Chapter 3

Feynman Path Integral

The aim of this chapter is to introduce the concept of the Feynman path integral. As well as developing the general construction scheme, particular emphasis is placed on establishing the interconnections between the quantum mechanical path integral, classical Hamiltonian mechanics and classical statistical mechanics. The practice of path integration is discussed in the context of several pedagogical applications: As well as the canonical examples of a quantum particle in a single and double potential well, we discuss the generalisation of the path integral scheme to tunneling of extended objects (quantum fields), dissipative and thermally assisted quantum tunneling, and the quantum mechanical spin.

In this chapter we will temporarily leave the arena of many–body physics and second quantisation and, at least superficially, return to single–particle quantum mechanics. By establishing the path integral approach for ordinary quantum mechanics, we will set the stage for the introduction of functional field integral methods for many–body theories explored in the next chapter. We will see that the path integral not only represents a gateway to higher dimensional functional integral methods but, when viewed from an appropriate perspective, already represents a field theoretical approach in its own right. Exploiting this connection, various techniques and concepts of field theory, viz. stationary phase analyses of functional integrals, the Euclidean formulation of field theory, instanton techniques, and the role of topological concepts in field theory will be motivated and introduced in this chapter.

3.1 The Path Integral: General Formalism

Broadly speaking, there are two basic approaches to the formulation of quantum mechanics: the 'operator approach' based on the canonical quantisation of physical observables together with the associated operator algebra, and the Feynman[1] path integral.[2] Whereas canonical quantisation is usually taught first in elementary courses on quantum mechanics, path integrals seem to have acquired the reputation of being a sophisticated concept that is better reserved for advanced courses. Yet this treatment is hardly justified! In fact, the path integral formulation has many advantages most of which explicitly support an intuitive understanding of quantum mechanics. Moreover, integrals — even the infinite dimensional ones encountered below — are hardly more abstract than infinite dimensional linear operators. Further merits of the path integral include the following:

- ▷ Whereas the classical limit is not always easy to retrieve within the canonical formulation of quantum mechanics, it constantly remains visible in the path integral approach. In other words, the path integral makes explicit use of classical mechanics as a basic 'platform' on which to construct a theory of quantum fluctuations. The classical solutions of Hamilton's equation of motion always remain a central ingredient of the formalism.[3]

- ▷ Path integrals allow for an efficient formulation of *non–perturbative* approaches to the solution of quantum mechanical problems. For example, the 'instanton' formulation of quantum tunnelling discussed below — whose extension to continuum theories has led to some of the most powerful concepts of quantum field theory — makes extensive use of the classical equations of motion when it is tailored to a path integral formulation.

- ▷ The Feynman path integral represents a prototype of the higher dimensional functional field integrals to be introduced in the next chapter. However,...

- ▷ ...even in its 'zero–dimensional' form discussed in this chapter, the path integral is of relevance to a wide variety of applications in many–body physics: Very often, one encounters enviroments such as the superconductor, superfluid, or strongly correlated few electron devices where a macroscopically large number of degrees of

[1] Richard P. Feynman 1918–1988: 1965 Nobel Laureate in Physics (with Sin–Itiro Tomonaga, and Julian Schwinger) for fundamental work in quantum electrodynamics, with deep–ploughing consequences for the physics of elementary particles.

[2] For a more extensive introduction to the Feynman path integral, one can refer to one of the many standard texts including Refs. [9, 16, 20] or, indeed, one may turn to the original paper, R. P. Feynman, *Space–time approach to non–relativistic quantum mechanics*, Rev. Mod. Phys. **20**, 367 (1948). Historically, Feynman's development of the path integral was motivated by earlier work by Dirac on the connection between classical and quantum mechanics, P. A. M. Dirac, *On the analogy between classical and quantum mechanics*, Rev. Mod. Phys. **17**, 195 (1945).

[3] For this reason, path integration has turned out to be an indispensable tool in fields such as **quantum chaos** where the quantum manifestations of classically non–trivial behaviour are investigated — for more details, see section 3.2.2 below.

freedom 'lock' to form a single collective variable. (For example, to a first approximation, the phase information carried by the order parameter field in moderatly large superconducting grains can often be described in terms of a *single* phase degree of freedom, i.e. a 'quantum particle' living on the complex unit circle.) Path integral techniques have proven ideally suited to the analysis of such systems.

What, then, is the basic idea of the path integral approach? More than any other formulation of quantum mechanics, the path integral formalism is based on connections to classical mechanics. The variational approach employed in chapter **??** relied on the fact that classically allowed trajectories in configuration space extremize an action functional. A principal constraint to be imposed on any such trajectory is energy conservation. By contrast, quantum particles have more freedom than their classical counterparts. In particular, by the Uncertainty Principle, energy conservation can be violated by an amount ΔE over a time $\sim \hbar/\Delta E$ (here, and throughout this chapter, we will reinstall $\hbar$ for clarity). The connection to action principles of classical mechanics becomes particularly apparent in problems of quantum tunneling: A particle of energy E may tunnel through a potential barrier of height $V > E$. However, this process is penalized by a damping factor $\sim \exp(i \int_{\text{barrier}} dx\, p/\hbar)$, where $p = \sqrt{2m(E-V)}$, i.e. the exponent of the (imaginary) action associated with the classically forbidden path.

These observations motivate the idea of a new formulation of quantum propagation: Could it be that, as in classical mechanics, the quantum amplitude A for propagation between any two points in coordinate space is again controlled by the action functional? — controlled in a relaxed sense where not just a single extremal path $x_{\text{cl}}(t)$, but an entire manifold of neighbouring paths contribute. More specifically, one might speculate that the quantum amplitude is obtained as $A \sim \sum_{x(t)} \exp(iS[x]/\hbar)$, where $\sum_{x(t)}$ symbolically stands for a summation over all paths compatible with the initial conditions of the problem, and S denotes the *classical* action. Although, at this stage, no formal justification for the path integral has been presented, with this *ansatz*, some features of quantum mechanics would obviously be born out correctly: Specifically, in the classical limit ($\hbar \to 0$), the quantum mechanical amplitude would become increasingly dominated by the contribution to the sum from the classical path $x_{\text{cl}}(t)$. This is because non–extremal configurations would be weighted by a rapidly oscillating amplitude associated with the large phase $S/\hbar$ and would, therefore, average to zero.[4] Secondly, quantum mechanical tunneling would be a natural element of the theory; non–classical paths do contribute to the net–amplitude, but at the cost of a damping factor specified by the imaginary action (as in the traditional formulation).

Fortunately, no fundamentally novel 'picture' of quantum mechanics needs to be declared to promote the idea of the path 'integral' $\sum_{x(t)} \exp(iS[x]/\hbar)$ to a working theory. As we will see in the next section, the new formulation can quantitatively be developed from the same principles of canonical quantization.

[4]More precisely, in the limit of small $\hbar$, the path sum can be evaluated by saddle–point methods, as detailed below.

3.2 Construction of the Path Integral

All information about any autonomous[5] quantum mechanical system is contained in the matrix elements of its time evolution operator. A formal integration of the time–dependent Schrödinger equation $i\hbar\partial_t|\Psi\rangle = \hat{H}|\Psi\rangle$ obtains the time evolution operator

$$|\Psi(t')\rangle = \hat{U}(t',t)|\Psi(t)\rangle, \qquad \hat{U}(t',t) = e^{-\frac{i}{\hbar}\hat{H}(t'-t)}\Theta(t'-t)\,. \tag{3.1}$$

The operator $\hat{U}(t',t)$ describes dynamical evolution under the influence of the Hamiltonian from a time t to time t'. Causality implies that $t' > t$ as indicated by the step or Heaviside Θ–function. In the real space representation we can write

$$\Psi(q',t') = \langle q'|\Psi(t')\rangle = \langle q'|\hat{U}(t',t)\Psi(t)\rangle = \int dq\ U(q',t';q,t)\Psi(q,t)\,,$$

where $U(q',t';q,t) = \langle q'|e^{-\frac{i}{\hbar}\hat{H}(t'-t)}|q\rangle\Theta(t'-t)$ defines the (q',q) component of the time evolution operator. As the matrix element expresses the probability amplitude for a particle to propagate between points q and q' in a time $t'-t$, it is sometimes known as the **propagator** of the theory.

The basic idea behind Feynman's path integral approach is easy to formulate. Rather than attacking the Schrödinger equation governing the time evolution for general times t, one may first attempt to solve the much simpler problem of describing the time evolution for infinitesimally small times Δt. In order to formulate this idea quantitatively one must first 'divide' the time evolution operator into $N \gg 1$ discrete 'time steps',

$$e^{-i\hat{H}t/\hbar} = \left[e^{-i\hat{H}\Delta t/\hbar}\right]^N\,, \tag{3.2}$$

where $\Delta t = t/N$. Albeit nothing more than a formal rewriting of Eq. (3.1), the representation (3.2) has the advantage that the factors $e^{-i\hat{H}\Delta t/\hbar}$ (or, rather, their expectation values) are small. (More precisely, if Δt is much smaller than the (reciprocal of the) eigenvalues of the Hamiltonian in the regime of physical interest, the exponents are small in comparison with unity and, as such, can be treated perturbatively.) A first simplification arising from this fact is that the exponentials can be factorised into two pieces each of which can be readily diagonalised. To achieve this factorisation, we make use of the identity

$$e^{-i\hat{H}\Delta t/\hbar} = e^{-i\hat{T}\Delta t/\hbar}e^{-i\hat{V}\Delta t/\hbar} + O(\Delta t^2)\,,$$

where the Hamiltonian $\hat{H} = \hat{T} + \hat{V}$ is the sum of a kinetic energy $\hat{T} = \hat{p}^2/2m$, and some potential energy operator $\hat{V}$.[6] (The following analysis, restricted for simplicity to a one–dimensional Hamiltonian, is easily generalised to arbitrary spatial dimension.) The key

[5] A system is classified as **autonomous** if its Hamiltonian does not explicitly depend on time. Actually the construction of the path integral can be straightforwardly extended so as to include time–dependent problems. However, in order to keep the introductory discussion as simple as possible, here we assume time–independence.

