonov-Bohm effect in Section 4.

We can repeat the division of the time interval T; let us divide it up into a large number N of time intervals of duration $\delta = T/N$. Then we can write for the propagator

$$A = \langle q'| \left(e^{-iH\delta}\right)^N |q\rangle = \langle q'| \underbrace{e^{-iH\delta}e^{-iH\delta}\cdots e^{-iH\delta}}_{N \text{ times}} |q\rangle .$$

We can again insert a complete set of states between each exponential, yielding

$$\begin{aligned} A &= \langle q'| e^{-iH\delta} \int dq_{N-1} |q_{N-1}\rangle \langle q_{N-1}| e^{-iH\delta} \int dq_{N-2} |q_{N-2}\rangle \langle q_{N-2}| \cdots \\ &\qquad \cdots \int dq_2 |q_2\rangle \langle q_2| e^{-iH\delta} \int dq_1 |q_1\rangle \langle q_1| e^{-iH\delta} |q\rangle \\ &= \int dq_1 \cdots dq_{N-1} \langle q'| e^{-iH\delta} |q_{N-1}\rangle \langle q_{N-1}| e^{-iH\delta} |q_{N-2}\rangle \cdots \\ &\qquad \cdots \langle q_1| e^{-iH\delta} |q\rangle \\ &\equiv \int dq_1 \cdots dq_{N-1} K_{q_N,q_{N-1}} K_{q_{N-1},q_{N-2}} \cdots K_{q_2,q_1} K_{q_1,q_0}, \end{aligned} \tag{3}$$

where we have defined $q_0 = q$, $q_N = q'$. (Note that these initial and final positions are *not* integrated over.) This expression says that the amplitude is the integral of the amplitude of all N-legged paths, as illustrated in Figure 1.

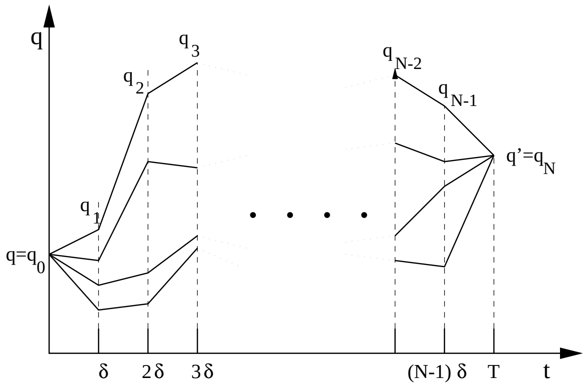

Figure 1: Amplitude as a sum over all N-legged paths.

Apart from mathematical details concerning the limit when $N \to \infty$, this is clearly going to become a sum over all possible paths of the amplitude for each path:

$$A = \sum_{\text{paths}} A_{\text{path}},$$

where

$$\sum_{\text{paths}} = \int dq_1 \cdots dq_{N-1}, \qquad A_{\text{path}} = K_{q_N,q_{N-1}} K_{q_{N-1},q_{N-2}} \cdots K_{q_2,q_1} K_{q_1,q_0}.$$

Let us look at this last expression in detail.

The propagator for one sub-interval is $K_{q_{j+1},q_j} = \langle q_{j+1}| e^{-iH\delta} |q_j\rangle$. We can expand the exponential, since δ is small:

$$\begin{aligned} K_{q_{j+1},q_j} &= \langle q_{j+1}| \left(1 - iH\delta - \frac{1}{2}H^2\delta^2 + \cdots\right) |q_j\rangle \\ &= \langle q_{j+1}| q_j\rangle - i\delta \langle q_{j+1}| H |q_j\rangle + o(\delta^2). \end{aligned} \tag{4}$$

The first term is a delta function, which we can write[3]

$$\langle q_{j+1}| q_j\rangle = \delta(q_{j+1} - q_j) = \int \frac{dp_j}{2\pi} e^{ip_j(q_{j+1}-q_j)}. \tag{5}$$

In the second term of (4), we can insert a factor 1 in the form of an integral over momentum eigenstates between H and $|q_j\rangle$; this gives

$$\begin{aligned} &-i\delta \langle q_{j+1}| \left(\frac{\hat{p}^2}{2m} + V(\hat{q})\right) \int \frac{dp_j}{2\pi} |p_j\rangle \langle p_j| q_j\rangle \\ &= -i\delta \int \frac{dp_j}{2\pi} \left(\frac{{p_j}^2}{2m} + V(q_{j+1})\right) \langle q_{j+1}| p_j\rangle \langle p_j| q_j\rangle \\ &= -i\delta \int \frac{dp_j}{2\pi} \left(\frac{{p_j}^2}{2m} + V(q_{j+1})\right) e^{ip_j(q_{j+1}-q_j)}, \end{aligned} \tag{6}$$

using $\langle q| p\rangle = \exp ipq$. In the first line, we view the operator $\hat{p}$ as operating to the right, while $V(\hat{q})$ operates to the left.

The expression (6) is asymmetric between q_j and q_{j+1}; the origin of this is our choice of putting the factor 1 to the right of H in the second term of (4). Had we put it to the left instead, we would have obtained $V(q_j)$ in (6). To not play favourites, we should choose some sort of average of these two. In what follows I will simply write $V(\bar{q}_j)$ where $\bar{q}_j = \frac{1}{2}(q_j + q_{j+1})$. (The exact choice does not matter in the continuum limit, which we will take eventually; the above is a common choice.) Combining (5) and (6), the sub-interval propagator is

$$\begin{aligned} K_{q_{j+1},q_j} &= \int \frac{dp_j}{2\pi} e^{ip_j(q_{j+1}-q_j)} \left(1 - i\delta\left(\frac{{p_j}^2}{2m} + V(\bar{q}_j)\right) + o(\delta^2)\right) \\ &= \int \frac{dp_j}{2\pi} e^{ip_j(q_{j+1}-q_j)} e^{-i\delta H(p_j,\bar{q}_j)}(1 + o(\delta^2)). \end{aligned} \tag{7}$$

There are N such factors in the amplitude. Combining them, and writing $\dot{q}_j = (q_{j+1} - q_j)/\delta$, we get

$$A_{\text{path}} = \int \prod_{j=0}^{N-1} \frac{dp_j}{2\pi} \exp i\delta \sum_{j=0}^{N-1} (p_j\dot{q}_j - H(p_j, \bar{q}_j)), \tag{8}$$

where we have neglected a multiplicative factor of the form $(1 + o(\delta^2))^N$, which will tend toward one in the continuum limit. Then the propagator becomes

$$\begin{aligned} K &= \int dq_1 \cdots dq_{N-1} A_{\text{path}} \\ &= \int \prod_{j=1}^{N-1} dq_j \int \prod_{j=0}^{N-1} \frac{dp_j}{2\pi} \exp i\delta \sum_{j=0}^{N-1} (p_j\dot{q}_j - H(p_j, \bar{q}_j)). \end{aligned} \tag{9}$$

[3] Please do not confuse the delta function with the time interval, δ.

Note that there is one momentum integral for each interval (N total), while there is one position integral for each intermediate position ($N-1$ total).

If $N \to \infty$, this approximates an integral over all functions $p(t)$, $q(t)$. We adopt the following notation:

$$\boxed{K \equiv \int \mathcal{D}p(t)\mathcal{D}q(t) \exp i \int_0^T dt\,(p\dot{q} - H(p,q))\,.} \tag{10}$$

This result is known as the *phase-space path integral.* The integral is viewed as over all functions $p(t)$ and over all functions $q(t)$ where $q(0) = q$, $q(T) = q'$. But to actually perform an explicit calculation, (10) should be viewed as a shorthand notation for the more ponderous expression (9), in the limit $N \to \infty$.

If, as is often the case (and as we have assumed in deriving the above expression), the Hamiltonian is of the standard form, namely $H = p^2/2m + V(q)$, we can actually carry out the momentum integrals in (9). We can rewrite this expression as

$$K = \int \prod_{j=1}^{N-1} dq_j \exp -i\delta \sum_{j=0}^{N-1} V(\bar{q}_j) \int \prod_{j=0}^{N-1} \frac{dp_j}{2\pi} \exp i\delta \sum_{j=0}^{N-1} \left(p_j \dot{q}_j - {p_j}^2/2m\right).$$

The p integrals are all Gaussian, and they are uncoupled. One such integral is

$$\int \frac{dp}{2\pi} e^{i\delta(p\dot{q} - p^2/2m)} = \sqrt{\frac{m}{2\pi i\delta}}\, e^{i\delta m\dot{q}^2/2}.$$

(The careful reader may be worried about the convergence of this integral; if so, a factor $\exp -\epsilon p^2$ can be introduced and the limit $\epsilon \to 0$ taken at the end.)

The propagator becomes

$$\begin{aligned} K &= \int \prod_{j=1}^{N-1} dq_j \exp -i\delta \sum_{j=0}^{N-1} V(\bar{q}_j) \prod_{j=0}^{N-1} \left(\sqrt{\frac{m}{2\pi i\delta}} \exp i\delta \frac{m\dot{q}_j^2}{2}\right) \\ &= \left(\frac{m}{2\pi i\delta}\right)^{N/2} \int \prod_{j=1}^{N-1} dq_j \exp i\delta \sum_{j=0}^{N-1} \left(\frac{m\dot{q}_j^2}{2} - V(\bar{q}_j)\right). \end{aligned} \tag{11}$$

The argument of the exponential is a discrete approximation of the action of a path passing through the points $q_0 = q, q_1, \cdots, q_{N-1}, q_N = q'$. As above, we can write this in the more compact form

$$\boxed{K = \int \mathcal{D}q(t) e^{iS[q(t)]}.} \tag{12}$$

This is our final result, and is known as the *configuration space path integral.* Again, (12) should be viewed as a notation for the more precise expression (11), as $N \to \infty$.

2.2 Examples

To solidify the notions above, let us consider a few explicit examples. As a first example, we will compute the free particle propagator first using ordinary quantum mechanics and then via the PI. We will then mention some generalizations which can be done in a similar manner.

2.2.1 Free particle

Let us compute the propagator $K(q', T; q, 0)$ for a free particle, described by the Hamiltonian $H = p^2/2m$. The propagator can be computed straightforwardly using ordinary quantum mechanics. To this end, we write

$$\begin{aligned} K &= \langle q'|\, e^{-iHT}\, |q\rangle \\ &= \langle q'|\, e^{-iT\hat{p}^2/2m} \int \frac{dp}{2\pi}\, |p\rangle \, \langle p|\, q\rangle \\ &= \int \frac{dp}{2\pi} e^{-iTp^2/2m}\, \langle q'|\, p\rangle \, \langle p|\, q\rangle \\ &= \int \frac{dp}{2\pi} e^{-iT(p^2/2m)+i(q'-q)p}. \end{aligned} \tag{13}$$

The integral is Gaussian; we obtain

$$K = \left(\frac{m}{2\pi i T}\right)^{1/2} e^{im(q'-q)^2/2T}. \tag{14}$$

Let us now see how the same result can be attained using PIs. The configuration space PI (12) is

$$\begin{aligned} K &= \lim_{N\to\infty} \left(\frac{m}{2\pi i\delta}\right)^{N/2} \int \prod_{j=1}^{N-1} dq_j \, \exp i\frac{m\delta}{2} \sum_{j=0}^{N-1} \left(\frac{q_{j+1}-q_j}{\delta}\right)^2 \\ &= \lim_{N\to\infty} \left(\frac{m}{2\pi i\delta}\right)^{N/2} \int \prod_{j=1}^{N-1} dq_j \, \exp i\frac{m}{2\delta}\Big[(q_N - q_{N-1})^2 + (q_{N-1}-q_{N-2})^2 + \cdots \\ &\qquad\qquad + (q_2 - q_1)^2 + (q_1 - q_0)^2\Big], \end{aligned}$$

where $q_0 = q$ and $q_N = q'$ are the initial and final points. The integrals are Gaussian, and can be evaluated exactly, although the fact that they are coupled complicates matters significantly. The result is

$$\begin{aligned} K &= \lim_{N\to\infty} \left(\frac{m}{2\pi i\delta}\right)^{N/2} \frac{1}{\sqrt{N}} \left(\frac{2\pi i\delta}{m}\right)^{(N-1)/2} e^{im(q'-q)^2/2N\delta} \\ &= \lim_{N\to\infty} \left(\frac{m}{2\pi i N\delta}\right)^{1/2} e^{im(q'-q)^2/2N\delta}. \end{aligned}$$

But $N\delta$ is the total time interval T, resulting in

$$K = \left(\frac{m}{2\pi i T}\right)^{1/2} e^{im(q'-q)^2/2T},$$

in agreement with (14).

A couple of remarks are in order. First, we can write the argument of the exponential as $T \cdot \frac{1}{2}m((q'-q)/T)^2$, which is just the action $S[q_c]$ for a particle moving along the classical path (a straight line in this case) between the initial and final points.

Secondly, we can restore the factors of $\hbar$ if we want, by ensuring correct dimensions. The argument of the exponential is the action, so in order to make it a pure number we must divide by $\hbar$; furthermore, the propagator has the dimension of the inner product of two position eigenstates, which is inverse length; in order that the coefficient have this dimension we must multiply by $\hbar^{-1/2}$. The final result is

$$\boxed{K = \left(\frac{m}{2\pi i\hbar T}\right)^{1/2} e^{iS[q_c]/\hbar}.} \tag{15}$$

This result typifies a couple of important features of calculations in this subject, which we will see repeatedly in these lectures. First, the propagator separates into two factors, one of which is the phase $\exp iS[q_c]/\hbar$. Second, calculations in the PI formalism are typically quite a bit more lengthy than using standard techniques of quantum mechanics.

2.2.2 Harmonic oscillator

As a second example of the computation of a PI, let us compute the propagator for the harmonic oscillator using this method. (In fact, we will not do the entire computation, but we will do enough to illustrate a trick or two which will be useful later on.)

Let us start with the somewhat-formal version of the configuration-space PI, (12):

$$K(q', T; q, 0) = \int \mathcal{D}q(t) e^{iS[q(t)]}.$$

For the harmonic oscillator,

$$S[q(t)] = \int_0^T dt \left(\frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2\right).$$

The paths over which the integral is to be performed go from $q(0) = q$ to $q(T) = q'$. To do this PI, suppose we know the solution of the classical problem, $q_c(t)$:

$$\ddot{q}_c + \omega^2 q_c = 0, \qquad q_c(0) = q, \qquad q_c(T) = q'.$$

We can write $q(t) = q_c(t) + y(t)$, and perform a change of variables in the PI to $y(t)$, since integrating over all deviations from the classical path is equivalent to integrating over all possible paths. Since at each time q and y differ by a constant, the Jacobian of the transformation is 1. Furthermore, since q_c obeys the correct boundary conditions, the paths $y(t)$ over which we integrate go from $y(0) = 0$ to $y(T) = 0$. The action for the path $q_c(t)+y(t)$ can be written as a power series in y:

$$S[q_c(t) + y(t)] \quad = \quad \int_0^T dt \left(\frac{1}{2}m\dot{q}_c^2 - \frac{1}{2}m\omega^2 {q_c}^2\right) + \underbrace{(\text{linear in } y)}_{=0} + \int_0^T dt \left(\frac{1}{2}m\dot{y}^2 - \frac{1}{2}m\omega^2 y^2\right).$$

The term linear in y vanishes by construction: q_c, being the classical path, is that path for which the action is stationary! So we may write $S[q_c(t) + y(t)] = S[q_c(t)] + S[y(t)]$. We substitute this into (12), yielding

$$K(q', T; q, 0) = e^{iS[q_c(t)]} \int \mathcal{D}y(t) e^{iS[y(t)]}. \tag{16}$$

As mentioned above, the paths $y(t)$ over which we integrate go from $y(0) = 0$ to $y(T) = 0$: the only appearance of the initial and final positions is in the classical path, *i.e.*, in the classical action. Once again, the PI separates into two factors. The first is written in terms of the action of the classical path, and the second is a PI over deviations from this classical path. The second factor is independent of the initial and final points.

This separation into a factor depending on the action of the classical path and a second one, a PI which is independent of the details of the classical path, is a recurring theme, and an important one. Indeed, it is often the first factor which contains most of the useful information contained in the propagator, and it can be deduced without even performing a PI. It can be said that much of the work in the game of path integrals consists in avoiding having to actually compute one!

As for the evaluation of (16), a number of fairly standard techniques are available. One can calculate the PI directly in position space, as was done above for the harmonic oscillator (see Schulman, chap. 6). Alternatively, one can compute it in Fourier space (writing $y(t) = \sum_k a_k \sin(k\pi t/T)$ and integrating over the coefficients $\{a_k\}$). This latter approach is outlined in Feynman and Hibbs, Section 3.11. The result is

$$K(q', T; q, 0) = \left(\frac{m\omega}{2\pi i \sin \omega T}\right)^{1/2} e^{iS[q_c(t)]}. \tag{17}$$

The classical action can be evaluated straightforwardly (note that this is not a PI problem, nor even a quantum mechanics problem!); the result is

$$S[q_c(t)] = \frac{m\omega}{2 \sin \omega T} \left((q'^2 + q^2)\cos \omega T - 2q'q\right).$$

We close this section with two remarks. First, the PI for any quadratic action can be evaluated exactly, essentially since such a PI consists of Gaussian integrals; the general result is given in Schulman, Chapter 6. In Section 6, we will evaluate (to the same degree of completeness as the harmonic oscillator above) the PI for a forced harmonic oscillator, which will prove to be a very useful tool for computing a variety of quantities of physical interest.

Second, the following fact is not difficult to prove, and will be used below (Section 4.2.). $K(q', T; q, 0)$ (whether computed via PIs or not) is the amplitude to propagate from one point to another in a given time interval. But this is the response to the following question: If a particle is initially at position q, what is its wave function after the elapse of a time T? Thus, if we consider K as a function of the final position and time, it is none other than the wave function for a particle with a specific initial condition. As such, the propagator satisfies the Schroedinger equation at its final point.