[6] Although this *ansatz* already covers a wide class of quantum mechanical problems, many applications of practical importance (e.g. Hamiltonians involving spin or magnetic fields) do not fit into this framework. For a detailed exposition covering its realm of applicability, we refer to the specialist literature such as, e.g., Schulman's text [20].

advantage of this factorisation is that the eigenstates of each of the factors $e^{-i\hat{T}\Delta t/\hbar}$ and $e^{-i\hat{V}\Delta t/\hbar}$ are known independently. To exploit this fact we consider the time evolution operator factorised as a product,

$$\langle q_F| \left[e^{-i\hat{H}\Delta t/\hbar}\right]^N |q_I\rangle \simeq \langle q_F|_\wedge e^{-i\hat{T}\Delta t/\hbar} e^{-i\hat{V}\Delta t/\hbar}{}_\wedge \cdots {}_\wedge e^{-i\hat{T}\Delta t/\hbar} e^{-i\hat{V}\Delta t/\hbar}|q_I\rangle \tag{3.3}$$

and insert at each of the positions indicated by the symbol '$\wedge$' the resolution of identity

$$\text{id.} = \int dq_n \int dp_n |q_n\rangle\langle q_n|p_n\rangle\langle p_n|\,. \tag{3.4}$$

Here $|q_n\rangle$ and $|p_n\rangle$ represent a complete set of position and momentum eigenstates respectively, and $n = 1, \ldots, N$ serves as an index keeping track of the time steps at which the unit operator is inserted. The rational behind the particular choice (3.4) is clear. The unit operator is arranged in such a way that both $\hat{T}$ and $\hat{V}$ act on the corresponding eigenstates. Inserting (3.4) into (3.3), and making use of the identity $\langle q|p\rangle = \langle p|q\rangle^* = e^{iqp/\hbar}/(2\pi\hbar)$, one obtains

$$\langle q_F|e^{-i\hat{H}t/\hbar}|q_I\rangle \simeq \int_{q_N=q_F, q_0=q_I} \prod_{n=1}^{N-1} dq_n \prod_{n=1}^{N} \frac{dp_n}{2\pi\hbar} e^{-i\frac{\Delta t}{\hbar}\sum_{n=0}^{N-1}\left(V(q_n)+T(p_{n+1})-p_{n+1}\frac{q_{n+1}-q_n}{\Delta t}\right)}\,. \tag{3.5}$$

Thus, the matrix element of the time evolution operator has been expressed as a $2N - 1$ dimensional integral over eigenvalues. Up to corrections of higher order in $V\Delta t/\hbar$ and $T\Delta t/\hbar$, the expression (3.5) is exact. At each 'time step' $t_n = n\Delta t$, $n = 1, \ldots, N$ we are integrating over a pair of coordinates $x_n \equiv (q_n, p_n)$ parametrising the **classical phase space**. Taken together, the points $\{x_n\}$ form an N–point discretization of a path in this space (see Fig. 3.1).

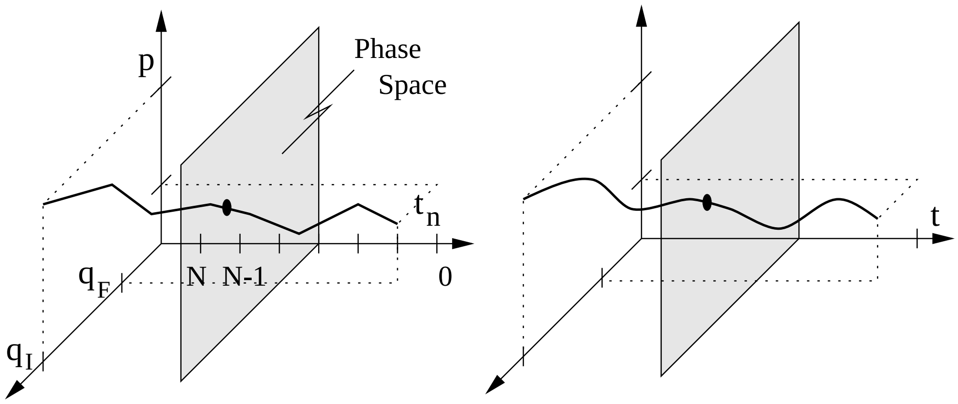


Figure 3.1: Left: visualisation of a set of phase space points contributing to the discrete time configuration integral (3.5). Right: in the continuum limit, the set of points becomes a smooth curve.

To make further progress, we need to develop some intuition for the behaviour of the integral (3.5). We first notice that rapid fluctuations of the integration arguments x_n as a function of the index n are strongly inhibited by the structure of the integrand. When taken together, contributions for which $(q_{n+1}-q_n)p_{n+1} > O(\hbar)$ (i.e. when the phase of the exponential exceeds 2π) tend to lead to a 'random phase cancellation'. In the language of wave mechanics, the 'incoherent' superposition of different Feynman paths destructively interferes. The *smooth* variation of the paths which contribute significantly motivate the application of a continuum limit analogous to that employed in chapter **??**.

To be specific, sending $N \to \infty$ whilst keeping $t = N\Delta t$ fixed, the formerly discrete set $t_n = n\Delta t,\ n = 1, \ldots, N$ becomes dense on the time interval $[0, t]$, and the set of phase space points $\{x_n\}$ becomes a continuous curve $x(t)$. In the same limit,

$$\Delta t \sum_{n=0}^{N-1} \mapsto \int_0^t dt', \qquad \frac{q_{n+1}-q_n}{\Delta t} \mapsto \partial_{t'} q\Big|_{t'=t_n} \equiv \dot{q}|_{t'=t_n},$$

while $[V(q_n) + T(p_{n+1})] \mapsto [T(p|_{t'=t_n}) + V(q|_{t'=t_n})] \equiv H(x|_{t'=t_n})$ denotes the *classical* Hamiltonian. In the limit $N \to \infty$, the fact that kinetic and potential energies are evaluated at neighbouring time slices, n and $n+1$, becomes irrelevant.[7] Finally,

$$\lim_{N\to\infty} \int_{q_N=q_F, q_0=q_I} \prod_{n=1}^{N-1} dq_n \prod_{n=1}^{N} \frac{dp_n}{2\pi\hbar} \equiv \int_{\substack{q(t)=q_F\\ q(0)=q_I}} Dx$$

defines the integration measure of the integral.

▷ INFO. Integrals extending over infinite dimensional integration measures like $D(q,p)$ are generally called **functional integrals** (recall our discussion of functionals in chapter **??**). The question of how functional integration can be rigorously defined is far from innocent and represents a subject of current, and partly controversial mathematical research. In this book — as in most applications in physics — we take a pragmatic point of view and deal with the infinite dimensional integration naively unless mathematical problems arise (which actually won't be the case!).

Then, applying these conventions to Eq. (3.5), one finally obtains

$$\boxed{\langle q_F | e^{-i\hat{H}t/\hbar} | q_I \rangle = \int_{\substack{q(t)=q_F\\ q(0)=q_I}} Dx \exp\left[\frac{i}{\hbar}\int_0^t dt' \left(p\dot{q} - H(p,q)\right)\right]} \tag{3.6}$$

[7]To see this formally, one may Taylor expand $T(p_{n+1}) = T(p(t'+\Delta t))|_{t'=n\Delta t}$ around $p(t')$. For smooth $p(t')$, all but the zeroth order contribution $T(p(t'))$, scale with powers of Δt, thereby becoming irrelevant. Note, however, that all of these arguments are based on the assertion that the dominant contributions to the path integral are smooth in the sense $q_{n+1} - q_n \sim O(\Delta t)$. A closer inspection, however, shows that in fact $q_{n+1} - q_n \sim O(\sqrt{\Delta t})$ [20]. In some cases, the most prominent one being the quantum mechanics of a particle in a magnetic field, the lowered power of Δt spoils the naive form of the continuity argument above, and more care must be applied in taking the continuum limit. In cases where a 'new' path integral description of a quantum mechanical problem is developed, it is imperative to delay taking the continuum limit until the fluctuation behaviour of the discrete integral across individual time slices has been thoroughly examined.

Eq. (3.6) represents the **Hamiltonian formulation of the path integral**: The integration extends over all possible paths through the classical phase space of the system which begin and end at the same *configuration* points q_I and q_F respectively (cf. Fig. 3.1). The contribution of each path is weighted by its Hamiltonian action.

Before we turn to the discussion of the path integral (3.6), it is first useful to recast the integral in an alternative form which will be both convenient in various applications and physically instructive. The search for an alternative formulation is motivated by the observation of the close resemblance of (3.6) with the Hamiltonian formulation of classical mechanics. Given that, classically, Hamiltonian and *Lagrangian* mechanics can be equally employed to describe dynamical evolution, it is natural to seek a Lagrangian analogue of (3.6). Until now, we have made no assumption about the momentum dependence of the kinetic energy $T(p)$. However, if we focus on Hamiltonians in which the dynamics is free, i.e. the kinetic energy dependence is quadratic in p, the Lagrangian form of the path integral can be inferred from (3.6) by Gaussian integration.

To make this point clear, let us rewrite the integral in a way that emphasises its dependence on the momentum variable p:

$$\langle q_F|e^{-i\hat{H}t/\hbar}|q_I\rangle = \int_{\substack{q(t)=q_F\\ q(0)=q_I}} Dq\ e^{-\frac{i}{\hbar}\int_0^t dt' V(q)} \int Dp\ e^{-\frac{i}{\hbar}\int_0^t dt'\left(\frac{p^2}{2m}-p\dot{q}\right)}\,. \tag{3.7}$$

The exponent of the integral is quadratic in the momentum variable or, equivalently, the integral is Gaussian in p. Carrying out the integration by means of Eq. (3.13) below, one obtains

$$\boxed{\langle q_F|e^{-i\hat{H}t/\hbar}|q_I\rangle = \int_{\substack{q(t)=q_F\\ q(0)=q_I}} Dq\ \exp\left[\frac{i}{\hbar}\int_0^t dt' L(q,\dot{q})\right]} \tag{3.8}$$

where $Dq = \lim_{N\to\infty}(\frac{Nm}{it2\pi\hbar})^{N/2}\prod_{n=1}^{N-1} dq_n$ denotes the functional measure of the remaining q–integration, and $L(q,\dot{q}) = m\dot{q}^2/2 - V(q)$ represents the classical Lagrangian. Strictly speaking, the (finite–dimensional) integral formula (3.13) is not directly applicable to the infinite dimensional Gaussian integral (3.7). This, however, does not represent a substantial problem as we can always rediscretise the integral (3.7), apply Eq. (3.13), and reinstate the continuum limit after integration (exercise).

Together Eqs. (3.6) and (3.8) represent the central results of this section. A quantum mechanical transition amplitude has been expressed in terms of an infinite dimensional integral extending over paths through phase space (3.6) or coordinate space (3.8). All paths begin (end) at the initial (final) coordinate of the matrix element. Each path is weighted by its *classical* action. Notice in particular that the quantum transition amplitude has been cast in a form which does not contain quantum mechanical operators. Nonetheless, quantum mechanics is still fully present! The point is that the integration extends over *all* paths and not just the subset of solutions of the classical equations of motion. (The distinguished role classical paths play in the path integral will be discussed below in section 3.2.2.) The two forms of the path integral, (3.6) and (3.8), represent the formal implementation of the 'alternative picture' of quantum mechanics proposed heuristically at the beginning of the chapter.

▷ INFO. **Gaussian Integration**: Apart from a few rare exceptions, all integrals encountered in this course will be of Gaussian[8] form. In most cases the dimension of the integrals will be large if not infinite. Yet, after a bit of practice, it will become clear that high dimensional representatives of Gaussian integrals are no more difficult to handle than their one–dimensional counterparts. Therefore, considering the important role played by Gaussian integration in field theory, we will here derive the principle formulae once and for all. Our starting point is the one–dimensional integral (both real and complex). The basic ideas underlying the proofs of the one–dimensional formulae, will provide the key to the derivation of more complex, multi–dimensional and functional identities which will be used liberally throughout the remainder of the text.

One–dimensional Gaussian integral: The basic ancestor of all Gaussian integrals is the identity

$$\boxed{\int_{-\infty}^{\infty} dx\, e^{-\frac{a}{2}x^2} = \sqrt{\frac{2\pi}{a}}, \qquad \mathrm{Re}\, a > 0} \tag{3.9}$$

In the following we will need various generalisations of Eq. (3.9). Firstly, we have $\int_{-\infty}^{\infty} dx\, e^{-ax^2/2}x^2 = \sqrt{2\pi/a^3}$, a result established either by substituting $a \to a + \epsilon$ in Eq. (3.9) and expanding both the left and the right side of the equation to leading order in ϵ, or by differentiating Eq. (3.9). Often one encounters integrals where the exponent is not purely quadratic from the outset but rather contains both quadratic and linear pieces. The generalisation of Eq. (3.9) to this case reads

$$\int_{-\infty}^{\infty} dx\, e^{-\frac{a}{2}x^2+bx} = \sqrt{\frac{2\pi}{a}} e^{\frac{b^2}{2a}}\,. \tag{3.10}$$

To prove this identity, one simply eliminates the linear term by means of the change of variables $x \to x + b/a$ which transforms the exponent to $ax^2/2 + bx \to -ax^2/2 + b^2/2a$. The constant factor scales out and we are left with Eq. (3.9). Note that Eq. (3.10) holds even for complex b. The reason is that by shifting the integration contour into the complex plane no singularities are encountered, i.e. the integral remains invariant.

Later, we will be concerned with the generalisation of the Gaussian integral to complex arguments. In this case, the extension of Eq. (3.9) reads

$$\int d(\bar{z}, z) e^{-\bar{z}wz} = \frac{\pi}{w}, \qquad \mathrm{Re}\, w > 0\,,$$

where $\bar{z}$ represents the complex conjugate of z. Here, $\int d(\bar{z}, z) \equiv \int_{-\infty}^{\infty} dxdy$ represents the independent integration over the real and imaginary parts of $z = x + iy$. The identity is easy to prove: Owing to the fact that $\bar{z}z = x^2 + y^2$, the integral factorizes into two pieces each of which is equivalent to Eq. (3.9) with $a = w$. Similarly, it may be checked that the complex

[8] **Johann Carl Friedrich Gauss** 1777-1855: worked in a wide variety of fields in both mathematics and physics incuding number theory, analysis, differential geometry, geodesy, magnetism, astronomy and optics. Portrait taken from the former German 10–Mark note. (Unfortunately, the subsequently introduced Euro notes no longer display Gauss' portrait.)

generalisation of Eq. (3.10) is given by

$$\int d(\bar{z},z)e^{-\bar{z}wz+\bar{u}z+\bar{z}v} = \frac{\pi}{w}e^{\frac{\bar{u}v}{w}}, \qquad \mathrm{Re}\, w > 0\,. \tag{3.11}$$

More importantly $\bar{u}$ and v may be *independent* complex numbers; they need not be related to each other by complex conjugation (exercise).

Gaussian integration in more than one dimension: All of the integrals above have higher dimensional counterparts. Although the real and complex versions of the N–dimensional integral formulae can be derived in a perfectly analogous manner, it is better to discuss them seperately in order not to confuse the notation.

(a) Real Case: The multi–dimensional generalisation of the prototype integral (3.9) reads

$$\int d\mathbf{v}e^{-\frac{1}{2}\mathbf{v}^T\mathbf{A}\mathbf{v}} = (2\pi)^{N/2}\det\mathbf{A}^{-1/2}, \tag{3.12}$$

where $\mathbf{A}$ is a positive definite real symmetric N–dimensional matrix and $\mathbf{v}$ is an N–component real vector. The proof makes use of the fact that $\mathbf{A}$ (by virtue of being symmetric) can be diagonalised by orthogonal transformation, $\mathbf{A} = \mathbf{O}^T\mathbf{D}\mathbf{O}$, where the matrix $\mathbf{O}$ is orthogonal, and all elements of the diagonal matrix $\mathbf{D}$ are positive. The matrix $\mathbf{O}$ can be absorbed into the integration vector by means of the variable transformation, $\mathbf{v}\mapsto\mathbf{O}\mathbf{v}$ which has unit Jacobian, $\det\mathbf{O} = 1$. As a result, we are left with a Gaussian integral with exponent $-\mathbf{v}^T\mathbf{D}\mathbf{v}/2$. Due to the diagonality of $\mathbf{D}$, the integral factorizes into N independent Gaussian integrals each of which contributes a factor $\sqrt{2\pi/d_i}$, where d_i, $i = 1,\ldots,N$ is the ith entry of the matrix $\mathbf{D}$. Noting that $\prod_{i=1}^N d_i = \det\mathbf{D} = \det\mathbf{A}$, (3.12) is derived.

The multi–dimensional generalization of (3.10) reads

$$\int d\mathbf{v}e^{-\frac{1}{2}\mathbf{v}^T\mathbf{A}\mathbf{v}+\mathbf{j}^T\cdot\mathbf{v}} = (2\pi)^{N/2}\det\mathbf{A}^{-1/2}e^{\frac{1}{2}\mathbf{j}^T\mathbf{A}^{-1}\mathbf{j}} \tag{3.13}$$

where $\mathbf{j}$ is an arbitrary N–component vector. Eq. (3.13) is proven by analogy with Eq. (3.10), i.e. by shifting the integration vector according to $\mathbf{v}\to\mathbf{v}+\mathbf{A}^{-1}\mathbf{j}$, which does not change the value of the integral but removes the linear term from the exponent, $-\frac{1}{2}\mathbf{v}^T\mathbf{A}\mathbf{v}+\mathbf{j}^T\cdot\mathbf{v}\to -\frac{1}{2}\mathbf{v}^T\mathbf{A}\mathbf{v}+\frac{1}{2}\mathbf{j}^T\mathbf{A}^{-1}\mathbf{j}$. The resulting integral is of the type (3.12), and we arrive at Eq. (3.13).

The integral (3.13) is not only of importance in its own right, but it also serves as a 'generator' of other useful integral identities. Applying the differentiation operation $\partial^2_{j_m j_n}|_{\mathbf{j}=0}$ to the left and the right hand side of Eq. (3.13), one obtains the identity[9] $\int d\mathbf{v}e^{-\frac{1}{2}\mathbf{v}^T\mathbf{A}\mathbf{v}}v_m v_n = (2\pi)^{N/2}\det\mathbf{A}^{-1/2}A^{-1}_{mn}$. This result can be more compactly formulated as

$$\langle v_m v_n\rangle = A^{-1}_{mn}, \tag{3.14}$$

where we have introduced the shorthand notation

$$\langle\ldots\rangle \equiv (2\pi)^{-N/2}\det\mathbf{A}^{1/2}\int d\mathbf{v}e^{-\frac{1}{2}\mathbf{v}^T\mathbf{A}\mathbf{v}}(\ldots), \tag{3.15}$$

suggesting an interpretation of the Gaussian weight as a probability distribution.

Indeed, the differentiation operation leading to (3.14) can be iterated: Differentiating four times, one obtains $\langle v_m v_n v_q v_p\rangle = A^{-1}_{mn}A^{-1}_{qp} + A^{-1}_{mq}A^{-1}_{np} + A^{-1}_{mp}A^{-1}_{nq}$. One way of memorising the

[9]Note that the notation A^{-1}_{mn} refers to the mn element of the matrix $\mathbf{A}^{-1}$.

structure of this — important — identity is that the Gaussian 'expectation' value $\langle v_m v_n v_p v_q \rangle$ is given by all 'pairings' of type (3.14) that can be formed from the four components v_m. This rule generalises to expectation values of arbitrary order: $2n$–fold differentiation of (3.13) yields

$$\boxed{\langle v_{i_1} v_{i_2} \dots v_{i_{2n}} \rangle = \sum_{\substack{\text{all possible} \\ \text{pairings of}\{i_1,\dots i_{2n}\}}} A^{-1}_{i_{k_1} i_{k_2}} \dots A^{-1}_{i_{k_{2n-1}} i_{k_{2n}}}} \tag{3.16}$$

This result is the mathematical identity underlying **Wick's theorem** (for real bosonic fields).