3 The Classical Limit: "Derivation" of the Principle of Least Action

Since the example calculations performed above are somewhat dry and mathematical, it is worth backing up a bit and staring at the expression for the configuration space PI, (12):

$$K = \int \mathcal{D}q(t) e^{iS[q(t)]/\hbar}.$$

This innocent-looking expression tells us something which is at first glance unbelievable, and at second glance *really* unbelievable. The first-glance observation is that a particle, in going from one position to another, takes all possible paths between these two positions. This is, if not actually unbelievable, at the very least least counter-intuitive, but we could argue away much of what makes us feel uneasy if we could convince ourselves that while all paths contribute, the classical path is the dominant one.

However, the second-glance observation is not reassuring: if we compare the contribution of the classical path (whose action is $S[q_c]$) with that of some other, arbitrarily wild, path (whose action is $S[q_w]$), we find that the first is $\exp iS[q_c]$ while the second is $\exp iS[q_w]$. They are both complex numbers of *unit magnitude*: each path taken in isolation is *equally important*. The classical path is no more important than any arbitrarily complicated path!

How are we to reconcile this *really* unbelievable conclusion with the fact that a ball thrown in the air has a more-or-less parabolic motion?

The key, not surprisingly, is in how different paths interfere with one another, and by considering the case where the rough scale of classical action of the problem is much bigger than the quantum of action, $\hbar$, we will see the emergence of the Principle of Least Action.

Consider two neighbouring paths $q(t)$ and $q'(t)$ which contribute to the PI (Figure 2). Let $q'(t) = q(t) + \eta(t)$, with $\eta(t)$ small. Then we can write the action as a functional Taylor expansion about the classical path:[4]

$$S[q'] = S[q+\eta] = S[q] + \int dt\, \eta(t) \frac{\delta S[q]}{\delta q(t)} + o(\eta^2).$$

The two paths contribute $\exp iS[q]/\hbar$ and $\exp iS[q']/\hbar$ to the PI; the combined contribution is

$$A \simeq e^{iS[q]/\hbar} \left(1 + \exp \frac{i}{\hbar} \int dt\, \eta(t) \frac{\delta S[q]}{\delta q(t)} \right),$$

where we have neglected corrections of order η^2. We see that the difference in phase between the two paths, which determines the interference between the two contributions, is $\hbar^{-1} \int dt\, \eta(t) \delta S[q]/\delta q(t)$.

We see that the smaller the value of $\hbar$, the larger the phase difference between two given paths. So even if the paths are *very* close together, so that the difference in actions is extremely small, for sufficiently small $\hbar$ the phase difference will still be large, and on average destructive interference occurs.

[4]The reader unfamiliar with manipulation of functionals need not despair; the only rule needed beyond standard calculus is the functional derivative: $\delta q(t)/\delta q(t') = \delta(t - t')$, where the last δ is the Dirac delta function.

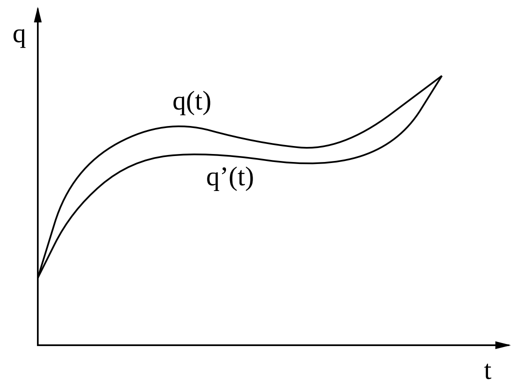

Figure 2: Two neighbouring paths.

However, this argument must be rethought for one exceptional path: that which extremizes the action, *i.e.*, the classical path, $q_c(t)$. For this path, $S[q_c + \eta] = S[q_c] + o(\eta^2)$. Thus the classical path and a very close neighbour will have actions which differ by much less than two randomly-chosen but equally close paths (Figure 3). This means that for fixed closeness

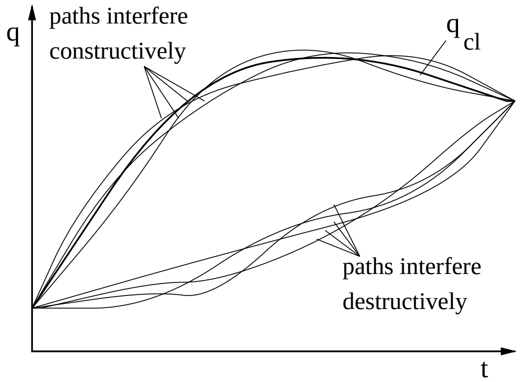

Figure 3: Paths near the classical path interfere constructively.

of two paths (I leave it as an exercise to make this precise!) and for fixed $\hbar$, paths near the classical path will on average interfere constructively (small phase difference) whereas for random paths the interference will be on average destructive.

Thus heuristically, we conclude that if the problem is classical (action $\gg \hbar$), the most important contribution to the PI comes from the region around the path which extremizes the PI. In other words, the particle's motion is governed by the principle that the action is stationary. This, of course, is none other than the Principle of Least Action from which the Euler-Lagrange equations of classical mechanics are derived.

4 Topology and Path Integrals in Quantum Mechanics: Three Applications

In path integrals, if the configuration space has holes in it such that two paths between the same initial and final point are not necessarily deformable into one another, interesting effects can arise. This property of the configuration space goes by the following catchy name: non-simply-connectedness. We will study three such situations: the Aharonov-Bohm effect, particle statistics, and magnetic monopoles and the quantization of electric charge.

4.1 Aharonov-Bohm effect

The Aharonov-Bohm effect is one of the most dramatic illustrations of a purely quantum effect: the influence of the electromagnetic potential on particle motion even if the particle is *perfectly* shielded from any electric or magnetic fields. While classically the effect of electric and magnetic fields can be understood purely in terms of the forces these fields create on particles, Aharonov and Bohm devised an ingenious thought-experiment (which has since been realized in the laboratory) showing that this is no longer true in quantum mechanics. Their effect is best illustrated by a refinement of Young's double-slit experiment, where particles passing through a barrier with two slits in it produce an interference pattern on a screen further downstream. Aharonov and Bohm proposed such an experiment performed with charged particles, with an added twist provided by a magnetic flux from which the particles are perfectly shielded passing between the two slits. If we perform the experiment

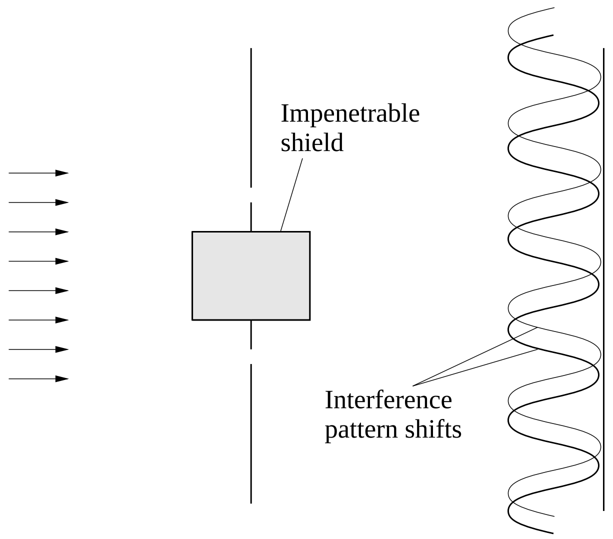

Figure 4: Aharonov-Bohm effect. Magnetic flux is confined within the shaded area; particles are excluded from this area by a perfect shield.

first with no magnetic flux and then with a nonzero and arbitrary flux passing through the shielded region, the interference pattern will change, *in spite of the fact that the particles are perfectly shielded from the magnetic field and feel no electric or magnetic force whatsoever.*

Classically we can say: no force, no effect. Not so in quantum mechanics. PIs provide a very attractive way of understanding this effect.

Consider first two representative paths $\mathbf{q}_1(t)$ and $\mathbf{q}_2(t)$ (in two dimensions) passing through slits 1 and 2, respectively, and which arrive at the same spot on the screen (Figure 5). Before turning on the magnetic field, let us suppose that the actions for these paths are $S[\mathbf{q}_1]$ and $S[\mathbf{q}_2]$. Then the interference of the amplitudes is determined by

$$e^{iS[\mathbf{q}_1]/\hbar} + e^{iS[\mathbf{q}_2]/\hbar} = e^{iS[\mathbf{q}_1]/\hbar}\left(1 + e^{i(S[\mathbf{q}_2]-S[\mathbf{q}_1])/\hbar}\right).$$

The relative phase is $\phi_{12} \equiv (S[\mathbf{q}_2] - S[\mathbf{q}_1])/\hbar$. Thus these two paths interfere constructively if $\phi_{12} = 2n\pi$, destructively if $\phi_{12} = (2n+1)\pi$, and in general there is partial cancellation between the two contributions.

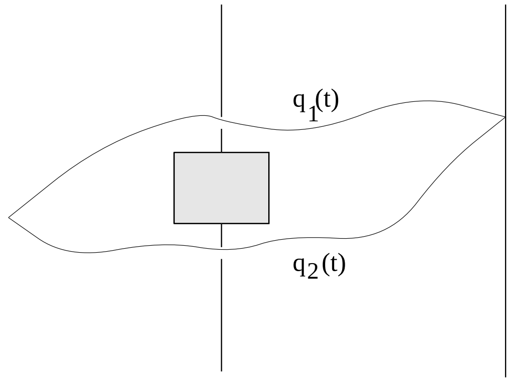

Figure 5: Two representative paths contributing to the amplitude for a given point on the screen.

How is this result affected if we add a magnetic field, $\mathbf{B}$? We can describe this field by a vector potential, writing $\mathbf{B} = \nabla \times \mathbf{A}$. This affects the particle's motion by the following change in the Lagrangian:

$$L(\dot{\mathbf{q}}, \mathbf{q}) \to L'(\dot{\mathbf{q}}, \mathbf{q}) = L(\dot{\mathbf{q}}, \mathbf{q}) - \frac{e}{c}\mathbf{v} \cdot \mathbf{A}(\mathbf{q}).$$

Thus the action changes by

$$-\frac{e}{c}\int dt\, \mathbf{v} \cdot \mathbf{A}(\mathbf{q}) = -\frac{e}{c}\int dt\, \frac{d\mathbf{q}(t)}{dt} \cdot \mathbf{A}(\mathbf{q}(t)).$$

This integral is $\int d\mathbf{q} \cdot \mathbf{A}(\mathbf{q})$, the line integral of $\mathbf{A}$ along the path taken by the particle. So including the effect of the magnetic field, the action of the first path is

$$S'[\mathbf{q}_1] = S[\mathbf{q}_1] - \frac{e}{c}\int_{\mathbf{q}_1(t)} d\mathbf{q} \cdot \mathbf{A}(\mathbf{q}),$$

and similarly for the second path.

Let us now look at the interference between the two paths, including the magnetic field.

$$\begin{aligned} e^{iS'[\mathbf{q}_1]/\hbar} + e^{iS'[\mathbf{q}_2]/\hbar} &= e^{iS'[\mathbf{q}_1]/\hbar}\left(1 + e^{i(S'[\mathbf{q}_2]-S'[\mathbf{q}_1])/\hbar}\right) \\ &= e^{iS'[\mathbf{q}_1]/\hbar}\left(1 + e^{i\phi'_{12}}\right), \end{aligned} \tag{18}$$

where the new relative phase is

$$\phi'_{12} = \phi_{12} - \frac{e}{\hbar c}\left(\int_{\mathbf{q}_2(t)} d\mathbf{q}\cdot\mathbf{A}(\mathbf{q}) - \int_{\mathbf{q}_1(t)} d\mathbf{q}\cdot\mathbf{A}(\mathbf{q})\right). \tag{19}$$

But the difference in line integrals in (19) is a contour integral:

$$\int_{\mathbf{q}_2(t)} d\mathbf{q}\cdot\mathbf{A}(\mathbf{q}) - \int_{\mathbf{q}_1(t)} d\mathbf{q}\cdot\mathbf{A}(\mathbf{q}) = \oint d\mathbf{q}\cdot\mathbf{A}(\mathbf{q}) = \Phi,$$

Φ being the flux inside the closed loop bounded by the two paths. So we can write

$$\phi'_{12} = \phi_{12} - \frac{e\Phi}{\hbar c}.$$

It is important to note that the change of relative phase due to the magnetic field is independent of the details of the two paths, as long as each passes through the corresponding slit. This means that the PI expression for the amplitude for the particle to reach a given point on the screen is affected by the magnetic field in a particularly clean way. Before the magnetic field is turned on, we may write $A = A_1 + A_2$, where

$$A_1 = \int_{\text{slit }1} \mathcal{D}\mathbf{q}\, e^{iS[\mathbf{q}]/\hbar},$$

and similarly for A_2. Including the magnetic field,

$$A'_1 = \int_{\text{slit }1} \mathcal{D}\mathbf{q}\, e^{i(S[\mathbf{q}]-(e/c)\int d\mathbf{q}\cdot\mathbf{A})/\hbar} = e^{-ie\int_1 d\mathbf{q}\cdot\mathbf{A}/\hbar c}A_1,$$

where we have pulled the line integral out of the PI since it is the same for all paths passing through slit 1 arriving at the point on the screen under consideration. So the amplitude is

$$\begin{aligned}
A &= e^{-ie\int_1 d\mathbf{q}\cdot\mathbf{A}/\hbar c}A_1 + e^{-ie\int_2 d\mathbf{q}\cdot\mathbf{A}/\hbar c}A_2 \\
&= e^{-ie\int_1 d\mathbf{q}\cdot\mathbf{A}/\hbar c}\left(A_1 + e^{-ie\oint d\mathbf{q}\cdot\mathbf{A}/\hbar c}A_2\right) \\
&= e^{-ie\int_1 d\mathbf{q}\cdot\mathbf{A}/\hbar c}\left(A_1 + e^{-ie\Phi/\hbar c}A_2\right).
\end{aligned}$$

The overall phase is irrelevant, and the interference pattern is influenced directly by the phase $e\Phi/\hbar c$. If we vary this phase continuously (by varying the magnetic flux), we can detect a shift in the interference pattern. For example, if $e\Phi/\hbar c = \pi$, then a spot on the screen which formerly corresponded to constructive interference will now be destructive, and vice-versa.

Since the interference is dependent only on the phase difference mod 2π, as we vary the flux we get a shift of the interference pattern which is periodic, repeating itself when $e\Phi/\hbar c$ changes by an integer times 2π.

4.2 Particle Statistics

The path integral can be used to see that particles in three dimensions must obey either Fermi or Bose statistics, whereas particles in two dimensions can have intermediate (or fractional) statistics. Consider a system of two identical particles; suppose that there is a short range, infinitely strong repulsive force between the two. We might ask the following question: if at $t = 0$ the particles are at $\mathbf{q}_1$ and $\mathbf{q}_2$, what is the amplitude that the particles will be at $\mathbf{q}'_1$ and $\mathbf{q}'_2$ at some later time T? We will first examine this question in three dimensions, and then in two dimensions.

4.2.1 Three dimensions

According to the PI description of the problem, this amplitude is

$$A = \sum_{\text{paths}} e^{iS[\mathbf{q}_1(t),\mathbf{q}_2(t)]},$$

where we sum over all two-particle paths going from $\mathbf{q}_1, \mathbf{q}_2$ to $\mathbf{q}'_1, \mathbf{q}'_2$.

However there is an important subtlety at play: if the particles are identical, then there are (in three dimensions!) two classes of paths (Figure 6).

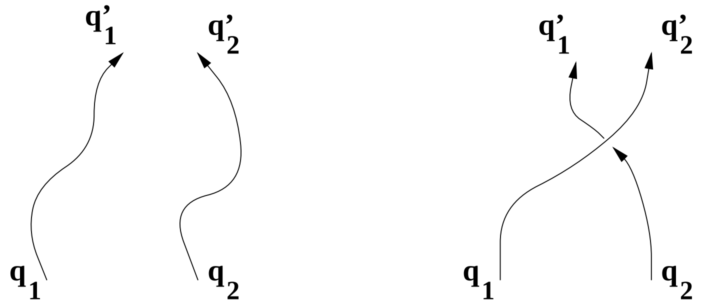

Figure 6: Two classes of paths.

Even though the second path involves an exchange of particles, the final configuration is the same due to the indistinguishability of the particles.

It is more economical to describe this situation in terms of the centre-of-mass position $\mathbf{Q} = (\mathbf{q}_1 + \mathbf{q}_2)/2$ and the relative position $\mathbf{q} = \mathbf{q}_2 - \mathbf{q}_1$. The movement of the centre of mass is irrelevant, and we can concentrate on the relative coordinate $\mathbf{q}$. We can also assume for simplicity that the final positions are the same as the initial ones. Then the two paths above correspond to the paths in relative position space depicted in Figure 7.

The point is, of course, that the relative positions $\mathbf{q}$ and $-\mathbf{q}$ represent the same configuration: interchanging $\mathbf{q}_1$ and $\mathbf{q}_2$ changes $\mathbf{q} \to -\mathbf{q}$.

We can elevate somewhat the tone of the discussion by introducing some amount of formalism. The configuration space for the relative position of two identical particles is *not* $\mathbf{R}^3 - \{0\}$,[5] as one would have naively thought, but $(\mathbf{R}^3 - \{0\})/\mathbf{Z}_2$. The division by the factor

[5] Recall that we have supposed that the particles have an infinite, short-range repulsion; hence the subtraction of the origin (which represents coincident points).

$\mathbf{Z}_2$ indicates that opposite points in the space $\mathbf{R}^3 - \{0\}$, namely any point $\mathbf{q}$ and the point diametrically opposite to it $-\mathbf{q}$, are to be *identified*: they represent the same configuration. We must keep this in mind when we attempt to draw paths: the second path of Figure 7 is a *closed* one.

A topological space such as our configuration space can be characterized as *simply connected* or as *non-simply connected* according to whether all paths starting and finishing at the same point can or cannot be contracted into the trivial path (representing no relative motion of the particles). It is clear from Figure 7 that the first path can be deformed to the trivial path, while the second one cannot, so the configuration space is not simply connected. Clearly any path which does not correspond to an exchange of the particles (a "direct" path) is topologically trivial, while any "exchange" path is not; we can divide the space of paths on the configuration space into two topological classes (direct and exchange).

Our configuration space is more precisely described as *doubly-connected*, since any two topologically nontrivial (exchange) paths taken one after the other result in a direct path, which is trivial. Thus the classes of paths form the elements of the group $\mathbf{Z}_2$ if we define the product of two paths to mean first one path followed by the other (a definition which extends readily to the product of classes).

One final bit of mathematical nomenclature: our configuration space, as noted above, is $(\mathbf{R}^3 - \{0\})/\mathbf{Z}_2$, which is not simply connected. We define the *(simply connected) covering space* as the simply connected space which looks locally like the original space. In our case, the covering space is just $\mathbf{R}^3 - \{0\}$.

At this point you might well be wondering: what does this have to do with PIs? We can rewrite the PI expression for the amplitude as the following PI in the covering space $\mathbf{R}^3 - \{0\}$:

$$\begin{aligned} A(\mathbf{q}, T; \mathbf{q}, 0) &= \sum_{\text{direct}} e^{iS[\mathbf{q}]} + \sum_{\text{exchange}} e^{iS[\mathbf{q}]} \\ &= \bar{A}(\mathbf{q}, T; \mathbf{q}, 0) + \bar{A}(-\mathbf{q}, T; \mathbf{q}, 0). \end{aligned} \tag{20}$$

The notation $\bar{A}$ is used to indicate that these PIs are in the covering space, while A is a PI in the configuration space. The first term is the sum over all paths from $\mathbf{q}$ to $\mathbf{q}$; the second is that for paths from $\mathbf{q}$ to $-\mathbf{q}$.

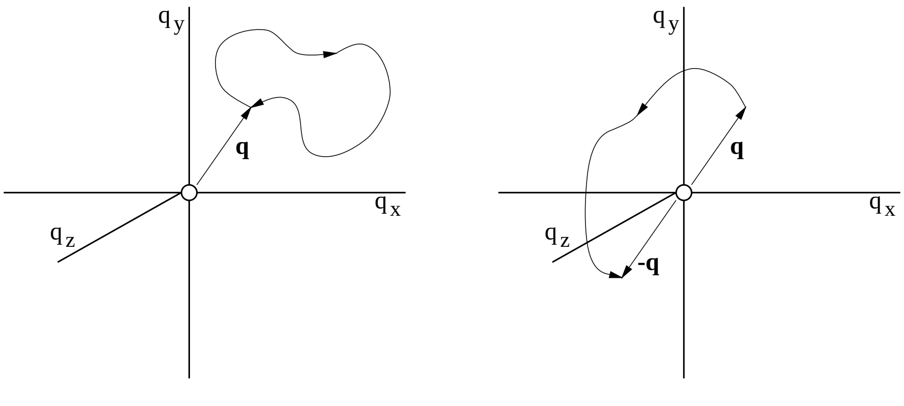

Figure 7: Paths in relative coordinate space.

Notice that each sub-path integral is a perfectly respectable PI in its own right: each would be a complete PI for the same dynamical problem but involving distinguishable particles. Since the PI can be thought of as a technique for obtaining the propagator in quantum mechanics, and since (as was mentioned at the end of Section 2) the propagator is a solution of the Schroedinger equation, either of these sub-path integrals also satisfies it. It follows that we can generalize the amplitude A to the following expression, which *still* satisfies the Schroedinger equation:

$$\begin{aligned} A(\mathbf{q},T;\mathbf{q},0) \to A^{\phi}(\mathbf{q},T;\mathbf{q},0) &= \sum_{\text{direct}} e^{iS[\mathbf{q}]} + e^{i\phi} \sum_{\text{exchange}} e^{iS[\mathbf{q}]} \\ &= \bar{A}(\mathbf{q},T;\mathbf{q},0) + e^{i\phi}\bar{A}(-\mathbf{q},T;\mathbf{q},0). \end{aligned} \tag{21}$$

This generalization might appear to be *ad hoc* and ill-motivated, but we will see shortly that it is intimately related to particle statistics.

There is a restriction on the added phase, ϕ. To see this, suppose that we no longer insist that the path be a closed one from $\mathbf{q}$ to $\mathbf{q}$. Then (21) generalizes to

$$A^{\phi}(\mathbf{q}',T;\mathbf{q},0) = \bar{A}(\mathbf{q}',T;\mathbf{q},0) + e^{i\phi}\bar{A}(-\mathbf{q}',T;\mathbf{q},0). \tag{22}$$

If we vary $\mathbf{q}'$ continuously to the point $-\mathbf{q}'$, we have

$$A^{\phi}(-\mathbf{q}',T;\mathbf{q},0) = \bar{A}(-\mathbf{q}',T;\mathbf{q},0) + e^{i\phi}\bar{A}(\mathbf{q}',T;\mathbf{q},0). \tag{23}$$

But since the particles are identical, the new final configuration $-\mathbf{q}'$ is identical to old one $\mathbf{q}'$. (22) and (23) are expressions for the amplitude for the same physical process, and can differ at most by a phase:

$$A^{\phi}(\mathbf{q}',T;\mathbf{q},0) = e^{i\alpha}A^{\phi}(-\mathbf{q}',T;\mathbf{q},0).$$

Combining these three equations, we see that

$$\bar{A}(\mathbf{q}',T;\mathbf{q},0) + e^{i\phi}\bar{A}(-\mathbf{q}',T;\mathbf{q},0) = e^{i\alpha}\left(\bar{A}(-\mathbf{q}',T;\mathbf{q},0) + e^{i\phi}\bar{A}(\mathbf{q}',T;\mathbf{q},0)\right).$$

Equating coefficients of the two terms, we have $\alpha = \phi$ (up to a 2π ambiguity), and

$$e^{i2\phi} = 1.$$

This equation has two physically distinct solutions: $\phi = 0$ and $\phi = \pi$. (Adding $2n\pi$ results in physically equivalent solutions.)

If $\phi = 0$, we obtain

$$A(\mathbf{q},T;\mathbf{q},0) = \bar{A}(\mathbf{q},T;\mathbf{q},0) + \bar{A}(-\mathbf{q},T;\mathbf{q},0), \tag{24}$$

the naive sum of the direct and exchange amplitudes, as is appropriate for Bose statistics.

If, on the other hand, $\phi = \pi$, we obtain

$$A(\mathbf{q},T;\mathbf{q},0) = \bar{A}(\mathbf{q},T;\mathbf{q},0) - \bar{A}(-\mathbf{q},T;\mathbf{q},0). \tag{25}$$

The direct and exchange amplitudes contribute with a relative minus sign. This case describes Fermi statistics.

In three dimensions, we see that the PI gives us an elegant way of seeing how these two types of quantum statistics arise.

4.2.2 Two dimensions

We will now repeat the above analysis in two dimensions, and will see that the difference is significant.

Consider a system of two identical particles in two dimensions, again adding a short-range, infinitely strong repulsion. Once again, we restrict ourselves to the centre of mass frame, since centre-of-mass motion is irrelevant to the present discussion. The amplitude that two particles starting at relative position $\mathbf{q} = (q_x, q_y)$ will propagate to a final relative position $\mathbf{q}' = (q'_x, q'_y)$ in time T is

$$A(\mathbf{q}', T; \mathbf{q}, 0) = \sum_{\text{paths}} e^{iS[\mathbf{q}_1(t), \mathbf{q}_2(t)]},$$

the sum being over all paths from $\mathbf{q}$ to $\mathbf{q}'$ in the configuration space.

Once again, the PI separates into distinct topological classes, but there are now an infinity of possible classes. To see this, consider the three paths depicted in Figure 8, where for simplicity we restrict to the case where the initial and final configurations are the same. It is important to remember that drawing paths in the plane is somewhat misleading: as in three dimensions, opposite points are identified, so that a path from any point to the diametrically opposite point is closed.

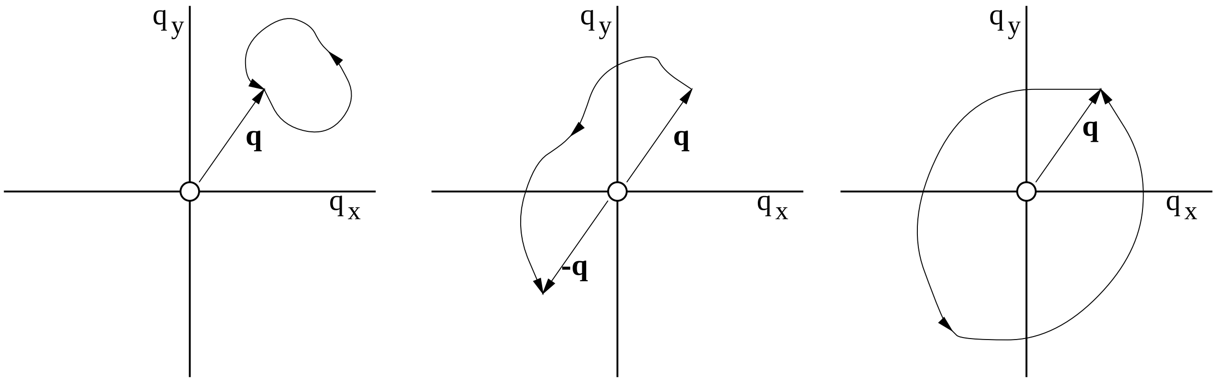

Figure 8: Three topologically distinct paths in two dimensions.

The first and second paths are similar to the direct and exchange paths of the three-dimensional problem. The third path, however, represents a distinct class of path in two dimensions. The particles circle around each other, returning to their starting points.

It is perhaps easier to visualize these paths in a three-dimensional space-time plot, where the vertical axis represents time and the horizontal axes represent space (Figure 9).

It is clear that in the third path the particle initially at $\mathbf{q}_1$ returns to $\mathbf{q}_1$, and similarly for the other particle: this path does not involve a permutation of the particles. It is also clear that this path cannot be continuously deformed into the first path, so it is in a distinct topological class. (It is critical here that we have excised the origin in relative coordinates – *i.e.*, that we have disallowed configurations where the two particles are at the same point in space.)

The existence of this third class of paths generalizes in an obvious way, and we are led to the following conclusion: the paths starting and finishing at relative position $\mathbf{q}$ can be divided into an infinite number of classes of paths in the plane (minus the origin); a class is specified by the number of interchanges of the particles (keeping track of the sense of each interchange). This is profoundly different from the three-dimensional case, where there were only two classes of paths: direct and exchange.

If we characterize a path by the polar angle of the relative coordinate, this angle is $n\pi$ in the n^{th} class, where n is an integer. (For the three paths shown above, $n = 0$, 1, and 2, respectively.)

We can write

$$A(\mathbf{q}, T; \mathbf{q}, 0) = \sum_{n=-\infty}^{\infty} \bar{A}_n(\mathbf{q}, T; \mathbf{q}, 0),$$

where $\bar{A}_n$ is the covering-space PI considering only paths of change of polar angle $n\pi$.

This path integral can again be generalized to

$$A(\mathbf{q}, T; \mathbf{q}, 0) = \sum_{n=-\infty}^{\infty} C_n \bar{A}_n(\mathbf{q}, T; \mathbf{q}, 0),$$

C_n being phases. Since each $\bar{A}_n$ satisfies the Schroedinger equation, so does this generalization.

Again, a restriction on the phases arises, as can be seen by the following argument. Let us relax the condition that the initial and final points are the same; let us denote the final point $\mathbf{q}'$ by its polar coordinates (q', θ'). Writing $A(q', \theta') \equiv A(\mathbf{q}', T; \mathbf{q}, 0)$, we have

$$A(q', \theta') = \sum_{-\infty}^{\infty} C_n \bar{A}_n(q', \theta').$$

Now, we can change continuously $\theta' \to \theta' + \pi$, yielding

$$A(q', \theta' + \pi) = \sum_{n=-\infty}^{\infty} C_n \bar{A}_n(q', \theta' + \pi). \tag{26}$$

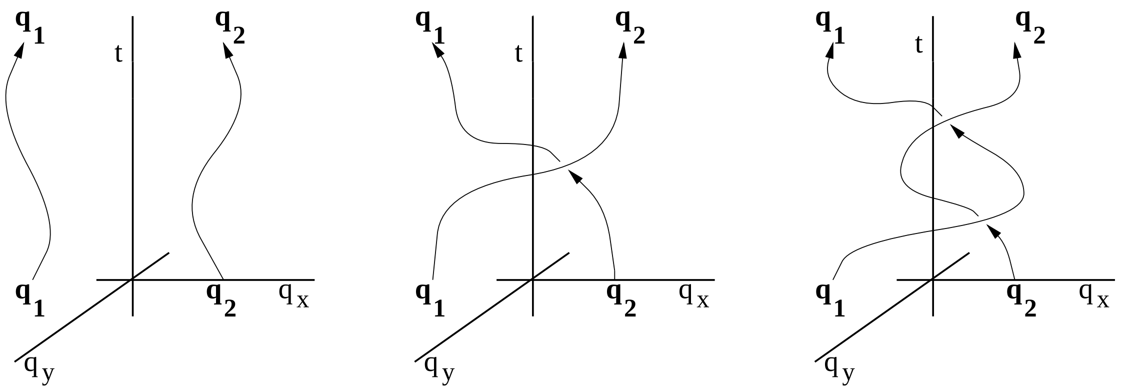

Figure 9: Spacetime depiction of the 3 paths in Figure 8.

Two critical observations can now be made. First, the final configuration is unchanged, so $A(q', \theta' + \pi)$ can differ from $A(q', \theta')$ by at most a phase:

$$A(q', \theta' + \pi) = e^{-i\phi} A(q', \theta').$$

Second, $\bar{A}_n(q', \theta' + \pi) = \bar{A}_{n+1}(q', \theta')$, since this is just two different ways of expressing exactly the same quantity. Applying these two observations to (26),

$$\begin{aligned} e^{-i\phi} \sum_{n=-\infty}^{\infty} C_n \bar{A}_n(q', \theta') &= \sum_{-\infty}^{\infty} C_n \bar{A}_{n+1}(q', \theta') \\ &= \sum_{-\infty}^{\infty} C_{n-1} \bar{A}_n(q', \theta'). \end{aligned} \tag{27}$$

Equating coefficients of $\bar{A}_n(q', \theta')$, we get

$$C_n = e^{i\phi} C_{n-1}.$$

Choosing $C_0 = 1$, we obtain for the amplitude

$$A = \sum_{-\infty}^{\infty} e^{in\phi} \bar{A}_n, \tag{28}$$

which is the two-dimensional analog of the three-dimensional result (21).

The most important observation to be made is that there is no longer a restriction on the angle ϕ, as was the case in three dimensions. We see that, relative to the "naive" PI (that with $\phi = 0$), the class corresponding to a net number n of counter-clockwise rotations of one particle around the other contributes with an extra phase $\exp in\phi$. If $\phi = 0$ or π, this collapses to the usual cases of Bose and Fermi statistics, respectively. However in the general case the phase relation between different paths is more complicated (not determined by whether the path is "direct" or "exchange"); this new possibility is known as fractional statistics, and particles obeying these statistics are known as anyons.

Anyons figure prominently in the accepted theory of the fractional quantum Hall effect, and were proposed as being relevant to high-temperature superconductivity, although that possibility seems not to be borne out by experimental results. Perhaps Nature has other applications of fractional statistics which await discovery.

4.3 Magnetic Monopoles and Charge Quantization

All experimental evidence so far tells us that all particles have electric charges which are integer multiples of a fundamental unit of electric charge, e.[6] There is absolutely nothing wrong with a theory of electrodynamics of particles of arbitrary charges: we could have particles of charge e and $\sqrt{17}e$, for example. It was a great mystery why charge was quantized in the early days of quantum mechanics.

In 1931, Dirac showed that the quantum mechanics of charged particles in the presence of magnetic monopoles is problematic, unless the product of the electric and magnetic charges

[6]The unit of electric charge is more properly $e/3$, that of the quarks; for simplicity, I will ignore this fact.

is an integer multiple of a given fundamental value. Thus, the existence of monopoles implies quantization of electric charge, a fact which has fueled experimental searches for and theoretical speculations about magnetic monopoles ever since.

We will now recast Dirac's argument into a modern form in terms of PIs. A monopole of charge g positioned at the origin has magnetic field

$$\mathbf{B} = g\frac{\hat{\mathbf{e}}_r}{r^2}.$$

As is well known, this field cannot be described by a normal (smooth, single-valued) vector potential: writing $\mathbf{B} = \nabla \times \mathbf{A}$ implies that the magnetic flux emerging from any closed surface (and thus the magnetic charge contained in any such surface) must be zero. This fact makes life difficult for monopole physics, for several reasons. Although classically the Maxwell equations and the Lorentz force equation form a complete set of equations for particles (both electrically and magnetically charged, with the simple addition of magnetic source terms) and electromagnetic fields, their derivation from an action principle requires that the electromagnetic field be described in terms of the electromagnetic potential, A_μ. Quantum mechanically, things are even more severe: one cannot avoid A_μ, because the coupling of a particle to the electromagnetic field is written in terms of A_μ, not the electric and magnetic fields.