(b) Complex Case: The results above are straightforwardly extended to multi–dimensional complex Gaussian integrals. The complex version of Eq. (3.12) is given by

$$\int d(\mathbf{v}^\dagger, \mathbf{v}) e^{-\mathbf{v}^\dagger \mathbf{A} \mathbf{v}} = \pi^N \det \mathbf{A}^{-1} \,, \tag{3.17}$$

where $\mathbf{v}$ is a complex N–component vector, $d(\mathbf{v}^\dagger, \mathbf{v}) \equiv \prod_{i=1}^N d\mathrm{Re}\, v_i \, d\mathrm{Im}\, v_i$, and $\mathbf{A}$ is a complex matrix with positive definite Hermitian part. (Remember that every matrix can be decomposed into a Hermitian and an anti–Hermitian component, $\mathbf{A} = \frac{1}{2}(\mathbf{A}+\mathbf{A}^\dagger)+\frac{1}{2}(\mathbf{A}-\mathbf{A}^\dagger)$.) For Hermitian $\mathbf{A}$, the proof of (3.17) is analogous to (3.12), i.e. $\mathbf{A}$ is unitarily diagonalisable, $\mathbf{A} = \mathbf{U}^\dagger \mathbf{A} \mathbf{U}$; the matrices $\mathbf{U}$ can be transformed into $\mathbf{v}$, the resulting integral factorises, etc. For non–Hermitian $\mathbf{A}$ the proof is more elaborate, if unedifying, and we refer to the literature for details. The generalization of Eq. (3.17) to exponents with linear contributions reads

$$\int d(\mathbf{v}^\dagger, \mathbf{v}) e^{-\mathbf{v}^\dagger \mathbf{A} \mathbf{v} + \mathbf{w}^\dagger \cdot \mathbf{v} + \mathbf{v}^\dagger \cdot \mathbf{w}'} = \pi^N \det \mathbf{A}^{-1} e^{\mathbf{w}^\dagger \mathbf{A}^{-1} \mathbf{w}'} \tag{3.18}$$

Note that $\mathbf{w}$ and $\mathbf{w}'$ may be *independent* complex vectors. The proof of this identity mirrors that of (3.13), i.e. by effecting the shift $\mathbf{v}^\dagger \to \mathbf{v}^\dagger + \mathbf{w}^\dagger$, $\mathbf{v} \to \mathbf{v} + \mathbf{w}'$.[10] As with Eq. (3.13), Eq. (3.18) may also serve as a generator of related integral identities. Differentiating the integral twice according to $\partial^2_{w_m, w'_n}|_{\mathbf{w}=\mathbf{w}'=0}$ gives

$$\langle \bar{v}_m v_n \rangle = A^{-1}_{nm} \,,$$

where $\langle \cdots \rangle \equiv \pi^{-N} \det \mathbf{A} \int d(\mathbf{v}^\dagger, \mathbf{v}) e^{-\mathbf{v}^\dagger \mathbf{A} \mathbf{v}} (\cdots)$. The iteration to more than two derivatives gives $\langle \bar{v}_n \bar{v}_m v_p v_q \rangle = A^{-1}_{pm} A^{-1}_{qn} + A^{-1}_{pn} A^{-1}_{qm}$ and, eventually,

$$\boxed{\langle \bar{v}_{i_1} \bar{v}_{i_2} \dots \bar{v}_{i_n} v_{j_1} v_{j_2} \dots v_{j_n} \rangle = \sum_P A^{-1}_{j_1 i_{P1}} \dots A^{-1}_{j_n i_{Pn}}}$$

where $\sum_P$ represents for the sum over all permutations of N integers.

Gaussian Functional Integration: With this preparation, we are in a position to investigate the main practice of quantum and statistical field theory — the method of Gaussian functional integration. Turning to Eq. (3.13), let us suppose that the components of the vector $\mathbf{v}$ parameterise the weight of a real scalar field on the sites of a one–dimensional lattice. In the continuum limit, the set $\{v_i\}$ translates to a function $v(x)$, and the matrix A_{ij} is replaced by an

[10] For an explanation of why $\mathbf{v}$ and $\mathbf{v}^\dagger$ may be shifted independently of each other, cf. the analyticity remarks made in connection with (3.11).

operator kernel or **propagator** $A(x, x')$. In this limit, the natural generalisation of Eq. (3.13) is

$$\int Dv(x) \exp\left[-\frac{1}{2}\int dx\ dx' v(x)A(x,x')v(x') + \int dx\ j(x)v(x)\right] \\ \propto (\det A)^{-1/2} \exp\left[\frac{1}{2}\int dx\ dx'\ j(x)A^{-1}(x,x')j(x')\right], \qquad (3.19)$$

where the inverse kernel $A^{-1}(x, x')$ satisfies the equation

$$\boxed{\int dx'\ A(x,x')A^{-1}(x',x'') = \delta(x-x'')} \qquad (3.20)$$

i.e. $A^{-1}(x, x')$ can be interpreted as the **Green function** of the operator $A(x, x')$. The notation $Dv(x)$ is used to denote the measure of the functional integral. Although the constant of proportionality, $(2\pi)^N$ left out of Eq. (3.19) is formally divergent in the thermodynamic limit $N \to \infty$, it does not effect averages that are obtained from derivatives of such integrals. For example, for Gaussian distributed functions, Eq. (3.14) has the generalisation

$$\boxed{\langle v(x)v(x')\rangle = A^{-1}(x,x')}$$

Accordingly, Eq. (3.16) assumes the form

$$\boxed{\langle v(x_1)v(x_2)\dots v(x_{2n})\rangle = \sum_{\substack{\text{all possible}\\ \text{pairings of}\{x_1,\dots x_{2n}\}}} A^{-1}(x_{k_1},x_{k_2})\dots A^{-1}(x_{k_{2n-1}},x_{k_{2n}})} \qquad (3.21)$$

The generalization of the other Gaussian averaging formulae discussed above should be obvious.

To make sense of Eq. (3.19) one must interpret the meaning of the determinant, $\det A$. When the variables entering the Gaussian integral were discrete, the latter simply represented the determinant of the (real symmetric) matrix. In the present case, one must interpret A as an Hermitian operator having an infinite set of eigenvalues. The determinant simply represents the product over this infinite set (see, e.g., section 3.3.1). This completes our discussion of the method of Gaussian integration. Although, in the following section, we will employ only a few of the integral identities above, later we will have occasion to draw on the properties of the 'field averages'.

Before turning to specific applications of the Feynman path integral, let us stay with the general structure of the formalism and identify two fundamental connections of the path integral to *classical point mechanics* and *classical and quantum statistical mechanics.*

3.2.1 Path Integral and Statistical Mechanics

The path integral reveals a connection between quantum mechanics and classical (and quantum) statistical mechanics whose importance to all areas of field theory and statistical physics can hardly be exaggerated. To reveal this link, let us for a moment forget about quantum mechanics and consider, by way of an example, a perfectly classical, one–dimensional continuum model describing a 'flexible string'. We assume that our string is held under constant tension, and confined to a 'gutter–like potential' (as shown in Fig. 3.2). For simplicity, we also assume that the mass density of the string is pretty high, so that its fluctuations are 'asymptotically slow' (the kinetic contribution to its energy is negligible). Transverse fluctuations of the string are then penalised by its line tension, and by the external potential.

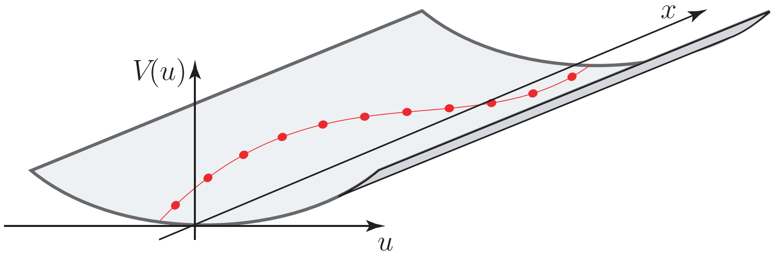


Figure 3.2: A string held under tension and confined to a potential well V.

Assuming that the transverse displacement of the string $u(x)$ is small, the potential energy stored in the string separates into two parts. The first arises form the line tension stored in the string, and the second comes from the external potential. Starting with the former, a transverse fluctuation of a line segment of length dx by an amount du, leads to a potential energy of magnitude $\delta V_{\text{tension}} = \sigma[(dx^2+du^2)^{1/2} - -dx] \simeq \sigma dx(\partial_x u)^2/2$, where σ denotes the tension. Integrated over the length of the string, one obtains $V_{\text{tension}}[\partial_x u] \equiv \int \delta V_{\text{tension}} = \frac{1}{2}\int_0^L dx\ \sigma(\partial_x u(x))^2$. The second contribution arising from the external potential is given by $V_{\text{external}}[u] \equiv \int_0^L dx\ V(u(x))$. Adding the two contributions, we find that the total energy of the string is given by $V = V_{\text{tension}} + V_{\text{external}} = \int_0^L dx[\frac{\sigma}{2}(\partial_x u)^2 + V(u)]$.

According to the general prinicples of statistical mechanics, the equilibrium properties of a system are encoded in the partition function $\mathcal{Z} = \text{tr}\left[e^{-\beta V}\right]$, where 'tr' denotes a summation over all possible configurations of the system and V is the total potential energy functional. Applied to the present case, $\text{tr} \to \int Du$, where $\int Du$ stands for the functional integration over all configurations of the string $u(x)$, $x \in [0, L]$. Thus, the partition function of the string is given by

$$\mathcal{Z} = \int Du\ \exp\left[-\beta\int_0^L dx\left(\frac{\sigma}{2}(\partial_x u)^2 + V(u)\right)\right]. \tag{3.22}$$

A comparison of this result with Eq. (3.8) shows that the partition function of the *classical* system coincides with the *quantum mechanical* amplitude

$$\boxed{\mathcal{Z} = \int dq\ \langle q|e^{i\hat{S}[q]/\hbar}|q\rangle\Big|_{t=-iL}}$$

evaluated at an imaginary 'time' $t \to -i\tau \equiv -iL$, where $\hat{H} = \hat{p}^2/2\sigma + V(q)$, and Planck's constant is identified with the 'temperature', $\hbar = 1/\beta$. (Here we have assumed that our string is subject to periodic boundary conditions.)

To see this explicitly, let us assume that we had reason to consider quantum propagation in imaginary time, i.e. $e^{-it\hat{H}/\hbar} \to e^{-\tau\hat{H}/\hbar}$, or $t \to -i\tau$. Assuming convergence (i.e. positivity of the eigenvalues of $\hat{H}$), a construction scheme perfectly analogous to the one outlined in section 3.1 would have led to a path integral formula of the structure (3.8). Formally, the only differece would be that (a) the Lagrangian would be integrated along the imaginary time axis $t' \to -i\tau' \in [0, -i\tau]$ and (b) that there would be a change of the sign of the kinetic energy term, viz. $(\partial_{t'} q)^2 \to -(\partial_{\tau'} q)^2$. After a suitable exchange of variables, $\tau \to L$, $\hbar \to 1/\beta$, the coincidence of the resulting expression with the partition function (3.22) is clear.