Dirac suggested that if a monopole exists, it could be described by an infinitely-thin and tightly-wound solenoid carrying a magnetic flux equal to that of the monopole. The solenoid is semi-infinite in length, running from the position of the monopole to infinity along an arbitrary path. The magnetic field produced by such a solenoid can be shown to be that of a monopole plus the usual field produced by a solenoid, in this case, an infinitely intense, infinitely narrow tube of flux running from the monopole to infinity along the position of the solenoid (Figure 10). Thus except inside the solenoid, the field produced is that of the monopole. The field inside the solenoid is known as the "Dirac string". The flux brought

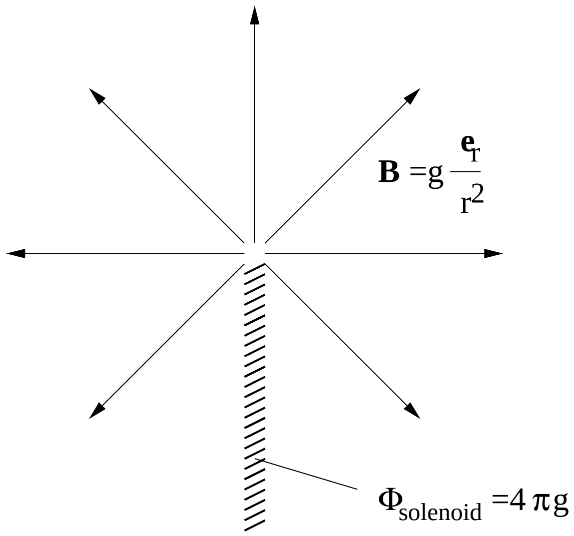

Figure 10: Monopole as represented by a semi-infinite, infinitely tightly-wound solenoid.

into any closed surface including the monopole is now zero, because the solenoid brings in a flux equal to that flowing out due to the monopole. Thus, the combined monopole-solenoid can be described by a vector potential.

However, in order for this to be a valid description of the monopole, we must somehow convince ourselves that the solenoid can be made invisible to any electrically charged particle passing by it. We can describe the motion of such a particle by a PI, and two paths passing on either side of the Dirac string can contribute to the PI (Figure 11). But the vector potential of the Dirac string will affect the action of each of these paths differently, as we have seen in the Aharonov-Bohm effect.

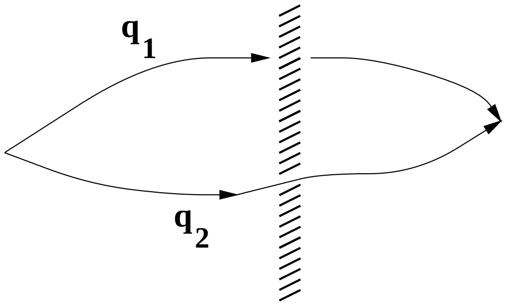

Figure 11: Paths contributing to the propagator in the presence of a monopole. The paths form a loop encircling the Dirac string.

In order that the interference of these paths be unaffected by the presence of the Dirac string, the relative phase must be an integral multiple of 2π. This phase is

$$-\frac{e}{\hbar c}\oint d\mathbf{q}\cdot\mathbf{A} = -\frac{e\Phi}{\hbar c} = -\frac{4\pi e g}{\hbar c}.$$

Setting this to $2\pi n$, in order for the motion of a particle of charge e to be unaffected by the presence of the Dirac string, the electric charge must be

$$e = \frac{2\pi\hbar c}{4\pi g}n = \frac{\hbar c}{2g}n. \tag{29}$$

Thus, the existence of magnetic monopoles requires the quantization of electric charge; the fundamental unit of electric charge is $2\pi\hbar c/g$.

In modern theories of fundamental physics, Grand Unified Theories also imply quantization of electric charge, apparently avoiding the necessity for magnetic monopoles. But any Grand Unified Theory actually has magnetic monopoles as well (though they are of a nature quite different to the "Dirac monopole"), so the intimate relation between magnetic monopoles and the quantization of electric charge is preserved, albeit in a form quite different from that suggested by Dirac.

5 Statistical Mechanics via Path Integrals

The path integral turns out to provide an elegant way of doing statistical mechanics. The reason for this is that, as we will see, the central object in statistical mechanics, the partition function, can be written as a PI. Many books have been written on statistical mechanics with emphasis on path integrals, and the objective in this lecture is simply to see the relation between the partition function and the PI.

The definition of the partition function is

$$Z = \sum_j e^{-\beta E_j}, \tag{30}$$

where $\beta = 1/k_B T$ and E_j is the energy of the state $|j\rangle$. We can write

$$Z = \sum_j \langle j| \, e^{-\beta H} \, |j\rangle = \mathrm{Tr} e^{-\beta H}.$$

But recall the definition of the propagator:

$$K(q', T; q, 0) = \langle q'| \, e^{-iHT} \, |q\rangle \, .$$

Suppose we consider T to be a complex parameter, and consider it to be pure imaginary, so that we can write $T = -i\beta$, where β is real. Then

$$\begin{aligned}
K(q', -i\beta; q, 0) &= \langle q'| \, e^{-iH(-i\beta)} \, |q\rangle \\
&= \langle q'| \, e^{-\beta H} \underbrace{\sum_j |j\rangle \, \langle j| \, |q\rangle}_{=1} \\
&= \sum_j e^{-\beta E_j} \, \langle q'| \, j\rangle \langle j \, |q\rangle \\
&= \sum_j e^{-\beta E_j} \langle j \, |q\rangle \, \langle q'| \, j\rangle.
\end{aligned}$$

Putting $q' = q$ and integrating over q, we get

$$\int dq \, K(q, -i\beta; q, 0) = \sum_j e^{-\beta E_j} \, \langle j| \underbrace{\int dq \, |q\rangle \, \langle q|}_{=1} |j\rangle = Z. \tag{31}$$

This is the central observation of this section: that the propagator evaluated at negative imaginary time is related to the partition function.

We can easily work out an elementary example such as the harmonic oscillator. Recall the path integral for it, (17):

$$K(q', T; q, 0) = \left(\frac{m\omega}{2\pi i \sin \omega T}\right)^{1/2} \exp\left\{ i \frac{m\omega}{2 \sin \omega T} \left(({q'}^2 + q^2) \cos \omega T - 2q'q \right) \right\}.$$

We can put $q' = q$ and $T = -i\beta$:

$$K(q, -i\beta; q, 0) = \left(\frac{m\omega}{2\pi \sinh(\beta\omega)}\right)^{1/2} \exp\left\{ -\frac{m\omega q^2}{\sinh(\beta\omega)} \left(\cosh(\beta\omega) - 1\right) \right\}.$$

The partition function is thus

$$
\begin{aligned}
Z &= \int dq\, K(q,-i\beta;q,0) = \left(\frac{m\omega}{2\pi\sinh(\beta\omega)}\right)^{1/2}\sqrt{\frac{\pi}{\frac{m\omega}{\sinh(\beta\omega)}\left(\cosh(\beta\omega)-1\right)}} \\
&= \left[2(\cosh(\beta\omega)-1)\right]^{-1/2} = \left[e^{\beta\omega/2}(1-e^{-\beta\omega})\right]^{-1} \\
&= \frac{e^{-\beta\omega/2}}{1-e^{-\beta\omega}} = \sum_{j=0}^{\infty} e^{-\beta(j+1/2)\omega}.
\end{aligned}
$$

Putting $\hbar$ back in, we get the familiar result

$$Z = \sum_{j=0}^{\infty} e^{-\beta(j+1/2)\hbar\omega}.$$

The previous calculation actually had nothing to do with PIs. The result for K was derived *via* PIs earlier, but it can be derived (more easily, in fact) in ordinary quantum mechaincs. However we can rewrite the partition function in terms of a PI. In ordinary (real) time,

$$K(q',T;q,0) = \int \mathcal{D}q(t)\exp i\int_0^T dt\left(\frac{m\dot{q}^2}{2} - V(q)\right),$$

where the integral is over all paths from $(q,0)$ to (q',T). With $q'=q$, $T\to -i\beta$,

$$K(q,-i\beta;q,0) = \int \mathcal{D}q(t)\exp i\int_0^{-i\beta} dt\left(\frac{m\dot{q}^2}{2} - V(q)\right).$$

where we now integrate along the negative imaginary time axis (Figure 12).

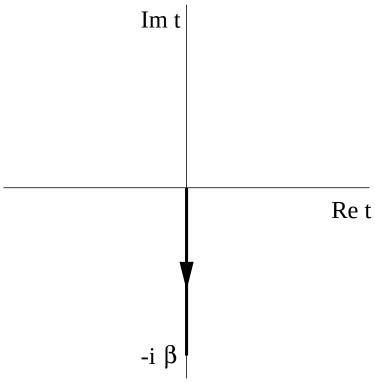

Figure 12: Path in the complex time plane.

Let us define a real variable for this integration, $\tau = it$. τ is called the imaginary time, since when the time t is imaginary, τ is real. (Kind of confusing, admittedly, but true.) Then the integral over τ is along *its* real axis: when $t: 0\to -i\beta$, then $\tau: 0\to\beta$. We can write q as a function of the variable τ: $q(t)\to q(\tau)$; then $\dot{q} = idq/d\tau$. The propagator becomes

$$K(q,-i\beta;q,0) = \int \mathcal{D}q(\tau)\exp - \int_0^{\beta} d\tau\left(\frac{m}{2}\left(\frac{dq}{d\tau}\right)^2 + V(q)\right). \tag{32}$$

The integral is over all functions $q(\tau)$ such that $q(0) = q(\beta) = q$.

The result (32) is an "imaginary-time" or "Euclidean" path integral, defined by associating to each path an amplitude (statistical weight) $\exp -S_E$, where S_E is the so-called Euclidean action, obtained from the usual ("Minkowski") action by changing the sign of the potential energy term.

The Euclidean PI might seem like a strange, unphysical beast, but it actually has many uses. One will be discussed in the next section, where use will be made of the fact that at low temperatures the ground state gives the dominant contribution to the partition function. It can therefore be used to find the ground state energy. We will also see the Euclidean PI in Section 9, when discussing the subject of instantons, which are used to describe phenomena such as quantum mechanical tunneling.

6 Perturbation Theory in Quantum Mechanics

We can use the Euclidean PI to compute a perturbation expansion for the ground state energy (among other things). This is not terribly useful in and of itself (once again, conventional techniques are a good deal easier), but the techniques used are very similar to those used in perturbation theory and Feynman diagrams in field theory. For this reason, we will discuss corrections to the ground state energy of an elementary quantum mechanical system in some detail.

From Z it is quite easy to extract the ground state energy. (This is a well-known fact of statistical mechanics, quite independent of PIs.) From the definition of Z,

$$Z(\beta) = \sum_j e^{-\beta E_j},$$

we can see that the contribution of each state decreases exponentially with β. However, that of the ground state decreases less slowly than any other state. So in the limit of large β (*i.e.*, low temperature), the ground state contribution will dominate. (This is mathematically straightforward, and also physically reasonable.) One finds

$$E_0 = -\lim_{\beta\to\infty} \frac{1}{\beta} \log Z. \tag{33}$$

In fact, we can extract E_0 from something slightly easier to calculate than Z. Rather than integrating over the initial (= final) position, as with Z, let us look at the Euclidean propagator from $q = 0$ to $q' = 0$ (the choice of zero is arbitrary).

$$K_E(0, \beta; 0, 0) = \langle q' = 0|\, e^{-\beta H}\, |q = 0\rangle\,.$$

We can insert a complete set of eigenstates of H:

$$\begin{aligned} K_E(0, \beta; 0, 0) &= \langle q' = 0|\, e^{-\beta H} \sum_j |j\rangle\, \langle j|\, q = 0\rangle \\ &= \sum_j e^{-\beta E_j} \phi_j(0)\phi_j^*(0), \end{aligned}$$

where ϕ_j are the wave functions of H. Again the ground state dominates as $\beta \to \infty$, and

$$E_0 = -\lim_{\beta\to\infty} \frac{1}{\beta} \log K_E(0, \beta; 0, 0). \tag{34}$$

(As $\beta \to \infty$, the difference between $\beta^{-1} \log Z$ and $\beta^{-1} \log K_E$ goes to zero.)

So let us see how we can calculate $K_E(0, \beta; 0, 0)$ perturbatively via the PI. The starting point is

$$K_E(0, \beta; 0, 0) = \int \mathcal{D}q\, e^{-S_E(\dot{q}, q)},$$

where the paths over which we integrate start and finish at $q = 0$, and where the Euclidean action is

$$S_E = \int_0^\beta d\tau \left(\frac{m\dot{q}^2}{2} + V(q) \right),$$

and with $\dot{q} = dq/d\tau$. As an example, consider the anharmonic oscillator, with quadratic and quartic terms in the potential:

$$K_E(0,\beta;0,0) = \int \mathcal{D}q \, \exp - \int d\tau \left(\frac{1}{2} m\dot{q}^2 + \frac{1}{2} m\omega^2 q^2 + \frac{\lambda}{4!} q^4 \right). \tag{35}$$

Clearly it is the quartic term which complicates life considerably; we cannot do the PI exactly.[7] But we can use the following trick to evaluate it perturbatively in λ. (This trick is far more complicated than necessary for this problem, but is a standard – and necessary! – trick in quantum field theory.) Define $K_E^0[J]$, the PI for a harmonic oscillator with a source term (which describes the action of an external force) added to the Lagrangian:

$$K_E^0[J] = \int \mathcal{D}q \, \exp - \int d\tau \left(\frac{1}{2} m\dot{q}^2 + \frac{1}{2} m\omega^2 q^2 - J(\tau) q(\tau) \right). \tag{36}$$

Unlike (35), this PI can be evaluated exactly; we will do this (as much as is necessary, at least) shortly. Once we have evaluated it, how does it help us to compute (35)? To see the use of $K_E^0[J]$, acting on it with a derivative has the effect of putting a factor q in the PI: for any time τ_1,

$$\frac{\delta K_E^0[J]}{\delta J(\tau_1)} = \int \mathcal{D}q \, q(\tau_1) \exp - \int d\tau \left(\frac{1}{2} m\dot{q}^2 + \frac{1}{2} m\omega^2 q^2 - J(\tau) q(\tau) \right).$$

A second derivative puts a second q in the PI:

$$\frac{\delta^2 K_E^0[J]}{\delta J(\tau_1) \delta J(\tau_2)} = \int \mathcal{D}q \, q(\tau_1) q(\tau_2) \exp - \int d\tau \left(\frac{1}{2} m\dot{q}^2 + \frac{1}{2} m\omega^2 q^2 - J(\tau) q(\tau) \right).$$

In fact, we can generalize this to an arbitrary functional F:

$$F\left[\frac{\delta}{\delta J}\right] K_E^0[J] = \int \mathcal{D}q \, F[q] e^{-S_E^0[J]}, \tag{37}$$

where $S_E^0[J]$ is the Euclidean action for the harmonic oscillator with source. (To prove (37), bring $F[\delta/\delta J]$ inside the PI; each $\delta/\delta J$ in F operating on $\exp -S_E^0[J]$ gives rise to a q in front of the exponential.)

Now, if we choose $F[q] = \exp - \int d\tau \frac{\lambda}{4!} q^4$, we get:

$$\begin{aligned} e^{-\int d\tau \frac{\lambda}{4!} \left(\frac{\delta}{\delta J}\right)^4} K_E^0[J] &= \int \mathcal{D}q \, \exp\left\{ - \int d\tau \frac{\lambda}{4!} q^4 \right\} e^{-S_E^0[J]} \\ &= \int \mathcal{D}q \, \exp - \int d\tau \left(\frac{1}{2} m\dot{q}^2 + \frac{1}{2} m\omega^2 q^2 + \frac{\lambda}{4!} q^4 - J(\tau) q(\tau) \right). \end{aligned}$$

[7] In fact, the situation is exactly like the evaluation of the ordinary integral

$$I = \int_{-\infty}^{\infty} dx \, \exp -\left(\frac{1}{2} x^2 + \frac{\lambda}{4!} x^4\right),$$

which looks innocent enough but which cannot be evaluated exactly. The technique which we will develop to evaluate (35) can also be used for this ordinary integral – an amusing and recommended exercise.

If we *now* put $J = 0$, we have the PI we started with. So the final result is:

$$\boxed{K_E(0,\beta;0,0) = \left(\exp\left\{-\int d\tau \frac{\lambda}{4!}\left(\frac{\delta}{\delta J}\right)^4\right\} K_E^0[J]\right)\Bigg|_{J=0}.} \tag{38}$$

We can, and will, calculate $K_E^0[J]$ as an explicit functional of J. If we then expand the exponential which operates on it in (38), we get a power series in λ:

$$\begin{aligned} K_E(0,\beta;0,0) &= \Bigg\{\Bigg(1 - \int d\tau \frac{\lambda}{4!}\left(\frac{\delta}{\delta J(\tau)}\right)^4 \\ &\qquad + \frac{1}{2!}\int d\tau \frac{\lambda}{4!}\left(\frac{\delta}{\delta J(\tau)}\right)^4 \int d\tau' \frac{\lambda}{4!}\left(\frac{\delta}{\delta J(\tau')}\right)^4 + \cdots\Bigg) K_E^0[J]\Bigg\}\Bigg|_{J=0} \\ &= K_E^0[J] - \frac{\lambda}{4!}\left(\int d\tau \left(\frac{\delta}{\delta J(\tau)}\right)^4 K_E^0[J]\right)\Bigg|_{J=0} + o(\lambda^2). \end{aligned}$$

Let us now evaluate $K_E^0[J]$,

$$K_E^0[J] = \int \mathcal{D}q \, \exp - \int d\tau \left(\frac{1}{2}m\dot{q}^2 + \frac{1}{2}m\omega^2 q^2 - J(\tau)q(\tau)\right).$$

To do this, suppose that we can find the *classical path* $q_{cJ}(\tau)$, the solution of

$$m\ddot{q} = m\omega^2 q - J(\tau), \qquad q(0) = q(\beta) = 0. \tag{39}$$

Once we have done this, we can perform a change of variables in the PI: we define $q(\tau) = q_{cJ}(\tau) + y(\tau)$, and integrate over paths $y(\tau)$. This is useful because

$$\begin{aligned} \int d\tau \left(\frac{1}{2}m\dot{q}^2 + \frac{1}{2}m\omega^2 q^2 - J(\tau)q(\tau)\right) &= \int d\tau \left(\frac{1}{2}m\dot{q}_{cJ}^2 + \frac{1}{2}m\omega^2 q_{cJ}^2 - J(\tau)q_{cJ}(\tau)\right) \\ &\quad + \underbrace{(\text{linear in } y)}_{=0} + \int d\tau \left(\frac{1}{2}m\dot{y}^2 + \frac{1}{2}m\omega^2 y^2\right). \end{aligned}$$

The linear term vanishes because q_{cJ} satisfies the equation of motion. So the PI becomes

$$K_E^0[J] = e^{-S_{Ec}[J]} \int \mathcal{D}y \, \exp - \int d\tau \left(\frac{1}{2}m\dot{y}^2 + \frac{1}{2}m\omega^2 y^2\right).$$

The crucial observation is that the resulting PI is *independent of* J: it is an irrelevant constant; call it C. (In fact, C is neither constant [it depends on β], nor entirely irrelevant [it is related to the unperturbed ground state energy, as we will see]. Crucial for the present purposes is that C is independent of J.)

$$K_E^0[J] = C\, e^{-S_{Ec}[J]},$$

where

$$S_{Ec}[J] = \int d\tau \left(\frac{1}{2}m\dot{q}_{cJ}^2 + \frac{1}{2}m\omega^2 q_{cJ}^2 - J(\tau)q_{cJ}(\tau)\right).$$

This can be simplified by integrating the first term by parts, yielding

$$S_{Ec}[J] = \int d\tau\, q_{cJ}\left(-\frac{1}{2}m\ddot{q}_{cJ} + \frac{1}{2}m\omega^2 q_{cJ} - J(\tau)\right).$$

Using the classical equation of motion (39), we get

$$S_{Ec}[J] = -\frac{1}{2}\int d\tau\, J(\tau)q_{cJ}(\tau).$$

We must still solve the classical problem, (39). The solution can be written in terms of the Green's function for the problem. Let $G(\tau,\tau')$ be the solution of

$$\begin{aligned} m\left(\frac{d^2}{d\tau^2} - \omega^2\right)G(\tau,\tau') &= \delta(\tau-\tau'),\\ G(0,\tau') = G(\beta,\tau') &= 0. \end{aligned}$$

Then we can immediately write

$$q_{cJ}(\tau) = \int_0^\beta d\tau' G(\tau,\tau')J(\tau'),$$

which can be proven by substution into (39). We can now write

$$\boxed{K_E^0[J] = C\,\exp\frac{1}{2}\int d\tau d\tau'\, J(\tau)G(\tau,\tau')J(\tau').} \tag{40}$$

We can find the Green's function easily in the limit $\beta\to\infty$. It is slightly more convenient to treat the initial and final times more symmetrically, so let us choose the time interval to be $(-\beta/2,+\beta/2)$; in the limit $\beta\to\infty$ we go from $-\infty$ to ∞. Then we have

$$m\left(\frac{d^2}{d\tau^2} - \omega^2\right)G(\tau,\tau') = \delta(\tau-\tau').$$

By taking the Fourier transform, we see that

$$G(\tau,\tau') = -\frac{1}{m}\int_{-\infty}^{\infty}\frac{dk}{2\pi}\frac{1}{(k^2+\omega^2)}e^{ik(\tau-\tau')}. \tag{41}$$

We can now compute the first-order correction to K_E (from which we get the first-order correction to the ground state energy). We have

$$K_E = K_E^0[0] - \frac{\lambda}{4!}\int d\tau\left(\frac{\delta}{\delta J(\tau)}\right)^4 K_E^0[J]\Bigg|_{J=0}. \tag{42}$$

Since in the second term we take four derivatives of $K_E^0[J]$ and then set $J=0$, only the piece of $K_E^0[J]$ which is quartic in J is relevant: fewer than four J's will be killed by the derivatives; more than four will be killed when setting $J = 0$.

$$\begin{aligned} K_E^0[J] &= C \exp \frac{1}{2} \int d\tau d\tau' \, J(\tau) G(\tau, \tau') J(\tau') \\ &= \text{irrelevant} + C \cdot \frac{1}{2} \left(\frac{1}{2} \int d\tau d\tau' \, J(\tau) G(\tau, \tau') J(\tau') \right)^2 \\ &= \frac{C}{8} \langle J_1 G_{12} J_2 \rangle \langle J_3 G_{34} J_4 \rangle, \end{aligned} \tag{43}$$

where we have used the compact notation $\langle J_1 G_{12} J_2 \rangle = \int d\tau_1 d\tau_2 \, J(\tau_1) G(\tau_1, \tau_2) J(\tau_2)$.

Substituting (43) into (42),

$$K_E = C \left(1 - \frac{\lambda}{4!} \frac{1}{8} \int d\tau \left(\frac{\delta}{\delta J(\tau)} \right)^4 \langle J_1 G_{12} J_2 \rangle \langle J_3 G_{34} J_4 \rangle + o(\lambda^2) \right). \tag{44}$$

To ensure that we understand the notation and how functional differentiation works, let us work out a slightly simpler example than the above. Consider

$$X \equiv \left(\frac{\delta}{\delta J(\tau)} \right)^2 \langle J_1 G_{12} J_2 \rangle = \left(\frac{\delta}{\delta J(\tau)} \right)^2 \int d\tau_1 d\tau_2 \, J(\tau_1) G(\tau_1, \tau_2) J(\tau_2).$$

The first derivative can act either on J_1 or J_2. In either case, it gives a delta function, which will make one of the integrals collapse:

$$\begin{aligned} X &= \frac{\delta}{\delta J(\tau)} \int d\tau_1 d\tau_2 \left(\delta(\tau - \tau_1) G(\tau_1, \tau_2) J(\tau_2) + J(\tau_1) G(\tau_1, \tau) \delta(\tau - \tau_2) \right) \\ &= \frac{\delta}{\delta J(\tau)} \left(\int d\tau_2 G(\tau, \tau_2) J(\tau_2) + \int d\tau_1 J(\tau_1) G(\tau_1, \tau) \right) \end{aligned}$$

In each term the remaining derivative acts similarly and kills the remaining integral; the result is

$$X = 2G(\tau, \tau).$$

The functional derivatives in (44) are a straightforward generalization of this; we find

$$K_E = C \left(1 - \frac{1}{8} \frac{\lambda}{4!} \int d\tau \, 4! G(\tau, \tau)^2 \right).$$

From (41) $G(\tau, \tau) = -1/2m\omega$; the τ integral is just the time interval β, and finally we get

$$K_E(0, \beta; 0, 0) = C \left(1 - \frac{\beta \lambda}{32 m^2 \omega^2} + o(\lambda^2) \right) = C e^{-\beta\lambda/32m^2\omega^2} \tag{45}$$

to order λ.

Now we can put this expression to good use, extracting the ground state energy from (34):

$$E_0 = -\lim_{\beta \to \infty} \frac{1}{\beta} \log K_E(0, \beta; 0, 0) = -\lim_{\beta \to \infty} \frac{1}{\beta} \left(\log C - \frac{\beta \lambda}{32 m^2 \omega^2} \right).$$

Recall that the constant C in (45) depends on β; this dependence must account for the ground state energy; the term linear in λ gives the first correction to the energy. Thus

$$E_0 = \frac{1}{2}\hbar\omega + \frac{\hbar^2\lambda}{32m^2\omega^2} + o(\lambda^2)$$

where we have reintroduced $\hbar$. We can check this result against standard perturbation theory (which is considerably easier!); the first-order correction to the ground state energy is

$$\int_{-\infty}^{\infty} dq\phi_0^*(q)\left(\frac{\lambda}{4!}q^4\right)\phi_0(q) = \cdots = \frac{\hbar^2\lambda}{32m^2\omega^2},$$

as above.

One's sanity would be called into question were it suggested that the PI calculation is a serious competitor for standard perturbation theory, although the latter itself gets rapidly more and more messy at higher orders. The technique above also gets messier, but it may well be that its messiness increases less quickly than that of standard perturbation theory. If so, the PI calculation could become competitive with standard perturbation theory at higher orders. But really the main motivation for discussing the above method is that it mimics in a more familiar setting standard perturbation techniques in quantum field theory.

To summarize this long and somewhat technical section, let us recall the main features of the above method. We express the ground state energy as an expression involving the propagator, (34). We separate the Lagrangian into a "free" (*i.e.*, quadratic) part and an "interacting" (beyond quadratic) part. Via the PI, we write the interacting propagator in terms of a free propagator with source term added, (38); this expression is amenable to a perturbation expansion. The free propagator can be evaluated explicitly, (40); then (38) can be computed to any desired order. From this we obtain directly the ground state energy to the same order.

7 Green's Functions in Quantum Mechanics

In quantum field theory we are interested in objects such as

$$\langle 0| T\hat{\phi}(x_1)\hat{\phi}(x_2)\cdots\hat{\phi}(x_n) |0\rangle \,,$$

the vacuum expectation value of a time-ordered product of Heisenberg field operators. This object is known as a Green's function or as a correlation function. The order of the operators is such that the earliest field is written last (right-most), the second earliest second last, *etc.* For example,

$$T\hat{\phi}(x_1)\hat{\phi}(x_2) = \begin{cases} \hat{\phi}(x_1)\hat{\phi}(x_2) & x_1^0 > x_2^0 \\ \hat{\phi}(x_2)\hat{\phi}(x_1) & x_2^0 > x_1^0 \end{cases}$$

Green's functions are related to amplitudes for physical processes such as scattering and decay processes. (This point is explained in most quantum field theory books.)

Let us look at the analogous object in quantum mechanics:

$$G^{(n)}(t_1, t_2, \cdots, t_n) = \langle 0| T\hat{q}(t_1)\hat{q}(t_2)\cdots\hat{q}(t_n) |0\rangle \,.$$

We will develop a PI expression for this.

First, we must recast the PI in terms of Heisenberg representation objects. The operator $\hat{q}(t)$ is the usual Heisenberg operator, defined in terms of the Schroedinger operator $\hat{q}$ by $\hat{q}(t) = e^{iHt}\hat{q}e^{-iHt}$. The eigenstates of the Heisenberg operator are $|q,t\rangle$: $\hat{q}(t)\,|q,t\rangle = q\,|q,t\rangle$. The relation with the time-independent eigenstates is $|q,t\rangle = e^{iHt}\,|q\rangle$.[8] Then we can write the PI:

$$K = \langle q'|\, e^{-iHT}\, |q\rangle = \langle q', T|\, q, 0\rangle = \int \mathcal{D}q\, e^{iS}.$$

We can now calculate the "2-point function" $G(t_1, t_2)$, via the PI. We will proceed in two steps. First, we will calculate the following expression:

$$\langle q', T|\, T\hat{q}(t_1)\hat{q}(t_2)\, |q, 0\rangle \,.$$

We will then devise a method for extracting the vacuum contribution to the initial and final states.

Suppose first that $t_1 > t_2$. Then

$$\begin{aligned} \langle q', T|\, T\hat{q}(t_1)\hat{q}(t_2)\, |q, 0\rangle &= \langle q', T|\, \hat{q}(t_1)\hat{q}(t_2)\, |q, 0\rangle \\ &= \int dq_1 dq_2 \langle q', T\, |q_1, t_1\rangle \underbrace{\langle q_1, t_1|\, \hat{q}(t_1)}_{\langle q_1, t_1| q_1} \underbrace{\hat{q}(t_2)\, |q_2, t_2\rangle}_{q_2 |q_2, t_2\rangle} \langle q_2, t_2|\, q, 0\rangle \\ &= \int dq_1 dq_2\, q_1 q_2 \langle q', T\, |q_1, t_1\rangle\, \langle q_1, t_1|\, q_2, t_2\rangle\, \langle q_2, t_2|\, q, 0\rangle. \end{aligned}$$

[8] There is a possible point of confusion here. We all know that Heisenberg states are independent of time, yet the eigenstates of $\hat{q}(t)$ depend on time. Perhaps the best way to view these states $|q,t\rangle$ is that they form, for any fixed time, a complete set of states. Just like the usual (time-independent) Heisenberg state $|q\rangle$ describes a particle which is localized at the point q *at time* $t = 0$, the state $|q,t\rangle$ describes a particle which is localized at the point q *at time* t.

Each of these matrix elements is a PI:

$$\langle q', T|\, T\hat{q}(t_1)\hat{q}(t_2)\, |q, 0\rangle = \int dq_1 dq_2\, q_1 q_2 \int_{q_1,t_1}^{q',T} \mathcal{D}q\, e^{iS} \int_{q_2,t_2}^{q_1,t_1} \mathcal{D}q\, e^{iS} \int_{q,0}^{q_2,t_2} \mathcal{D}q\, e^{iS}.$$

This expression consists of a first PI from the initial position q to an arbitrary position q_2, a second one from there to a second arbitrary position q_1, and a third one from there to the final position q'. So we are integrating over all paths from q to q', subject to the restriction that the paths pass through the intermediate points q_1 and q_2. We then integrate over the two arbitrary positions, so that in fact we are integrating over *all* paths: we can combine these three path integrals plus the integrations over q_1 and q_2 into one PI. The factors q_1 and q_2 in the above integral can be incorporated into this PI by simply including a factor $q(t_1)q(t_2)$ in the PI. So

$$\langle q', T|\, \hat{q}(t_1)\hat{q}(t_2)\, |q, 0\rangle = \int_{q,0}^{q',T} \mathcal{D}q\, q(t_1)q(t_2)e^{iS} \qquad (t_1 > t_2).$$

An identical calculation shows that exactly this same final expression is also valid for $t_2 < t_1$: magically, the PI does the time ordering automatically. Thus for all times

$$\langle q', T|\, T\hat{q}(t_1)\hat{q}(t_2)\, |q, 0\rangle = \int_{q,0}^{q',T} \mathcal{D}q\, q(t_1)q(t_2)e^{iS}.$$

As for how to obtain vacuum-to-vacuum matrix elements, our work on statistical mechanics provides us with a clue. We can expand the states $\langle q', T|$ and $|q, 0\rangle$ in terms of eigenstates of the Hamiltonian. If we evolve towards a *negative imaginary time*, the contribution of all other states will decay away relative to that of the ground state. We have (resetting the initial time to $-T$ for convenience)

$$\langle q', T|\, q, -T\rangle \propto \langle 0, T|\, 0, -T\rangle,$$

where on the right the "0" denotes the ground state. The proportionality involves the ground state wave function and an exponential factor $\exp 2iE_0T = \exp -2E_0|T|$.

We could perform all calculations in a Euclidean theory and analytically continue to real time when computing physical quantities (many books do this), but to be closer to physics we can also consider T not to be pure imaginary and negative, but to have a small negative imaginary phase: $T = |T|e^{-i\epsilon}$ $(\epsilon > 0)$. *In what follows, I will simply write T, but please keep in mind that it has a negative imaginary part!* With this,

$$\langle 0, T|\, 0, -T\rangle \propto \langle q', T|\, q, -T\rangle = \int \mathcal{D}q\, e^{iS}.$$

To compute the Green's functions, we must simply add $T\hat{q}(t_1)\hat{q}(t_2)\cdots\hat{q}(t_n)$ to the matrix element, and the corresponding factor $q(t_1)q(t_2)\cdots q(t_n)$ inside the PI:

$$\langle 0, T|\, T\hat{q}(t_1)\hat{q}(t_2)\cdots\hat{q}(t_n)\, |0, -T\rangle \propto \int \mathcal{D}q\, q(t_1)q(t_2)\cdots q(t_n)e^{iS}.$$

The proportionality sign is a bit awkward; fortunately, we can rid ourselves of it. To do this, we note that the left hand expression is not exactly what we want: the vacua $|0,\pm T\rangle$ differ by a phase. We wish to eliminate this phase; to this end, the Green's functions are defined

$$\begin{aligned} G^{(n)}(t_1,t_2,\cdots,t_n) &= \langle 0|\, T\hat{q}(t_1)\hat{q}(t_2)\cdots\hat{q}(t_n)\,|0\rangle \\ &\equiv \frac{\langle 0,T|\, T\hat{q}(t_1)\hat{q}(t_2)\cdots\hat{q}(t_n)\,|0,-T\rangle}{\langle 0,T|\,0,-T\rangle} \\ &= \frac{\int \mathcal{D}q\, q(t_1)q(t_2)\cdots q(t_n)e^{iS}}{\int \mathcal{D}q\, e^{iS}}, \end{aligned}$$

with no proportionality sign. The wave functions and exponential factors in the numerator and denominator cancel.

To compute the numerator, we can once again use the trick we used in perturbation theory in quantum mechanics, namely, adding a source to the action. We define

$$Z[J] = \frac{\int \mathcal{D}q\, e^{i(S+\int dt\, J(t)q(t))}}{\int \mathcal{D}q\, e^{iS}} = \frac{\langle 0|\,0\rangle_J}{\langle 0|\,0\rangle_{J=0}}.$$

If we operate on $Z[J]$ with $i^{-1}\delta/\delta J(t_1)$, this gives

$$\begin{aligned} \left.\left(\frac{1}{i}\frac{\delta}{\delta J(t_1)}Z[J]\right)\right|_{J=0} &= \left.\left(\frac{\int \mathcal{D}q\, q(t_1)e^{i(S+\int dt\, J(t)q(t))}}{\int \mathcal{D}q\, e^{iS}}\right)\right|_{J=0} \\ &= \frac{\int \mathcal{D}q\, q(t_1)e^{iS}}{\int \mathcal{D}q\, e^{iS}} \\ &= \frac{\langle 0,T|\,\hat{q}(t_1)\,|0,-T\rangle}{\langle 0,T|\,0,-T\rangle} = \langle 0|\,\hat{q}(t_1)\,|0\rangle \end{aligned}$$

(The expectation values are evaluated in the *absence* of J.)

Repeating this procedure, we obtain a PI with several q's in the numerator. This ordinary product of q's in the PI corresponds, as discussed earlier in this section, to a time-ordered product in the matrix element. So we make the following conclusion:

$$\left.\left(\frac{1}{i}\frac{\delta}{\delta J(t_1)}\cdots\frac{1}{i}\frac{\delta}{\delta J(t_n)}Z[J]\right)\right|_{J=0} = \frac{\int \mathcal{D}q\, q(t_1)\cdots q(t_n)e^{iS}}{\int \mathcal{D}q\, e^{iS}} = \langle 0|\, T\hat{q}(t_1)\cdots\hat{q}(t_1)\,|0\rangle\,.$$

For obvious reasons, the functional $Z[J]$ is called the *generating functional* for Green's functions; it is a very handy tool in quantum field theory and in statistical mechanics.