The connection between quantum mechanics and classical statistical mechanics outlined above generalises to **higher dimensions**: There are close analogies between quantum field theories in d dimensions and classical statistical mechanics in $d+1$. (The equality of the path integral above with the one–dimensional statistical model is merely the $d = 0$ version of this connection.) In fact, this connection turned out to be one of the major driving forces behind the success of path integral techniques in modern field theory/statistical mechanics. It offered, for the first time, a possibility to draw connections between systems which had seemed unrelated.

However, the concept of imaginary times not only provides a bridge between quantum and classical statistical mechanics, but also plays a role within a purely quantum mechanical context. Consider the *quantum* partition function of a *single particle* quantum mechanical system,

$$\boxed{\mathcal{Z} = \mathrm{tr}[e^{-\beta\hat{H}}] = \int dq\, \langle q|e^{-\beta\hat{H}}|q\rangle}$$

The partition function can be interpreted as a trace over the transition amplitude $\langle q|e^{-i\hat{H}t/\hbar}|q\rangle$ evaluated at an imaginary time $t = -i\hbar\beta$. Thus, real time dynamics and quantum statistical mechanics can be treated on the same footing, provided that we allow for the appearance of imaginary times.

Later we will see that the concept of imaginary or even generalized complex times plays an important role in all of field theory. There is even some nomenclature regarding imaginary times. The transformation $t \to -i\tau$ is denoted as a **Wick rotation** (alluding to the fact that a multiplication with the imaginary unit can be interpreted as a $\pi/2$–rotation in the complex plane). Imaginary time representations of Lagrangian actions are termed **Euclidean**, whereas the real time forms are called **Minkowski**[11] **actions**.

▷ INFO. The origin of this terminology can be understood by considering the structure of the action of, say, the phonon model (1.2). Forgetting for a moment about the magnitude of the

[11] **Hermann Minkowski** 1864–1909: A pure mathematician credited with the development of a four–dimensional treatment of electrodynamics and, separately, the geometry of numbers.

coupling constants, we see that the action has the bilinear structure $\sim x_\mu g^{\mu\nu} x_\nu$, where $\mu = 0, 1$, the vector $x_\mu = \partial_\mu \phi$ and the diagonal matrix $g = \mathrm{diag}(-1, 1)$ is the two dimensional version of a Minkowski metric. (In three spatial dimensions, g would take the form of the standard Minkowski metric of special relativity.) Wick rotating time, the -1 in the metric changes sign and g becomes a positive definite Euclidean metric. The nature of this transformation motivates the notation above.

Once one has grown accustomed to the idea that the interpretation of time as an imaginary quantity can be useful, yet more general concepts can be conceived. For example, one may contemplate quantum propagation along temporal contours that are neither purely real nor purely imaginary but rather are *generally complex*. Indeed, it has turned out that path integrals with curvelinear integration contours in the complex 'time plane' find numerous applications in statistical and quantum field theory.

3.2.2 Semiclassics from the Path Integral

In deriving the two path integral representations (3.6) and (3.8) no approximations were made. Yet the vast majority of quantum mechanical problems cannot be solved in closed form, and it would be hoping for too much to expect that within the path integral approach this situation would be any different. In fact no more than the path integrals of problems with a quadratic Hamiltonian — corresponding to the quantum mechanical harmonic oscillator and generalisations thereof — can be carried out in closed form. Yet what counts more than the (rare) availability of exact solutions is the flexibility with which *approximation* schemes can be developed. As for the path integral formulation, it is particularly strong in cases where **semiclassical limits of quantum theories** are explored. Here, by 'semiclassical', we mean the limit $\hbar \to 0$, i.e. the case where the theory is expected to be largely governed by classical structures with quantum fluctuations superimposed.

To see more formally how classical structures enter the path integral approach, let us explore Eqs. (3.6) and (3.8) in the limit of small $\hbar$. In this case the path integrals are dominated by path configurations with stationary action. (Non–stationary contributions to the integral imply massive phase fluctuations which largely average to zero.) Now, since the exponents of the two path integrals (3.6) and (3.8) involve the classical action functionals in their Hamiltonian respectively Lagrangian form, the extremal path configurations are simply the solutions of the classical equations of motion, viz.

$$\begin{aligned} \text{Hamiltonian}: \quad & \delta S[x] = 0 \quad \Rightarrow d_t x = \{H(x), x\} \equiv \partial_p H \partial_q x - \partial_q H \partial_p x, \\ \text{Lagrangian}: \quad & \delta S[q] = 0 \quad \Rightarrow (d_t \partial_{\dot q} - \partial_q)\, L(q, \dot q) = 0\,. \end{aligned} \tag{3.23}$$

These equations are to be solved subject to the boundary conditions $q(0) = q_I$ and $q(t) = q_F$. (Note that these boundary conditions do not uniquely specify a solution, i.e. in general there are many solutions to the equations (3.23). As an exercise, one may try to invent examples!)

Now the very fact that the stationary phase configurations are classical does not imply that quantum mechanics has disappeared completely. As with saddle–point approximations in general, it is not just the saddle–point itself that matters but also the fluctuations around it. At least it is necessary to integrate out Gaussian (quadratic) fluctuations around the point of stationary phase. In the case of the path integral, fluctuations of the action around the stationary phase configurations involve non–classical (in that they do not solve the classical equations of motion) trajectories through phase or coordinate space. Before exploring how this mechanism works in detail, let us consider the stationary phase analysis of functional integrals in general.

▷ INFO. **Stationary Phase Approximation**: Consider a general functional integral $\int Dx\, e^{-F[x]}$ where $Dx = \lim_{N\to\infty}\prod_{n=1}^{N} dx_n$ represents a functional measure resulting from taking the continuum limit of some finite dimensional integration space, and the 'action' $F[x]$ may be an arbitrary complex functional of x (leading to convergence of the integral). The function resulting from taking the limit of infinitely many discretisation points, $\{x_n\}$ is denoted by $x : t \mapsto x(t)$ (where t plays the role of the formerly discrete index n). Evaluating the integral above within a stationary phase approximation amounts to performing the following steps:

1. Firstly, find the 'points' of stationary phase, i.e. configurations $\bar{x}$ qualified by the condition of vanishing functional derivative,
$$D_x F = 0 \Leftrightarrow \forall t : \left.\frac{\delta F[x]}{\delta x(t)}\right|_{x=\bar{x}} = 0.$$
Although there may, in principle, be one or many solutions, for clarity, we first discuss the case in which the stationary phase configuration $\bar{x}$ is unique.

2. Secondly, Taylor expand the functional to second order around $\bar{x}$, viz.
$$F[x] = F[\bar{x}+y] = F[\bar{x}] + \frac{1}{2}\int dt \int dt' y(t')A(t,t')y(t) + \dots \tag{3.24}$$
where $A(t,t') = \left.\frac{\delta^2 F[x]}{\delta x(t)\delta x(t')}\right|_{x=\bar{x}}$ denotes the second functional derivative. Due to the stationarity of $\bar{x}$, one may note that no first order contribution can appear.

3. Thirdly, check that the *operator* $\hat{A} \equiv \{A(t,t')\}$ is positive definite. If it is not, there is a problem — the integration over the Gaussian fluctuations y below diverges. (In practice, where the analysis is rooted in a physical context, such eventualities arise only rarely. In situations were problems do occur, the resolution can usually be found in a judicious rotation of the integration contour.) For positive definite $\hat{A}$, however, the functional integral over y can be performed after which one obtains $\int Dx\, e^{-F[x]} \simeq e^{-F[\bar{x}]}\det(\frac{\hat{A}}{2\pi})^{-1/2}$, (cf. the discussion of Gaussian integrals above and, in particular, Eq. (3.19)).

4. Finally, if there are many stationary phase configurations, $\bar{x}_i$, the individual contributions have to be added:
$$\int Dx\, e^{-F[x]} \simeq \sum_i e^{-F[\bar{x}_i]} \det\left(\frac{\hat{A}_i}{2\pi}\right)^{-1/2}. \tag{3.25}$$

Eq. (3.25) represents the most general form of the stationary phase evaluation of a (real) functional integral.

▷ EXERCISE. Applied to the Gamma function, $\Gamma(z+1) = \int_0^\infty dx\, x^z e^{-x}$, with z complex, show that the stationary phase approximation is consistent with Stirling's approximation, viz. $\Gamma(s+1) = \sqrt{2\pi s} e^{s(\ln s - 1)}$.

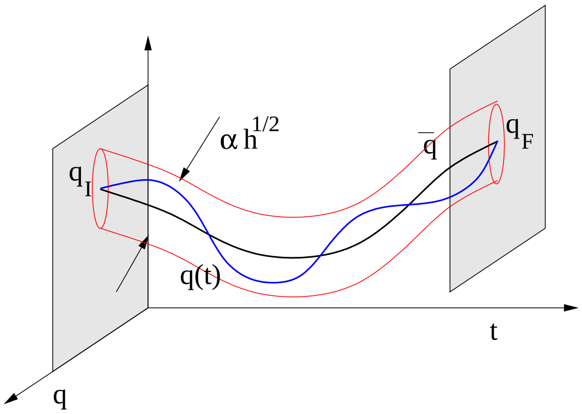


Figure 3.3: Quantum fluctuations around a classical path in coordinate space (here we assume a set of two–dimensional coordinates). Non–classical paths q fluctuating around a classical solution $q_{\rm cl}$ typically extend a distance $O(h^{1/2})$. *All* paths begin and end at q_I and q_F, respectively.

Applied to the Lagrangian form of the Feynman path integral, this program can be implemented directly. In this case, the extremal field configuration $\bar{q}(t)$ is identified as the classical solution associated with the Lagrangian, i.e. $\bar{q}(t) \equiv q_{\rm cl}(t)$. Defining $r(t) = q(t) - q_{\rm cl}(t)$ as the deviation of a general path, $q(t)$, from a nearby classical path, $q_{\rm cl}(t)$ (see Fig. 3.3), and assuming for simplicity that there exists only one classical solution connecting q_I with q_F in a time t, a stationary phase analysis obtains

$$\langle q_F | e^{-i\hat{H}t/\hbar} | q_I \rangle \simeq e^{iS[q_{\rm cl}]/\hbar} \int_{r(0)=r(t)=0} Dr \exp\left[\frac{i}{2\hbar} \int_0^t dt' dt'' r(t') \left. \frac{\delta^2 S[q]}{\delta q(t') \delta q(t'')} \right|_{q=q_{\rm cl}} r(t'') \right] \quad (3.26)$$

as the Gaussian approximation to the path integral (cf. Eq. (3.24)). For free Lagrangians of the form, $L(q, \dot{q}) = m\dot{q}^2/2 - V(q)$, the second functional derivative of the action can be straightforwardly computed by means of the rules of functional differentiation formulated in chapter 1. Alternatively, one can obtain this result by simply expanding the action as a Taylor series in the deviation $r(t)$. As a result, one obtains (exercise)

$$\frac{1}{2} \int_0^t dt \int dt' r(t) \left. \frac{\delta^2 S[q]}{\delta q(t) \delta q(t')} \right|_{q=q_{\rm cl}} r(t') = -\frac{1}{2} \int dt\; r(t) \left[m\partial_t^2 + V''(q_{\rm cl}(t)) \right] r(t)\,, \quad (3.27)$$

where $V''(q_{\rm cl}(t)) \equiv \partial_q^2 V(q)|_{q=q_{\rm cl}}$ represents the ordinary (second) derivative of the potential function evaluated at $q_{\rm cl}(t)$. Thus, the Gaussian integration over r yields the square root of the determinant of the operator $m\partial_t^2 + V''(q_{\rm cl}(t))$ — interpreted as an operator acting in the space of functions $r(t)$ with boundary conditions $r(0) = r(t) = 0$. (Note that, as we are dealing with a differential operator, the issue of boundary conditions is crucial.)

▷ INFO. More generally, Gaussian integration over fluctuations around the stationary phase configuration obtains the formal expression

$$\langle q_F|e^{-i\hat{H}t/\hbar}|q_I\rangle \simeq \det\left(\frac{i}{2\pi\hbar}\frac{\partial^2 S[q_{\rm cl}]}{\partial q_I \partial q_F}\right)^{1/2} e^{\frac{i}{\hbar}S[q_{\rm cl}]}\,, \tag{3.28}$$

as the final result for the **transition amplitude evaluated in the semiclassical approximation**. (In cases where there is more than one classical solution, the individual contributions have to be added.) To derive this expression, one shows that the operator controlling the quadratic action (3.27) fulfills some differential relations which can be, again, related back to the classical action. While a detailed formulation of this calculation (see, for example, Ref. [20], page 94) is beyond the scope of the present text, the heuristic interpretation of the result is straightforward:

According to the rules of quantum mechanics $p(q_F, q_I, t) = |\langle q_F|e^{-i\hat{H}t/\hbar}|q_I\rangle|^2$ defines the probability density for a particle injected at coordinate q_I to arrive at coordinate q_F after a time t. In the semiclassical approximation, the probability density assumes the form $p(q_F, q_I, t) = |\det(\frac{1}{2\pi\hbar}\frac{\partial^2 S[q_{\rm cl}]}{\partial q_I \partial q_F})|$. We can gain some physical insight into this expression from the following consideration: For a fixed inital coordinate q_I, the final coordinate $q_F(q_I, p_I)$ becomes a function of the initial *momentum* p_I. The classical probability density $p(q_I, q_F)$ can then be related to the probabilty density $\tilde{p}(q_I, p_I)$ for a particle to leave from the initial phase space coordinate (q_I, p_I) according to

$$p(q_I, q_F)dq_I dq_F = p(q_I, q_F)\left|\det\left(\frac{\partial q_F}{\partial p_I}\right)\right| dq_I dp_I = \tilde{p}(q_I, p_I)dq_I dp_I.$$

Now, when we say that our particle actually left at the phase space coordinate (q_I, p_I), $\tilde{p}$ becomes singular at (q_I, p_I) while being zero everywhere else. In quantum mechanics, however, all we can say is that our particle was initially confined to a *Planck cell* centered around (q_I, p_I): $\tilde{p}(q_I, p_I) = 1/(2\pi\hbar)^d$. We thus conclude that $p(q_I, q_F) = |\det(\partial p_I/\partial q_F)|\hbar^{-d}$. Finally, noticing that $p_I = -\partial_{q_I} S$ we arrive at the result of the semiclassical analysis above.

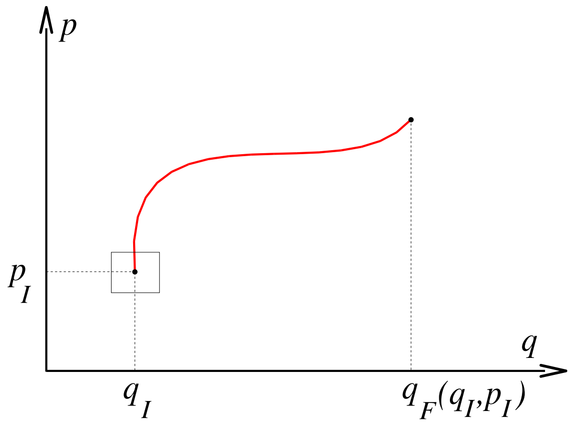


Figure 3.4: Trajectory in two-dimensional phase space: For fixed initial coordinate q_I, the final coordinate $q_F(q_I, p_I)$ becomes a function of the initial momentum. In quantum mechanics, the Planck cell $\hbar^d$ (indicated by the rectangle) limits the accuracy at which the initial coordinate can be set.

In deriving (3.28) we have restricted ourselves to the consideration of *quadratic* fluctuations around the classical paths — the essence of the semiclassical approximation. Under what conditions is this approximation justified? Unfortunately there is no rigorous and generally applicable answer to this question: For finite $\hbar$, the quality of the approximation depends largely on the sensitivity of the action to path variations. Whether or not the approximation is legitimate is a question that has to be judged from case to case. However, the *asymptotic* stability of the semiclassical approximation in the limit $\hbar \to 0$, can be deduced simply from power counting. From the structure of Eq. (3.28) it is clear that the typical magnitude of fluctuations $r(t)$ scales as $r \sim (\hbar/\delta_q^2 S)^{1/2}$, where $\delta_q^2 S$ is a symbolic shorthand for the functional variation of the action. (Variations larger than that lead to phase fluctuations $> 2\pi$, thereby being negligible.) Non–Gaussian contributions to the action would have the structure $\sim \hbar^{-1} r^n \delta_q^n S$, $n > 2$. For a typical r, this is of the order $\sim \delta_q^n S/(\delta_q^2 S)^{n/2} \hbar^{n/2-1}$. Since the S-dependent factors are classical ($\hbar$–independent), these contributions scale to zero as $\hbar \to 0$.

This concludes the conceptual part of the chapter. Before turning to the discussion of specific applications of the path integral, let us first briefly recapitulate the main steps taken in its construction:

3.2.3 Construction Recipe of the Path Integral

Consider a general quantum transition amplitude $\langle\psi|e^{-i\hat{H}t/\hbar}|\psi'\rangle$, where t may be real, purely imaginary or generally complex. To construct a functional integral representation of the amplitude:

1. Partition the time interval into $N \gg 1$ steps,
$$e^{-i\hat{H}t/\hbar} = \left[e^{-i\hat{H}\Delta t/\hbar}\right]^N, \qquad \Delta t = t/N\,.$$

2. Regroup the operator content appearing in the expansion of each factor $e^{-i\hat{H}\Delta t/\hbar}$ according to the relation
$$e^{-i\hat{H}\Delta t/\hbar} = 1 + \Delta t \sum_{mn} c_{mn} \hat{A}^m \hat{B}^n + O(\Delta t^2)\,,$$
where the eigenstates $|a\rangle, |b\rangle$ of $\hat{A}, \hat{B}$ are known and the coefficients c_{mn} are c–numbers. (In the quantum mechanical application above $\hat{A} = \hat{p}$, $\hat{B} = \hat{q}$.) This 'normal ordering' procedure emphasizes that many distinct quantum mechanical systems are associated with the same classical action.

3. Insert resolutions of identity according to
$$\begin{aligned} e^{-i\hat{H}\Delta t/\hbar} &= \sum_{a,b} |a\rangle\langle a|(1 + \Delta t \sum_{mn} c_{mn}\hat{A}^m\hat{B}^n + O(\Delta t^2))|b\rangle\langle b| \\ &= \sum_{a,b} |a\rangle\langle a|e^{-iH(a,b)\Delta t/\hbar}|b\rangle\langle b| + O(\Delta t^2)\,, \end{aligned}$$
where $H(a,b)$ is the Hamiltonian evaluated at the eigenvalues of $\hat{A}$ and $\hat{B}$.

4. Regroup terms in the exponent: Due to the 'mismatch' of the eigenstates at neighbouring time slices n and $n+1$, not only the Hamiltonians $H(a,b)$, but also sums over differences of eigenvalues appear (cf. the last term in the action (3.5)).

5. Take the continuum limit.

3.3 Applications of the Feynman Path Integral

Having introduced the general machinery of path integration we now turn to the discussion of specific applications. Our starting point will be an investigation of a low energy quantum particle confined to a single potential well, and the phenomenon of tunneling in a double well. With the latter, we will become acquainted with instanton techniques and the role of topology in field theory. The ideas developed in this section will be generalised further to the investigation of quantum mechanical decay and quantum dissipation. Finally, we will turn our attention to the development of the path integral for quantum mechanical spin and, as a case study, explore the semiclassical trace formulae for quantum chaos.

The simplest example of a quantum mechanical problem is that of a **free particle** ($\hat{H} = \hat{p}^2/2m$). Yet, within the framework of the path integral, this example, which can be dealt with straightforwardly by elementary means, is far from trivial: the Gaussian functional integral engaged in its construction involves divergences which must be regularised by rediscretising the path integral. Nevertheless, its knowledge will be useful as a means to normalise the path integral in the applications below. Therefore, we leave it as an exercise to show[12]

$$\boxed{G_{\text{free}}(q_F, q_I; t) \equiv \langle q_F | e^{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}t} | q_I \rangle \Theta(t) = \left(\frac{m}{2\pi i\hbar t}\right)^{1/2} e^{\frac{i}{\hbar}\frac{m}{2t}(q_F - q_I)^2}\Theta(t)} \tag{3.29}$$

where the Heaviside Θ–function reflects causality.[13]

▷ Exercise. Derive Eq. (3.29) by the standard methodology of quantum mechanics. Hint: insert a resolution of identity and perform a Gaussian integral.