How do we calculate $Z[J]$? Let us examine the numerator:

$$N \equiv \int \mathcal{D}q\, e^{i(S+\int dt\, J(t)q(t))}.$$

Suppose initially that S is the harmonic oscillator action (denoted S_0):

$$S_0 = \int dt\, \left(\frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2\right),$$

Then the corresponding numerator, N_0, is the non-Euclidean (*i.e.*, real-time) version of the propagator $K_E^0[J]$ we used in Section 6. We can calculate $N_0[J]$ in the same way as $K_E^0[J]$. Since the calculation repeats much of that of $K_E^0[J]$, we will be succinct.

By definition,

$$N_0 = \int \mathcal{D}q(t) \exp i \int dt \left(\frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2 + Jq\right).$$

We do the path integral over a new variable y, defined by $q(t) = q_c(t) + y(t)$, where q_c is the classical solution. Then the PI over y is a constant (independent of J) and we can avoid calculating it. (It will cancel against the denominator in $Z[J]$.) Calling it C, we have

$$N_0 = Ce^{iS_{0J}[q_c]},$$

where

$$S_{0J}[q_c] = \int dt \left(\frac{1}{2}m\dot{q}_c^2 - \frac{1}{2}m\omega^2 q_c^2 + Jq_c\right) = \frac{1}{2}\int dt J(t) q_c(t),$$

using the fact that q_c satisfies the equation of motion. We can write the classical path in terms of the Green's function (to be determined shortly), defined by

$$\left(\frac{d^2}{dt^2} + \omega^2\right) G(t,t') = -i\delta(t-t'). \tag{46}$$

Then

$$q_c(t) = -i \int dt' G(t,t') J(t').$$

We can now write

$$N_0 = C \exp \frac{1}{2} \int dt dt' \, J(t) G(t,t') J(t').$$

Dividing by the denominator merely cancels the factor C, giving our final result:

$$\boxed{Z[J] = \exp \frac{1}{2} \int dt dt' \, J(t) G(t,t') J(t').}$$

We can solve (46) for the Green's function by going into momentum space; the result is

$$G(t,t') = G(t-t') = \int \frac{dk}{2\pi} \frac{i}{k^2 - \omega^2} e^{-ik(t-t')}.$$

However, there are poles on the axis of integration. (This problem did not arise in Euclidean space; see (41).) The Green's function is ambiguous until we give it a "pole prescription", *i.e.*, a boundary condition. But remember that our time T has a small, negative imaginary part. We require that G go to zero as $T \to \infty$. The correct pole prescription then turns out to be

$$\boxed{G(t-t') = \int \frac{dk}{2\pi} \frac{i}{k^2 - \omega^2 + i\epsilon} e^{-ik(t-t')}.} \tag{47}$$

We could at this point do a couple of practice calculations to get used to this formalism. Examples would be to compute perturbatively the generating functional for an action which has terms beyond quadratic (for example, a q^4 term), or to compute some Green's function in either the quadratic or quartic theory. But since these objects aren't really useful in quantum mechanics, without further delay we will go directly to the case of interest: quantum field theory.

8 Green's Functions in Quantum Field Theory

It is easy to generalize the PI to many degrees of freedom; we have in fact already done so in Section 4, where particles move in two or three dimensions. It is simply a matter of adding a new index to denote the different degrees of freedom (be they the different coordinates of a single particle in more than one dimension or the particle index for a system of many particles).

One of the most important examples of a system with many degrees of freedom is a field theory: $q(t) \to \phi(\mathbf{x}, t) = \phi(x)$. Not only is this a system of *many* degrees of freedom, but one of a continuum of degrees of freedom. The passage from a discrete to continuous system in path integrals can be done in the same way as in ordinary classical field theory: we can discretize the field (modeling it by a set of masses and springs, for instance), do the usual path integral manipulations on the discrete system, and take the continuum limit at the end of the calculation. The final result is a fairly obvious generalization of the one-particle results, so I will not dwell on the mundane details of discretization and subsequent taking of the continuum limit.

The analog of the quantum mechanical propagator is the transition amplitude to go from one field configuration $\phi(\mathbf{x})$ at $t = 0$ to another $\phi'(\mathbf{x}')$ at $t = T$:

$$K(\phi'(\mathbf{x}'), T; \phi(\mathbf{x}), 0) = \int \mathcal{D}\phi e^{iS[\phi]}, \tag{48}$$

where S is the field action, for instance

$$S[\phi] = \int d^4x \left(\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2 \right) \tag{49}$$

for the free scalar field. In (48) the integral is over all field configurations $\phi(x)$ obeying the stated initial and final conditions.

In field theory, we are not really interested in this object. Rather (as mentioned earlier), we are interested in Green's functions. Most of the work required to translate (48) into an expression for a Green's function (generating functional of Green's functions, more precisely) has already been done in the last section, so let us study a couple of cases.

8.1 Free scalar field.

For the free scalar field, whose action is given by (49), the generating functional is

$$Z_0[J] = \frac{\langle 0|\, 0\rangle_J}{\langle 0|\, 0\rangle_{J=0}}.$$

Both numerator and denominator can be written in terms of PIs. The numerator is

$$N_0 = \int \mathcal{D}\phi \, \exp i \int d^4x \left(\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2} m^2 \phi^2 + J\phi \right).$$

We write $\phi = \phi_c + \varphi$, where ϕ_c is the classical field configuration, and integrate over the deviation from ϕ_c. The action can be written

$$S[\phi_c + \varphi] = \int d^4x \left(\frac{1}{2}(\partial_\mu \phi_c)^2 - \frac{1}{2} m^2 \phi_c^2 + J\phi_c \right) + \int d^4x \left(\frac{1}{2}(\partial_\mu \varphi)^2 - \frac{1}{2} m^2 \varphi^2 \right),$$

where as usual there is no term linear in φ since ϕ_c by definition extremizes the classical action. So

$$N_0 = C \exp i \int d^4x \left(\frac{1}{2}(\partial_\mu \phi_c)^2 - \frac{1}{2}m^2\phi_c^2 + J\phi_c\right),$$

where

$$C = \int \mathcal{D}\varphi \exp i \int d^4x \left(\frac{1}{2}(\partial_\mu \varphi)^2 - \frac{1}{2}m^2\varphi^2\right).$$

C is independent of J and will cancel in Z. (Indeed, the denominator is *equal to* C.)

Using the fact that ϕ_c obeys the classical equation

$$(\partial^2 + m^2)\phi_c = J,$$

we can write

$$N_0 = C \exp \frac{i}{2} \int d^4x\, J(x)\phi_c(x).$$

Finally, we can write ϕ_c in terms of the Klein-Gordon Green's function, defined by

$$(\partial^2 + m^2)\Delta_F(x, x') = -i\delta^4(x - x').$$

It is

$$\phi_c(x) = i \int d^4x\, \Delta_F(x, x')J(x'),$$

so

$$\boxed{Z_0 = \frac{N_0}{C} = \exp -\frac{1}{2} \int d^4x d^4x'\, J(x)\Delta_F(x, x')J(x').}$$

The Green's function is found by solving its equation in 4-momentum space; the result is

$$\Delta_F(x, x') = \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} e^{-ik\cdot(x-x')} = \Delta_F(x - x'),$$

adopting the same pole prescription as in (47). Note that Δ_F is an even function, $\Delta_F(x - x') = \Delta_F(x' - x)$.

Let us calculate a couple of Green's functions. These calculations are reminiscent of those at the end of Section 6. As a first example, consider

$$G_0^{(2)}(x_1, x_2) = \langle 0|\, T\hat{\phi}(x_1)\hat{\phi}(x_2)\, |0\rangle = \frac{1}{i^2}\left(\frac{\delta^2}{\delta J(x_1)\delta J(x_2)} Z_0[J]\right)\Bigg|_{J=0}.$$

Expanding Z_0 in powers of J,

$$Z_0[J] = 1 - \frac{1}{2} \int d^4x d^4x'\, J(x)\Delta_F(x - x')J(x') + o(J^4).$$

The term *quadratic* in J is the only one that survives both differentiation (which kills the "1") and the setting of J to zero (which kills all higher-order terms). So

$$G_0^{(2)}(x_1, x_2) = \frac{1}{i^2}\frac{\delta^2}{\delta J(x_1)\delta J(x_2)}\left(-\frac{1}{2}\int d^4x d^4x'\, J(x)\Delta_F(x - x')J(x')\right).$$

There arise two identical terms, depending on which derivative acts on which J. The result is

$$G_0^{(2)}(x_1, x_2) = \Delta_F(x_1 - x_2).$$

So the Green's function (or two-point function) in the quantum field theory sense is also the Green's function in the usual differential-equations sense.

As a second example, the four-point Green's function is

$$G_0^{(4)}(x_1, x_2, x_3, x_4) = \frac{1}{i^4}\left(\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_4)} \exp -\frac{1}{2}\int d^4x d^4x' \, J(x)\Delta_F(x, x')J(x')\right)\Bigg|_{J=0}.$$

This time the only part of the exponential that contributes is the term with four J's.

$$G_0^{(4)}(x_1, x_2, x_3, x_4) = \frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_4)} \frac{1}{2}\left(-\frac{1}{2}\int d^4x d^4x' \, J(x)\Delta_F(x, x')J(x')\right)^2.$$

There are $4! = 24$ terms, corresponding to the number of ways of associating the derivatives with the J's. In 8 of them, the Green's functions which arise are $\Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4)$, and so on. The result is

$$\begin{aligned} G_0^{(4)}(x_1, x_2, x_3, x_4) &= \Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4) + \Delta_F(x_1 - x_3)\Delta_F(x_2 - x_4) \\ &\quad + \Delta_F(x_1 - x_4)\Delta_F(x_2 - x_3), \end{aligned} \tag{50}$$

which can be represented diagramatically as in Figure 13.

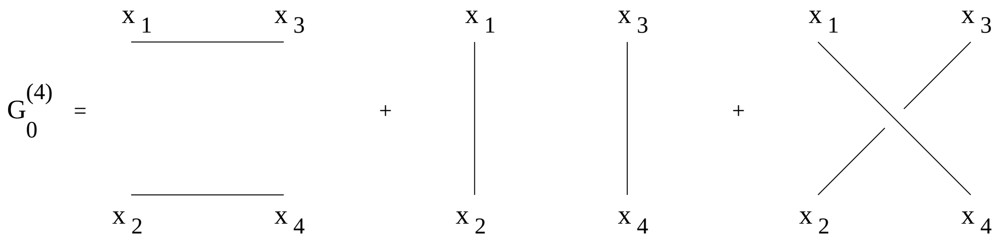

Figure 13: Diagrammatic representation of (50). Each line counts as a factor Δ_F with argument corresponding to the endpoints of the line.

8.2 Interacting scalar field theory.

Usually, if the Lagrangian has a term beyond quadratic we can no longer evaluate exactly the functional integral, and we must resort to perturbation theory. The generating functional method is tailor-made to do this in a systematic fashion. To be specific, consider ϕ^4 theory, defined by the Lagrangian density

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4.$$

Then the generating functional is (up to an unimportant constant: we will normalize ultimately so that $Z[J = 0] = 1$)

$$Z[J] = C\int \mathcal{D}\phi \, \exp i\int d^4x \left(\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 + J\phi\right).$$

Because of the quartic term, we cannot evaluate the functional integral exactly. But we can use a trick we first saw when discussing perturbation theory in quantum mechanics: replacing the higher-order term by a functional derivative with respect to J:

$$\begin{aligned} Z[J] &= C\int \mathcal{D}\phi \exp\left\{-i\frac{\lambda}{4!}\int d^4x\,\phi^4\right\}\exp i\int d^4x\left(\frac{1}{2}(\partial_\mu\phi)^2-\frac{1}{2}m^2\phi^2+J\phi\right) \\ &= C\int \mathcal{D}\phi \exp\left\{-i\frac{\lambda}{4!}\int d^4x\,\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4\right\}\exp i\int d^4x\left(\frac{1}{2}(\partial_\mu\phi)^2-\frac{1}{2}m^2\phi^2+J\phi\right). \end{aligned}$$

We can pull the first exponential out of the integral; the functional integral which remains is that for Z_0. Adjusting the constant C so that $Z[J=0]=0$, we get

$$Z[J]=\frac{\exp\left\{-i\frac{\lambda}{4!}\int d^4x\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4\right\}\exp -\frac{1}{2}\int d^4xd^4x'\,J(x)\Delta_F(x,x')J(x')}{\left(\exp\left\{-i\frac{\lambda}{4!}\int d^4x\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4\right\}\exp -\frac{1}{2}\int d^4xd^4x'\,J(x)\Delta_F(x,x')J(x')\right)\Big|_{J=0}}.$$

This expression now enables us to compute a perturbative expansion for any Green's function we desire. This is a rather mechanical job, and the only way to learn it is by doing lots of examples. To illustrate the method, let us look at $G^{(2)}(x_1,x_2)$ to the first nontrivial order in λ.

We have

$$G^{(2)}(x_1,x_2)=\frac{\left(\frac{1}{i^2}\frac{\delta^2}{\delta J(x_1)\delta J(x_2)}\exp\left\{-i\frac{\lambda}{4!}\int d^4x\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4\right\}\exp -\frac{1}{2}\langle J_a\Delta_{Fab}J_b\rangle\right)\Big|_{J=0}}{\left(\exp\left\{-i\frac{\lambda}{4!}\int d^4x\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4\right\}\exp -\frac{1}{2}\langle J_a\Delta_{Fab}J_b\rangle\right)\Big|_{J=0}},$$

where as in Section 6 $\langle\cdots\rangle$ implies integration over the positions of the J's. In both numerator and denominator, we can expand both exponentials. The only terms that survive are those that have the same total number of derivatives and J's. Let us look at the term linear in λ in the numerator. There are six derivatives, so we need the term from the expansion of the second exponential with six J's. For this term, we get the following expression for the numerator:

$$-\frac{\delta^2}{\delta J(x_1)\delta J(x_2)}\left(-i\frac{\lambda}{4!}\right)\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\frac{1}{3!}\left(-\frac{1}{2}\right)^3\langle J_a\Delta_{Fab}J_b\rangle\langle J_c\Delta_{Fcd}J_d\rangle\langle J_e\Delta_{Fef}J_f\rangle.$$

There are now a total of $6!=720$ terms! However, only two distinct analytical expressions result. The first of these arises if the derivatives at x_1 and x_2 act on different $\langle\cdots\rangle$'s. A little combinatorial head-scratching tells us that there are 576 such terms, yielding the following expression:

$$-\frac{i\lambda}{2}\int d^4x\,\Delta_F(x_1-x)\Delta_F(x-x)\Delta_F(x-x_2), \tag{51}$$

which can be represented pictorially as in Figure 14 (a).

The only other expression arises when the derivatives at x_1 and x_2 act on the same $\langle\cdots\rangle$. This accounts for the remaining 144 terms; the analytic form which results is

$$-\frac{i\lambda}{8}\Delta_F(x_1-x_2)\int d^4x\,\Delta_F(x-x)^2, \tag{52}$$

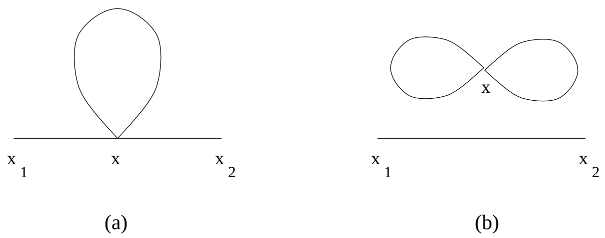

Figure 14: Diagrammatic representation of (a) (51) and (b) (52). Intersection points represent vertices, and count as a factor $-i\lambda \int d^4x$. Symmetry factors complete the association of an analytic expression with each diagram.

corresponding to the diagram in Figure 14(b).

The denominator can be evaluated in a similar fashion; the Green's function to order λ is

$$G^{(2)}(x_1, x_2) = \frac{\begin{array}{c}\left\{\Delta_F(x_1 - x_2) - \frac{i\lambda}{2}\int d^4x\, \Delta_F(x_1 - x)\Delta_F(x - x)\Delta_F(x - x_2)\right. \\ \left. -\frac{i\lambda}{8}\Delta_F(x_1 - x_2)\int d^4x\, \Delta_F(x - x)^2 + o(\lambda^2)\right\}\end{array}}{1 - \frac{i\lambda}{8}\int d^4x\, \Delta_F(x - x)^2 + o(\lambda^2)}.$$

Since we have only computed the numerator and denominator to order λ, we can rewrite this expression in the following way:

$$G^{(2)}(x_1, x_2) = \frac{\left(\begin{array}{c}\left\{\Delta_F(x_1 - x_2) - \frac{i\lambda}{2}\int d^4x\, \Delta_F(x_1 - x)\Delta_F(x - x)\Delta_F(x - x_2) + o(\lambda^2)\right\} \\ \times\left\{1 - \frac{i\lambda}{8}\int d^4x\, \Delta_F(x - x)^2 + o(\lambda^2)\right\}\end{array}\right)}{1 - \frac{i\lambda}{8}\int d^4x\, \Delta_F(x - x)^2 + o(\lambda^2)}.$$

We can now cancel the second factor in the numerator against the denominator, *to order* λ, resulting in

$$G^{(2)}(x_1, x_2) = \Delta_F(x_1 - x_2) - \frac{i\lambda}{2}\int d^4x\, \Delta_F(x_1 - x)\Delta_F(x - x)\Delta_F(x - x_2) + o(\lambda^2).$$

This factorization of the numerator into a part containing no factors independent of the external position times the denominator occurs to all orders, as can be proven fairly cleanly via a combinatoric argument. The conclusion is that so-called disconnected parts (parts of diagrams not connected to any external line) cancel from Green's functions, a fact which simplifies greatly the calculation of these objects.