▷ Exercise. Using the path integral, obtain a perturbative expansion for the scattering amplitude $\langle \mathbf{p}' | U(t \to \infty, t' \to -\infty) | \mathbf{p} \rangle$ of a free particle from a short–ranged central potential $V(r)$. In particular, show that the first order term in the expansion recovers the Born scattering amplitude $-i\hbar e^{-i(t-t')E(p)/\hbar} \delta(E(p) - E(p')) \langle \mathbf{p}' | V | \mathbf{p} \rangle$.

[12] Compare this result to the solution of a classical diffusion equation.

[13] Motivated by its interpretation as a Green function, in the following we will refer to the quantum transition probability amplitude by the symbol G (as opposed to U used above).

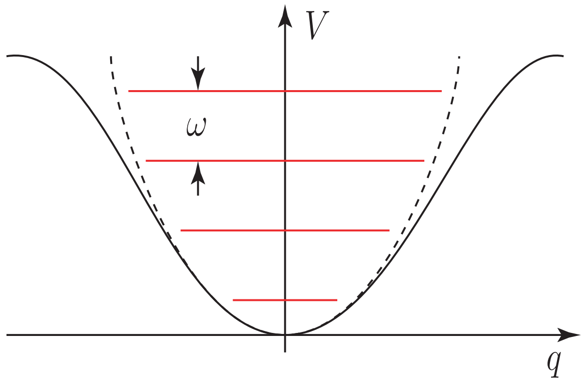


Figure 3.5: Solid: Potential well. Dashed: Quadratic fit approximating the potential shape close to the minimum.

3.3.1 Quantum Particle in a Well

As a first application of the path integral, let us consider the problem of a quantum particle in a one–dimensional potential well (see the figure). The discussion of this example will illustrate how the semiclassical evaluation scheme discussed above works in practice. For simplicity we assume the potential to be symmetric, $V(q) = V(-q)$ with $V(0) = 0$. The quantity we wish to compute is the probability amplitude that a particle injected at $q = 0$ returns after a time t, i.e. with $\hat{H} = \hat{p}^2/2m + V(\hat{q})$, $G(0,0;t) \equiv \langle q_F = 0|e^{-i\hat{H}t/\hbar}|q_I = 0\rangle\Theta(t)$. Drawing on our previous discussion, the path integral representation of the transition amplitude is given by

$$G(0,0;t) = \int_{q(t)=q(0)=0} Dq \exp\left[\frac{i}{\hbar}\int_0^t dt' L(q,\dot{q})\right],$$

where $L = m\dot{q}^2/2 - V(q)$ represents the corresponding Lagrangian.

Now, for a generic potential $V(q)$, the path integral can not be evaluated exactly. Instead, we wish to invoke the semiclassical analysis outlined conceptually above. Accordingly, we must first find solutions to the classical equation of motion. Minimising the action with respect to variations of $q(t)$, one obtains the Euler–Lagrange equation of motion $m\ddot{q} = -V'(q)$ where, as usual, we have used the shorthand $V'(q) \equiv \partial_q V(q)$. According to the Feynman path integral, this equation must be solved subject to the boundary conditions $q(t) = q(0) = 0$. One solution is obvious, viz. $q_{\rm cl}(t) = 0$. Assuming that this is in fact the only solution,[14] we obtain (cf. Eqs. (3.26) and (3.27))

$$G(0,0;t) \simeq \int_{r(0)=r(t)=0} Dr \exp\left[-\frac{i}{\hbar}\int_0^t dt' r(t')\frac{m}{2}\left(\partial_{t'}^2 + \omega^2\right) r(t')\right],$$

[14] In general, this assumption is wrong: For smooth potentials $V(q)$, a Taylor expansion of V at small q obtains the harmonic oscillator potential, $V(q) = V_0 + m\omega^2 q^2/2 + \cdots$. For times t that are commensurate with π/ω, one has periodic solutions, $q_{\rm cl}(t) \propto \sin(\omega t)$ that start out from the origin at time $t = 0$ and revisit it at just the right time t. In the next section we will see why the restriction to just the trivial solution was nonetheless legitimate (for arbitrary times t).

where, by definition, $m\omega^2 \equiv V''(0)$ is the second derivative of the potential at the origin.[15] Note that, in this case, the contribution to the action from the stationary phase field configuration vanishes $S[q_{\rm cl}] = 0$. Following the discussion of section 3.2, Gaussian functional integration over r then leads to the semiclassical expansion

$$G(0,0;t) \simeq J\det\left(-m(\partial_t^2 + \omega^2)/2\right)^{-1/2}, \tag{3.30}$$

where the prefactor J absorbs various constant prefactors.

Operator determinants are usually most conveniently obtained by presenting them as a product over eigenvalues. In the present case, the eigenvalues ϵ_n are determined by the equation

$$-\frac{m}{2}\left(\partial_t^2 + \omega^2\right) r_n = \epsilon_n r_n ,$$

which is to be solved subject to the boundary condition $r_n(t) = r_n(0) = 0$. A complete set of solutions to this equation is given by[16] $r_n(t') = \sin(n\pi t'/t)$, $n = 1, 2, \ldots$, with eigenvalues $\epsilon_n = m[(n\pi/t)^2 - \omega^2]/2$. Applied to the determinant, one therefore finds

$$\det\left(-m(\partial_t^2 + \omega^2)/2\right)^{-1/2} = \prod_{n=1}^{\infty}\left[\frac{m}{2}\left(\left(\frac{n\pi}{t}\right)^2 - \omega^2\right)\right]^{-1/2} .$$

To interpret this result, one must first make sense of the infinite product (which even seems divergent for times commensurate with π/ω!). Moreover the value of the constant J has yet to be properly determined. To resolve these difficulties, one may exploit the fact that (a) we do know the transition amplitude (3.29) of the *free* particle system, and (b) the latter coincides with the transition amplitude G in the special case where the potential $V \equiv 0$. In other words, had we computed $G_{\rm free}$ via the path integral, we would have obtained the same constant J and, more importantly, an infinite product like the one above, but with $\omega = 0$. This allows the transition amplitude to be regularised as

$$G(0,0;t) \equiv \frac{G(0,0;t)}{G_{\rm free}(0,0;t)} G_{\rm free}(0,0;t) = \prod_{n=1}^{\infty}\left[1 - \left(\frac{\omega t}{n\pi}\right)^2\right]^{-1/2} \left(\frac{m}{2\pi i\hbar t}\right)^{1/2} \Theta(t) .$$

Then, with the help of the mathematical identity $\prod_{n=1}^{\infty}[1 - (x/n\pi)^2]^{-1} = x/\sin x$, one finally arrives at the result

$$G(0,0;t) \simeq \sqrt{\frac{m\omega}{2\pi i\hbar \sin(\omega t)}}\,\Theta(t) . \tag{3.31}$$

In the case of the harmonic oscillator, the expansion of the potential necessarily truncates at quadratic order and, in this case, the expression above is exact. (For a more ranging discussion of the path integral for the quantum harmonic oscillator system, see

[15] Those who are uncomfortable with functional differentiation can arrive at the same expression simply by substituting $q(t) = q_{\rm cl}(t) + r(t)$ into the action and expanding in r.

[16] To find the solutions of this equation, recall the structure of the Schrödinger equation of a particle in a one–dimensional box of width $L = t$!

problem 3.5.) For a general potential, the semiclassical approximation effectively involves the replacement of $V(q)$ by a quadratic potential with the same curvature. The calculation above also illustrates how coordinate space fluctuations around a completely static solution may reinstate the zero–point fluctuations characteristic of quantum mechanical bound states.

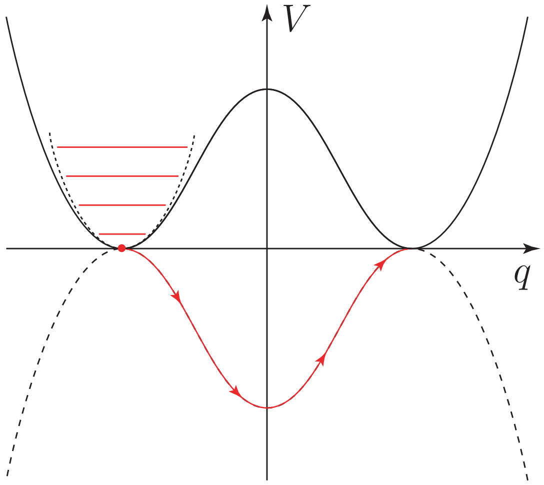


Figure 3.6: Solid: Double well potential. Dashed: Inverted potential

3.3.2 Double Well Potential: Tunneling and Instantons

As a second application of the path integral let us now consider the motion of a particle in a double well potential (see the figure). Our aim will be to estimate the quantum probability amplitude for a particle to either stay at the bottom of one of the local minima *or* to go from one minimum to the other. In doing so, it is understood that the energy range accessible to the particle (i.e. $\Delta E \sim \hbar/t$) is well below the potential barrier height, i.e. quantum mechanical transfer between minima is by *tunnelling*. Here, in contrast to the single well system, it is far from clear what kind of classical stationary phase solutions may serve as a basis for a description of quantum tunnelling; there appear to be no classical paths connecting the two minima. Of course one may think of particles 'rolling' over the potential hill. Yet, these are singular and, by assumption, energetically inaccessible.

The key to resolving these difficulties is an observation, already made above, that the time argument appearing in the path integral should be considered as a general complex quantity that can (according to convenience) be sent to any value in the complex plane. In the present case, a Wick rotation to imaginary times will reveal a stationary point of the action. At the end of the calculation, the *real* time amplitudes we seek can be straightforwardly obtained by analytic continuation.