It cannot be overemphasized that there are only three ways to get accustomed to this formalism: practice, practice, and practice. Other reasonable exercises are the calculation of $G^{(2)}$ to order λ^2 and the calculation of $G^{(4)}$ to order λ^2. ϕ^3 theory is also a useful testing ground for the techniques discussed in this section.

9 Instantons in Quantum Mechanics

9.1 General discussion

It has already been briefly mentioned that in quantum mechanics certain aspects of a problem can be overlooked in a perturbative treatment. One example occurs if we have a harmonic oscillator with a cubic anharmonic term: $V(q) = \frac{1}{2}m\omega^2 q^2 + \lambda q^3$ (Figure 15).

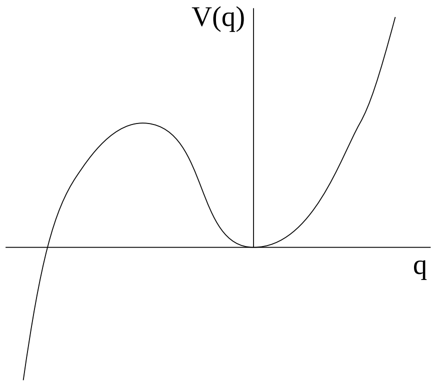

Figure 15: $V(q) = \frac{1}{2}m\omega^2 q^2 + \lambda q^3$.

We can calculate corrections to harmonic oscillator wave functions and energies perturbatively in λ, to any desired order, blissfully ignorant of a serious pathology in the model. As can be seen from Figure 15, this model has no ground state: the potential energy is unbounded as $q \to -\infty$, a point completely invisible to perturbation theory.

A second example is the double-well potential, $V(q) = \frac{\lambda}{4!}(q^2 - a^2)^2$ (Figure 16). There

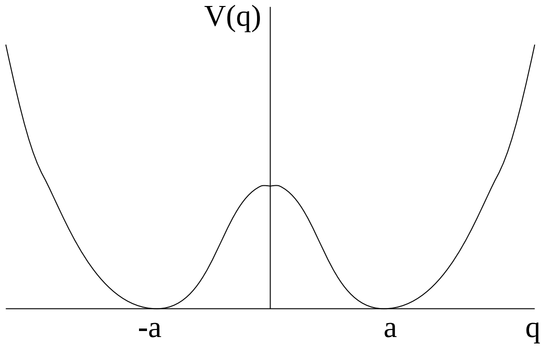

Figure 16: $V(q) = \frac{\lambda}{4!}(q^2 - a^2)^2$.

are two classical ground states. We can ignore this fact and expand V about one of the minima; it then takes the form of a harmonic oscillator about that minimum plus anharmonic terms (both cubic and quartic). We can then compute perturbative corrections to the wave functions and energies, and never see any evidence of the other minimum. Were we to expand about the other minimum, we would produce an identical set of perturbative corrections. By symmetry the ground state energies calculated perturbatively to *any order* will be the same for the expansions about the two minima, so it appears that we have degenerate ground states. But in fact the ground state is not degenerate: a nonperturbative energy splitting separates the true ground state (an even function of q) from the first excited state (an odd function); this splitting is not seen in perturbation theory.

We will examine this second example using PIs, the main goal being to calculate the energy splitting between the two candidate ground states.

Let us first recall the PI expression for the Euclidean propagator:

$$K_E(q', \frac{\beta}{2}; q, -\frac{\beta}{2}) = \langle q'|\, e^{-\beta H/\hbar}\, |q\rangle = \int \mathcal{D}q\, e^{-S_E/\hbar},$$

where

$$S_E = \int_{-\beta/2}^{\beta/2} d\tau \left(\frac{1}{2}m\dot{q}^2 + V(q)\right).$$

Henceforth, we will set $m \to 1$. K_E is useful because we can write it as

$$K_E = \sum_n \langle q'|\, n\rangle\langle n\, |q\rangle\, e^{-\beta E_n/\hbar}; \tag{53}$$

in the limit $\beta \to \infty$, this term will be dominated by the lowest-energy states. I say "states" here rather than "state" because we must calculate the two lowest-energy eigenvalues to get the splitting of the (perturbatively degenerate) lowest-energy states in the double-well potential.

We will evaluate the PI using an approximation known as the semiclassical approximation, or alternatively as the method of steepest descent. To illustrate it, consider the following integral

$$I = \int_{-\infty}^{\infty} dx\, e^{-S(x)/\hbar},$$

where $S(x)$ is a function with several local minima (Figure 17).

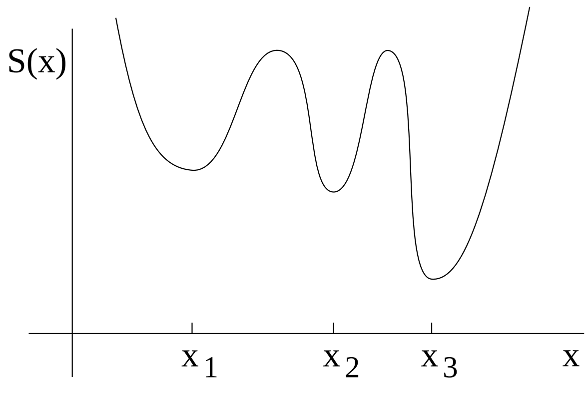

Figure 17: Potential with several minima.

Suppose we are interested in this integral as $\hbar \to 0$. Then the integral will be dominated by the minima of S; we can approximate it by a series of Gaussian integrations, one for each minimum of S. If x_i is such a minimum, then in its vicinity $S(x) \simeq S(x_i) + \frac{1}{2}(x - x_i)^2 S''(x_i)$; we can write

$$I \simeq I_1 + I_2 + I_3 + \cdots, \tag{54}$$

where

$$\begin{aligned} I_i &= \int_{-\infty}^{\infty} dx \, \exp -[S(x_i) + \frac{1}{2}(x - x_i)^2 S''(x_i)]/\hbar \\ &= e^{-S(x_i)/\hbar} \sqrt{\frac{2\pi\hbar}{S''(x_i)}}. \end{aligned}$$

Anharmonicities of S appear as corrections of order $\hbar$ to I. (This can be easily seen, for example, by considering a specific case such as $S(x) = ax^2 + bx^4$.)

We will compute the PI (53) in the semi-classical approximation, where the analog of the x_i in the above example will be classical paths (extremum of the action $S_E[q]$).

Suppose, then, that $q_c(\tau)$ is the classical solution to the problem

$$\frac{d^2}{d\tau^2} q = \frac{\partial V(q)}{\partial q}, \qquad q(-\beta/2) = q, \qquad q(\beta/2) = q'.$$

We can write $q(\tau) = q_c(\tau) + y(\tau)$; the action is

$$\begin{aligned} S_E[q_c + y] &= \int_{-\beta/2}^{\beta/2} d\tau \left(\frac{1}{2}\dot{q}_c + \dot{y}^2 + V(q_c + y) \right) \\ &= \int_{-\beta/2}^{\beta/2} d\tau \left(\frac{1}{2}\dot{q}_c^2 + V(q_c) \right) + (\text{linear in } y) \\ &\qquad + \int_{-\beta/2}^{\beta/2} d\tau \left(\frac{1}{2}\dot{y}^2 + \frac{1}{2}V''(q_c)y^2 \right) + \cdots . \end{aligned} \tag{55}$$

The term linear in y vanishes for the usual reason, and the higher order terms not written down are of cubic or higher order in y. Neglecting these (which give order $\hbar$ corrections to the PI), the propagator becomes

$$K_E = \int \mathcal{D}q \, e^{-S_E/\hbar} = e^{-S_E[q_c]/\hbar} \int \mathcal{D}y \, \exp - \int d\tau \left(\frac{1}{2}\dot{y}^2 + \frac{1}{2}V''(q_c)y^2 \right) /\hbar.$$

The functions $y(\tau)$ over which we integrate satisfy the boundary conditions $y(-\beta/2) = y(\beta/2) = 0$. The PI, being Gaussian, can be done exactly; it is not as straightforward as the harmonic oscillator PI since $V''(q_c)$ depends on τ. While we have often managed to *avoid* evaluating PIs, here we must evaluate it. (Unfortunately, this is rather difficult.)

To this end, we can use a generalization of the Fourier expansion technique mentioned in Section 2.2.2. We can rewrite the action as

$$S_E = \int d\tau \left(\frac{1}{2}\dot{y}^2 + \frac{1}{2}V''(q_c)y^2 \right) = \frac{1}{2} \int d\tau \, y \left(-\frac{d^2}{d\tau^2} + V''(q_c) \right) y. \tag{56}$$

The Schroedinger-like equation

$$\left(-\frac{d^2}{d\tau^2} + V''(q_c) \right) y = \lambda y, \qquad y(-\beta/2) = y(\beta/2) = 0$$

has a complete, orthonormal set of solutions; let the solutions and eigenvalues be $y_k(\tau)$ and λ_k, respectively. The orthonormality relation is

$$\int_{-\beta/2}^{\beta/2} d\tau y_k(\tau) y_l(\tau) = \delta_{kl}.$$

Then we can substitute $y(\tau) = \sum_k a_k y_k(\tau)$ in (56), giving

$$S_E = \frac{1}{2}\int d\tau \sum_k a_k y_k \left(-\frac{d^2}{d\tau^2} + V''(q_c)\right) \sum_l a_l y_l = \frac{1}{2}\sum_{k,l} a_k a_l \lambda_l \int d\tau y_k y_l = \frac{1}{2}\sum_k a_k^2 \lambda_k.$$

The PI can now be written as an integral over all possible values of the coefficients $\{a_k\}$. This gives

$$K_E = J' \int \prod_k da_k \, e^{-\sum_k a_k^2 \lambda_k / 2\hbar}, \tag{57}$$

where J' is the Jacobian of the transformation from $y(\tau)$ to $\{a_k\}$. (57) is a product of uncoupled Gaussian integrals; the result is

$$K_E = J' \prod_k \left(\frac{2\pi\hbar}{\lambda_k}\right)^{1/2} = J' \prod_k (2\pi\hbar)^{1/2} (\prod_k \lambda_k)^{-1/2} = J' \prod_k (2\pi\hbar)^{1/2} \mathrm{det}^{-1/2}\left(-\frac{d^2}{d\tau^2} + V''(q_c)\right),$$

where we have written the product of eigenvalues as the determinant of the Schroedinger operator on the space of functions vanishing at $\pm\beta/2$.

We can write $J = J' \prod_k (2\pi\hbar)^{1/2}$, giving

$$K_E = J \mathrm{det}^{-1/2}\left(-\frac{d^2}{d\tau^2} + V''(q_c)\right)(1 + o(\hbar)),$$

where the $o(\hbar)$ corrections can in principle be computed from the neglected beyond-quadratic terms in (55). We will not be concerned with these corrections, and henceforth we will drop the $(1 + o(\hbar))$.

9.2 Single Well in the Semiclassical Approximation

Before looking at the double well, it is worthwhile examining the single well, defined be

$$V(q) = \frac{1}{2}\omega^2 q^2 + \frac{\lambda}{4!} q^4.$$

The classical equation is

$$\frac{d^2}{d\tau^2} q = V'(q).$$

Note that this is the equation of motion for a particle moving in a potential $-V(q)$. If we choose the initial and final points $q = q' = 0$, then the classical solution is simply $q_c(\tau) = 0$; furthermore, $V''(q_c) = V''(0) = \omega^2$, and

$$K_E = J \mathrm{det}^{-1/2}\left(-\frac{d^2}{d\tau^2} + \omega^2\right).$$

The evaluation of the determinant is not terribly difficult (the eigenvalues can be easily found; their product can be found in a table of mathematical identities); the result, for large β, is

$$K_E = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\beta\omega/2}.$$

From (53), we can extract the ground state energy since, for large β, $K_E \sim \exp -E_0\beta/\hbar$. We find $E_0 = \hbar\omega/2$ up to corrections of order $\lambda\hbar^2$. We have discovered an incredibly complicated way of calculating the ground state energy of the harmonic oscillator!

9.3 Instantons in the Double Well Potential

Let us now study a problem of much greater interest: the double well. We will see that configurations known as "instantons" make a non-perturbative correction to the energies. We wish to evaluate the PI

$$K_E = \int_{q,-\beta/2}^{q',\beta/2} \mathcal{D}q \, e^{-S_E},$$

where

$$S_E = \int d\tau \left(\frac{1}{2}\dot{q}^2 + \frac{\lambda}{4!}(q^2 - a^2)^2\right),$$

for $\beta \to \infty$. As explained above, the PI is dominated by minima of S_E, *i.e.*, by classical solutions. The classical equation corresponds to a particle moving in the potential $-V(q)$ (Figure 18); the "energy" $E = \frac{1}{2}\dot{q}^2 - V(q)$ is conserved.

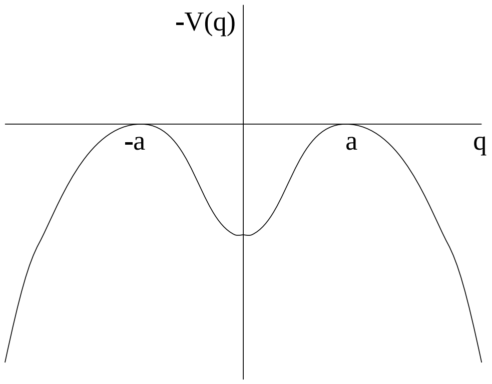

Figure 18: Inverted double-well potential.

Let us examine classical solutions, taking the boundary values q, q' of the classical solution corresponding to the maxima of $-V$, $\pm a$. In the limit $\beta \to \infty$, these will be solutions of zero "energy", since as $\tau \to \pm\infty$ both the kinetic and potential "energy" vanish.

First, if $q = q' = a$ (an identical argument applies if $q = q' = -a$), the obvious classical solution is $q(\tau) = a$; a quadratic approximation about this constant solution would be identical to the single-well case discussed above.

But what if $q = -a$ and $q' = a$ (or vice-versa)? Then the obvious classical solution corresponds to the particle initially sitting atop the maximum of $-V$ at $-a$, rolling towards

$q = 0$ after a very long (infinite, in the limit $\beta \to \infty$) time, and ending up at rest at the other maximum of $-V$ as $\tau \to \infty$ (Figure 19).

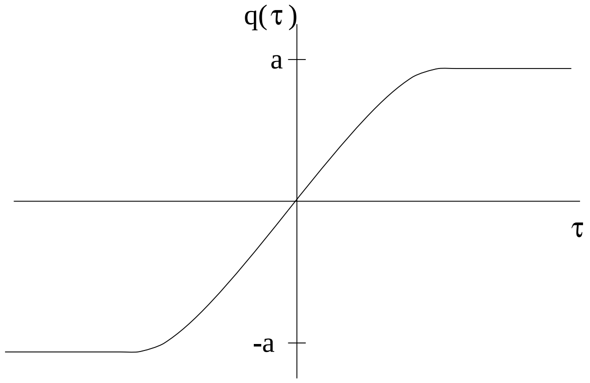

Figure 19: Instanton in the double-well potential.

We can get the analytical form of this solution: setting $E \to 0$, we have

$$\frac{1}{2}\dot{q}^2 = V(q), \qquad \text{or} \qquad \frac{dq}{d\tau} = \pm\sqrt{\frac{\lambda}{12}}(q^2 - a^2).$$

There are a family of solutions interpolating between $-a$ and a:

$$q(\tau) = a \tanh \frac{\omega}{2}(\tau - \tau_0), \tag{58}$$

where $\omega = \sqrt{\lambda a^2/3}$ and where τ_0 is an integration constant which corresponds to the time at which the solution crosses $q = 0$.

This solution is much like a topological soliton in field theory, except that it is localized in time rather than in space. One could argue that the solution doesn't appear to be localized: q goes to different values as $\tau \to \pm\infty$. But these are just different, but physically equivalent, ground states, so we can say that the instanton is a configuration which interpolates between two ground states; the system is in a ground state except for a brief time – an "instant". For this reason, the solution is known as an *instanton*.

I called the two solutions $q(\tau) = a$ and $q(\tau) = a \tanh \frac{\omega}{2}(\tau - \tau_0)$ the *obvious* classical solutions because there are an infinite number of approximate classical solutions which are potentially important in the PI. Since the instanton is localized in time, and since the total time interval β is very large (in particular, much larger than the instanton width), a series of widely-separated instantons and anti-instantons (configurations interpolating between $+a$ and $-a$) is also a solution, up to exponentially small interactions between neighbouring instantons and anti-instantons. Such a configuration is shown in Figure 20, where the horizontal scale has been determined by the duration of imaginary time β; on this scale the instanton and anti-instanton appear as step functions.

It is clear than an instanton must be followed by an anti-instanton, and that if the asymptotic values of the position are $+a$ and $+a$ the classical solution must contain anti-instanton-instanton pairs whereas if they are $-a$ and $+a$ we need an extra instanton at the beginning.

Let us choose first limiting values $q(-\beta/2) = q(+\beta/2) = +a$. Then we are interested in

$$K_E = \int_{a,-\beta/2}^{a,\beta/2} \mathcal{D}q\, e^{-S_E}.$$

In the spirit of (54), in the steepest-descent approximation K_E is equal to the sum of PIs evaluated about *all* classical solutions. The classical solutions are: $q_c(\tau) = a$; q_c =anti-instanton-instanton$\equiv$ AI; $q_c = AIAI$; *etc.*, where the positions of the As and Is are not determined, and must be integrated over. Schematically, we may write

$$K_E = K_E^0 + K_E^2 + K_E^4 + \cdots, \tag{59}$$

where the superscript denotes the total number of Is or As. Let us discuss the first couple of contributions in some detail.
$q_c = a$: This case is essentially equivalent to the single-well case discussed above, and we get

$$K_E^0 = \sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2},$$

where $\omega = (\lambda a^2/3)^{1/2}$ is the frequency of small oscillations about the minimum of V.
$q_c = AI$: This case is rather more interesting (that is to say, complicated!). Let us suppose that the classical solution around which we expand consists of an anti-instanton at time τ_1 and an instanton at τ_2 (Figure 21); clearly $\tau_2 > \tau_1$.

Then we can write $q = q_c + y$, and

$$S_E[q] = S_E[q_c] + S_E^{\rm quad}[y].$$

We can evaluate $S_E[q_c]$: it is twice the action of a single instanton (assuming the I and A are sufficiently far apart that any interaction is negligible): $S_E[q_c] = 2S_E^{\rm inst}$. The one-instanton action $S_E^{\rm inst}$ is

$$S_E^{\rm inst} = \int d\tau \left(\frac{1}{2}\dot{q}^2 + V(q)\right)\Big|_{\rm inst} = 2\int d\tau V(q)\Big|_{\rm inst}.$$

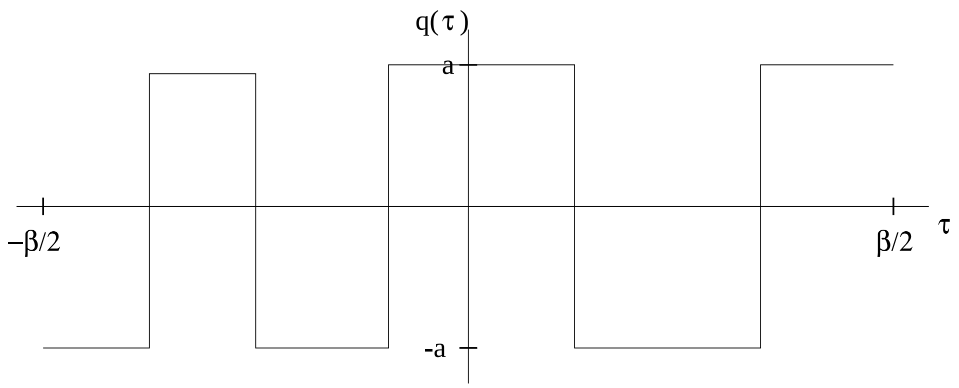

Figure 20: Multi-instanton configuration.

With the instanton profile given by (58), the result is

$$S_E^{\text{inst}} = \sqrt{\frac{\lambda}{3}}\frac{2a^3}{3}.$$

To evaluate the PI with the action $S_E^{\text{quad}}[y]$, let us divide the imaginary time interval into two semi-infinite regions I and II, where the boundary between the two regions is between and well away from the A and the I (Figure 22).

Then we can write

$$K_E^2 = \frac{\beta^2}{2} e^{-2S_E^{\text{inst}}} \int_{I+II} \mathcal{D}y\, e^{-S_E^{\text{quad}}/\hbar}. \tag{60}$$

Here the first factor represents integration over the positions of the A and I (remember that the A must be to the left of the I!). The quadratic action can be written

$$S_E^{\text{quad}} = {S_E^{\text{quad}}}_I + {S_E^{\text{quad}}}_{II},$$

where ${S_E^{\text{quad}}}_I$ is the quadratic action in the presence of an anti-instanton and ${S_E^{\text{quad}}}_{II}$ is that in the presence of an instanton.

Then the PI separates into two factors:

$$\int_{I+II} \mathcal{D}y\, e^{-S_E^{\text{quad}}/\hbar} = \int_I \mathcal{D}y\, e^{-{S_E^{\text{quad}}}_I/\hbar} \cdot \int_{II} \mathcal{D}y\, e^{-{S_E^{\text{quad}}}_{II}/\hbar}, \tag{61}$$

where there is an implied integration over the intermediate position at the boundary of the two regions. The quadratic no-instanton PI also separates into two factors:

$$\int \mathcal{D}y\, e^{-S_E^{\text{quad},0}/\hbar} = \int_I \mathcal{D}y\, e^{-{S_E^{\text{quad},0}}_I/\hbar} \times \int_{II} \mathcal{D}y\, e^{-{S_E^{\text{quad},0}}_{II}/\hbar}, \tag{62}$$

where the superscript "0" denotes that this is the PI about a no-instanton (constant) background. We can combine (61) and (62) to give:

$$\int_{I+II} \mathcal{D}y\, e^{-S_E^{\text{quad}}/\hbar} = \int \mathcal{D}y\, e^{-S_E^{\text{quad},0}/\hbar} \frac{\int_I \mathcal{D}y\, e^{-{S_E^{\text{quad}}}_I/\hbar}}{\int_I \mathcal{D}y\, e^{-{S_E^{\text{quad},0}}_I/\hbar}} \frac{\int_{II} \mathcal{D}y\, e^{-{S_E^{\text{quad}}}_{II}/\hbar}}{\int_{II} \mathcal{D}y\, e^{-{S_E^{\text{quad},0}}_{II}/\hbar}}. \tag{63}$$

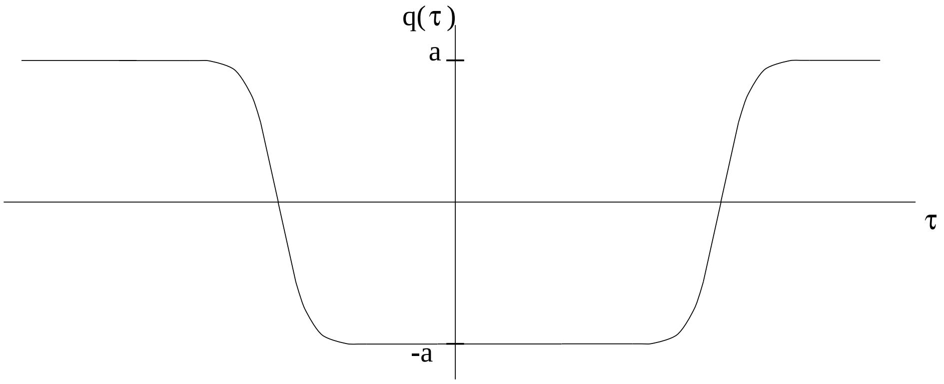

Figure 21: Anti-instanton-instanton.

But

$$\frac{\int_I \mathcal{D}y \, e^{-S_E^{\rm quad}{}_I/\hbar}}{\int_I \mathcal{D}y \, e^{-S_E^{\rm quad,0}{}_I/\hbar}} = \frac{\int \mathcal{D}y \, e^{-S_E^{\rm quad}/\hbar}}{\int \mathcal{D}y \, e^{-S_E^{\rm quad,0}/\hbar}} \tag{64}$$

and similarly for the last factor in (63), so we obtain

$$\int_{I+II} \mathcal{D}y \, e^{-S_E^{\rm quad}/\hbar} = \sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2} R^2,$$

where R is the ratio of the PI in the presence and absence of an instanton (or, equivalently, anti-instanton) given in (64). Substituting this into (60),

$$K_E^2 = e^{-2S_E^{\rm inst}/\hbar} \sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2} R^2 \frac{\beta^2}{2}.$$

A similar argument gives

$$K_E^4 = e^{-4S_E^{\rm inst}/\hbar} \sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2} R^4 \frac{\beta^4}{4!},$$

and so on for subsequent terms in the expansion (59).

Summing these contributions, we get

$$\begin{aligned} K_E &= \sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2} \left(1 + \frac{\left(\beta R e^{-S_E^{\rm inst}/\hbar}\right)^2}{2!} + \frac{\left(\beta R e^{-S_E^{\rm inst}/\hbar}\right)^4}{4!} + \cdots \right) \\ &= \sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2} \cosh\left(\beta R e^{-S_E^{\rm inst}/\hbar}\right) \\ &= \frac{1}{2}\sqrt{\frac{\omega}{\pi\hbar}} e^{-\beta\omega/2} \left(e^{\beta R e^{-S_E^{\rm inst}/\hbar}} + e^{-\beta R e^{-S_E^{\rm inst}/\hbar}} \right). \end{aligned}$$

Now we must recall why we're calculating this object in the first place. The propagator can be written as in (53):

$$K_E = \sum_n \langle a | \, n \rangle \langle n \, | a \rangle \, e^{-\beta E_n/\hbar}.$$

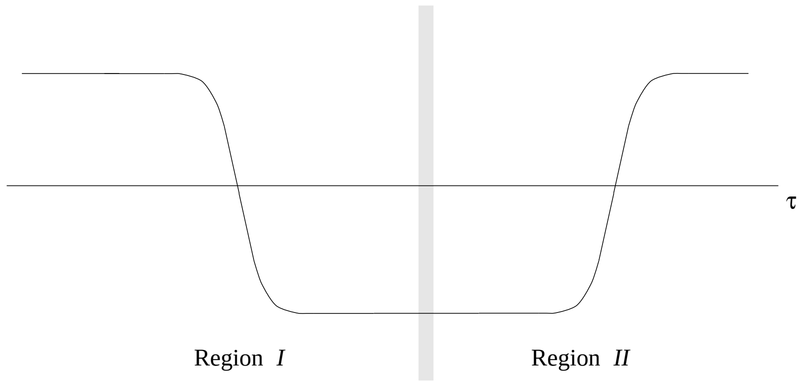

Figure 22: Division of imaginary time into two regions, one containing the anti-instanton, the other containing the instanton.

By comparing these two expressions we see that the lowest two energies are

$$\frac{\hbar\omega}{2} - \hbar R e^{-S_E^{\text{inst}}/\hbar} \qquad \text{and} \qquad \frac{\hbar\omega}{2} + \hbar R e^{-S_E^{\text{inst}}/\hbar}.$$

So the energy splitting is given by

$$\boxed{\Delta E = 2\hbar R e^{-S_E^{\text{inst}}/\hbar}.} \tag{65}$$

ΔE is clearly non-perturbative: it cannot be expanded as a power series in $\hbar$ (or, equivalently, in λ).

In principle, we should calculate the ratio

$$R = \frac{\text{(instanton background PI)}}{\text{(constant background PI)}} \sim \text{ratio of determinants},$$

but I don't know how to compute it other than by doing a very arduous, technical calculation; luckily, time will not permit it. The interested reader can consult the book by Sakita for a discussion of this calculation.

As a final note, we have calculated the PI with $q = q' = a$; a good exercise is to do the analogous calculation for $q = -a$, $q' = a$.

9.4 Instantons in a Periodic Potential

Consider a particle moving in a one-dimensional periodic potential (Figure 23).

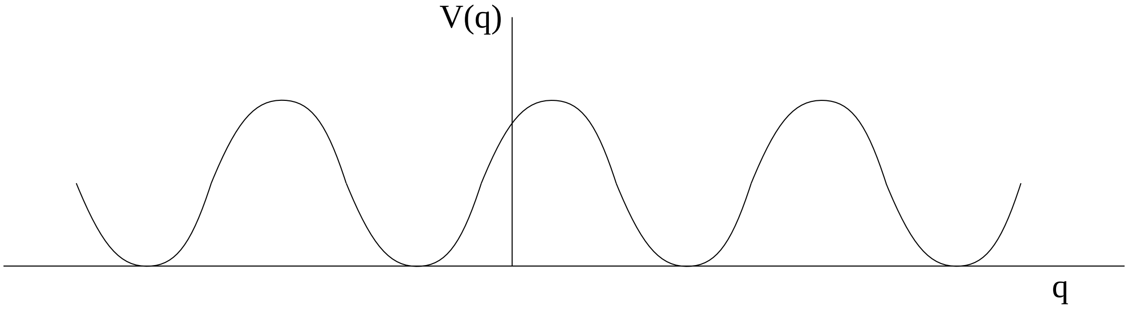

Figure 23: One-dimensional periodic potential.

With two minima, as we have seen, instantons enable us to calculate the energy splitting between the lowest-energy states of even and odd parity. In a periodic potential, we will see that a continuum of energies arise.

Let us label the classical minima of V by an integer, j. Clearly this model will have solutions analogous to the instantons above, going from any minimum of V to the adjacent minimum. We define an instanton as the classical solution going from any j to $j+1$, and an anti-instanton as that going from j to $j-1$. Then the Euclidean PI to go from $j=0$ to $j=0$, for instance, can be computed in a manner similar to the calculation of the previous section. This time any number and any order of instantons and anti-instantons are possible, subject to the constraint that $n_I = n_A$.

A calculation similar to that of the previous section results in the following expression for the propagator:

$$
\begin{aligned}
K_E(0,\beta/2;0,-\beta/2) &= \sqrt{\frac{\omega}{\pi\hbar}}e^{-\beta\omega/2}\sum_{n=0}^{\infty}\frac{1}{n!^2}\Bigg(\underbrace{e^{-S_E^{\mathrm{inst}}/\hbar}R\beta}_{\equiv Q}\Bigg)^{2n} \\
&= \sqrt{\frac{\omega}{\pi\hbar}}e^{-\beta\omega/2}\sum_{n,n'=0}^{\infty}\frac{Q^n}{n!}\frac{Q^{n'}}{n'!}\int_0^{2\pi}\frac{d\theta}{2\pi}e^{i\theta(n-n')} \\
&= \sqrt{\frac{\omega}{\pi\hbar}}e^{-\beta\omega/2}\int_0^{2\pi}\frac{d\theta}{2\pi}\sum_{n=0}^{\infty}\frac{(Qe^{i\theta})^n}{n!}\sum_{n'=0}^{\infty}\frac{(Qe^{-i\theta})^{n'}}{n'!} \\
&= \sqrt{\frac{\omega}{\pi\hbar}}e^{-\beta\omega/2}\int_0^{2\pi}\frac{d\theta}{2\pi}\exp\{Qe^{i\theta}\}\exp\{Qe^{-i\theta}\} \\
&= \sqrt{\frac{\omega}{\pi\hbar}}e^{-\beta\omega/2}\int_0^{2\pi}\frac{d\theta}{2\pi}\exp\left\{2\beta Re^{-S_E^{\mathrm{inst}}/\hbar}\cos\theta\right\}.
\end{aligned}
$$

The derivation of this is a worthwhile exercise. From it, we can read the energies:

$$E(\theta) = \frac{\hbar\omega}{2} - 2\hbar Re^{-S_E^{\mathrm{inst}}/\hbar}\cos\theta.$$

The second factor is an expression of the well-known result that the degeneracy is broken nonperturbatively; the energies form a continuum, depending on the value of θ.

Another application of instantons in quantum mechanics is the phenomenon of tunneling (barrier penetration). The instanton method can be used to calculate the lifetime of a metastable state in a potential of the form depicted in Figure 15. We will not discuss this application.

Instantons also appear in (and are by far most useful in) field theory. In certain field theories the space of finite-Euclidean-action configurations separates into distinct topological classes. An instanton is a nontrivial configuration of this type. The necessary topological requirements for this to occur are not hard to satisfy, and the list of theories that have instantons includes the Abelian Higgs model in 1+1 dimensions, the O(3) nonlinear σ-model in 1+1 dimensions, the Skyrme model in 2+1 dimensions, and (most significantly) QCD. Instantons give rise to a host of interesting phenomena depending on the model, including confinement (not in QCD though!), θ-vacua, a solution of the U(1) problem in strong interactions, and the decay of a metastable vacuum. Unfortunately time does not permit discussion of these fascinating phenomena.

10 Summary and Gross Omissions

In this set of lectures the subject of path integrals has been covered starting from scratch, emphasizing explicit calculations in quantum mechanics. This emphasis has its price: I have not had time to cover several things I would have liked to discuss. My hope is that having been subjected to calculations in gory detail for the most part in the relatively familiar context of quantum mechanics, you will be able to study more complicated and interesting applications on your own.

Here is a list of some of the important aspects and applications of this subject which I didn't have time to discuss:

1. Fermi fields and Grassmann functional integration;
2. Gauge theories (gauge fixing and ghosts arise in a particularly elegant way);
3. Feynman's variational method and application to the polaron (electron moving in a crystal environment);
4. Derivation of the Landau-Ginsburg theory, including application to superconductivity;
5. Instantons in field theory;
6. Critical phenomena.

I hope that, in spite of these unforgivable omissions, these lectures have been worthwhile.

11 Acknowledgements

I wish to thank the Rencontres du Vietnam and the NCST, Hanoi for the invitation to give these lectures, and especially Patrick Aurenche for his boundless energy before, during, and after the School. I am grateful to Jean-Sébastien Caux and Manu Paranjape for discussions and advice during the preparation of these lectures, and to Patrick Irwin, who spotted many weaknesses and errors in this manuscript after I had considered it fit for human consumption.

This work was financed in part by the Natural Sciences and Engineering Research Council of Canada.

12 References

1. P.A.M. Dirac, Physikalische Zeitschrift der Sowjetunion, Band 3, Heft 1 (1933).
2. R.P. Feynman, Reviews of Modern Physics 20, 367 1948.
3. R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, 1965.
4. R.P. Feynman, *Statistical Mechanics: A Set of Lectures*, Benjamin, 1972.
5. G. Baym, *Lectures on Quantum Mechanics*, Benjamin-Cummings, 1973.
6. L.S. Schulman, *Techniques and Applications of Path Integration*, John Wiley and Sons, 1981.
7. D.J. Amit, *Field Theory, the Renormalization Group and Critical Phenomena*, 2nd Edition, World Scientific, 1984.
8. L.H. Ryder, *Quantum Field Theory*, Cambridge University Press, 1985.
9. P. Ramond, *Field Theory: A Modern Primer, Second Edition*, Addison-Wesley, 1990.
10. E.S. Abers and B.W. Lee, Physics Reports 9C, 1, 1973.
11. B. Sakita, *Quantum Theory of Many-Variable Systems and Fields*, World Scientific, 1985.
12. S. Coleman, *Aspects of Symmetry*, Cambridge University Press, 1985.
13. A.M. Tsvelik, *Quantum Field Theory in Condensed Matter Physics*, Cambridge University Press, 1995.
14. V.N. Popov, *Functional Integrals and Collective Excitations*, Cambridge University Press, 1987.
15. E. Fradkin, *Field Theories of Condensed Matter Systems*, Addison-Wesley, 1991.