To be specific, let us consider the imaginary time transition amplitudes

$$G_E(a, \pm a; \tau) \equiv \langle \pm a| \exp\left[-\frac{\tau}{\hbar}\hat{H}\right] |a\rangle = G(-a, \mp a; \tau) \tag{3.32}$$

where the coordinates $\pm a$ coincide with the two minima of the potential. From (3.32) the real time amplitudes $G(a, \pm a; t)$ can be recovered by the analytic continuation $\tau \to it$. According to section 3.2.1, the **Euclidean path integral** formulation of the transition amplitudes is given by

$$\boxed{G(a, \pm a; \tau) = \int_{q(0)=\pm a, q(\tau)=a} Dq \; \exp\left[-\frac{1}{\hbar}\int_0^\tau d\tau' \left(\frac{m}{2}\dot{q}^2 + V(q)\right)\right]} \tag{3.33}$$

where the function q now depends on imaginary time. From (3.33) we obtain the stationary phase (or saddle–point) equations

$$-m\ddot{q} + V'(q) = 0. \tag{3.34}$$

From this result, one can infer that, as a consequence of the Wick rotation, there is an effective *inversion* of the potential, $V \to -V$ (shown dashed in the figure above). The crucial point is that, within the inverted potential landscape, the barrier has become a sink, i.e. within the new formulation, there *are* classical solutions connecting the two points, $\pm a$. More precisely, there are three different types of classical solutions which fulfill the condition to be at coordinates $\pm a$ at times 0 and/or τ: (a) The solution wherein the particle rests permanently at a,[17] (b) the corresponding solution staying at $-a$ and, most importantly, (c) the solution in which the particle leaves its initial position at $\pm a$, accelerates through the minimum at 0 and eventually reaches the final position $\mp a$ at time τ. In computing the transition amplitudes, all three types of paths have to be taken into account. As for (a) and (b), by computing quantum fluctuations around these solutions, one can recover the physics of the zero–point motion described in section 3.3.1 for each well individually (exercise: convince yourself that this is true!). Now let us see what happens if the paths connecting the two coordinates are added to this picture.

The Instanton Gas

The classical solution of the Euclidean equation of motion that connects the two potential maxima is called an **instanton solution** while a solution traversing the same path but in the opposite direction (‘$-a \to a$’ $\rightsquigarrow$ ‘$a \to -a$’) is called an anti–instanton. The name ‘instanton’ was invented by ’t Hooft[18] with the idea that these objects are very similar in

[17] Note that the potential inversion answers a question that arose above, i.e. whether or not the classical solution staying at the bottom of the single well was actually the only one to be considered. As with the double well, we could have treated the single well within an imaginary time representation, whereupon the well would have become a hill. Clearly there is just one classical solution being at two different times at the top of the hill, viz. the solution that stays there forever. By formulating the semiclassical expansion around that path, we would have obtained (3.31) with $t \to -i\tau$, which, upon analytic continuation, would have led back to the real time result.

[18]

Gerardus ’t Hooft 1946– : 1999 Nobel Laureate in Physics for elucidating the quantum structure of electroweak interactions in physics.

their mathematical structure to 'solitons', particle–like solutions of classical field theories. However, unlike solitons, they are structures in time (albeit Euclidean time); thus the 'instant–'. As another etymographic remark, note that the syllable '–on' in 'instanton' hints to an interpretation of these states as a kind of particle. The background is that, as a function of the time coordinate, instantons are almost everywhere constant save for a short region of variation (see below). Alluding to the interpretation of time as something akin to a spatial dimension, these states can be interpreted as a well–localised excitation or, according to standard field theoretical practice, a *particle*.[19]

To proceed, we must first compute the classical action associated with a single instanton solution. Multiplying (3.34) by $\dot{q}_{\rm cl}$, integrating over time (i.e. performing the first integral of the equation of motion), and using the fact that at $q_{\rm cl} = \pm a$, $\dot{q}_{\rm cl} = 0$ and $V = 0$, one finds that

$$\frac{m}{2}\dot{q}_{\rm cl}^2 = V(q_{\rm cl}). \tag{3.35}$$

With this result, one obtains the instanton action

$$S_{\rm inst} = \int_0^\tau d\tau' \, \overbrace{\left(\frac{m}{2}\dot{q}_{\rm cl}^2 + V(q_{\rm cl})\right)}^{m\dot{q}_{\rm cl}^2} = \int d\tau' \, \frac{dq_{\rm cl}}{d\tau'}(m\dot{q}_{\rm cl}) = \int_{-a}^{a} dq \, (2mV(q))^{1/2}. \tag{3.36}$$

Notice that $S_{\rm inst}$ is determined solely by the functional profile of the potential V (i.e. does not depend on the structure of the solution $q_{\rm cl}$).

Secondly, let us explore the structure of the instanton as a function of time. Defining the second derivative of the potential at $\pm a$ by $V''(\pm a) = m\omega^2$, Eq. (3.35) implies that for large times (where the particle is close to the right maximum), $\dot{q}_{\rm cl} = -\omega(q_{\rm cl} - a)$ which integrates to $q_{\rm cl}(\tau) \stackrel{\tau\to\infty}{\longrightarrow} a - e^{-\tau\omega}$. Thus the temporal extension of the instanton is set by the oscillator frequencies of the local potential minima (the maxima of the inverted potential) and, in cases where tunnelling takes place on time scales much larger than that, can be regarded as short (see Fig. 3.7).

The confinement of the instanton configuration to a narrow interval of time has an important implication — there must exist *approximate* solutions of the stationary equation involving further anti–instanton/instanton pairs (physically, the particle repeatedly bouncing to and fro in the inverted potential). According to the general philosophy of the saddle–point scheme, the path integral is obtained by summing over all solutions of the saddle–point equations and hence over all instanton configurations. The summation over multi–instanton configurations — termed the '**instanton gas**' — is substantially simplified by the fact that individual instantons have short temporal support (events of overlapping configurations are rare) and that not too many instantons can be accommodated in a finite time interval (the instanton gas is dilute). The actual density is dictated by the competition between the configurational 'entropy' (favouring high density), and

[19] In addition to the original literature, the importance that has been attached to the instanton method has inspired a variety of excellent and pedagogical reviews of the field. Of these, the following are highly recommended: A. M. Polyakov, *Quark confinement and topology of gauge theories*, Nucl. Phys. **B120**, 429 (1977); S. Coleman, in *Aspects of symmetry — selected Erice lectures*, (Cambridge University Press 1985) chapter 7.

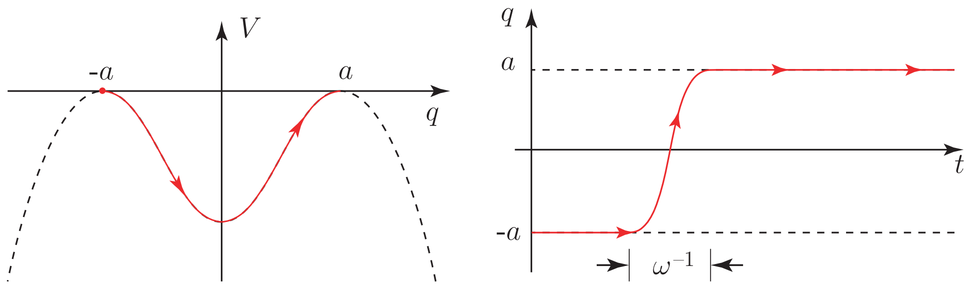


Figure 3.7: Single instanton configuration.

the 'energetics', the exponential weight implied by the action (favouring low density) — see the estimate below.

In practice, multi–instanton configurations imply a transition amplitude

$$G(a, \pm a; \tau) \simeq \sum_{n \text{ even / odd}} K^n \int_0^\tau d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{n-1}} d\tau_n A_n(\tau_1, \ldots, \tau_n), \tag{3.37}$$

where A_n denotes the amplitude associated with n instantons, and we have taken into account the fact that in order to connect a with $\pm a$, the number of instantons must be even/odd. The n instanton bounces contributing to each A_n can take place at arbitrary times $\tau_i \in [0, \tau]$, $i = 1, \ldots, n$ and all these possibilities have to be added (i.e. integrated). Here K denotes a (dimensionful) constant absorbing the temporal dimension $[\text{time}]^n$ introduced by the time integrations, and $A_n(\tau_1, \ldots, \tau_n)$ is the transition amplitude, evaluated within the semiclassical approximation around a configuration of n instanton bounces at times $0 \le \tau_1 < \tau_2 < \ldots < \tau_n \le \tau$ (see Fig. 3.8). In the following, we will first focus on the transition amplitude A_n which controls the exponential dependence of the tunneling amplitude returning later to consider the prefactor K.

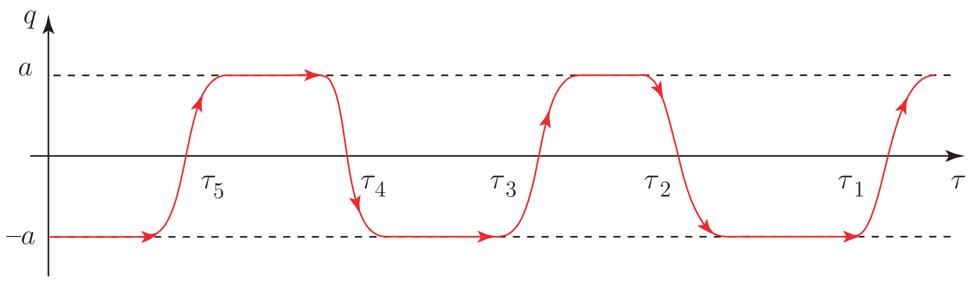


Figure 3.8: Dilute instanton gas configuration.

According to the general semiclassical principle, each amplitude $A_n = A_{n,\text{cl.}} \times A_{n,\text{qu.}}$ factorises into two parts, a classical contribution $A_{n,\text{cl.}}$ accounting for the action of the instanton configuration; and a quantum contribution $A_{n,\text{qu.}}$ resulting from quadratic fluctuations around the classical path. Focusing initially on $A_{n,\text{cl.}}$ we note that, at intermediate times, $\tau_i \ll \tau' \ll \tau_{i+1}$, where the particle rests on top of either of the maxima at $\pm a$, no action accumulates (cf. the previous section). However, each instanton bounce has a finite action S_{inst} (see Eq. (3.36)) and these contributions add up to give the full classical action,

$$A_{n,\text{cl.}}(\tau_1,\ldots,\tau_n) = e^{-nS_{\text{inst}}/\hbar}, \tag{3.38}$$

which is independent of the time coordinates τ_i. (The individual instantons 'don't know of each other'; their action is independent of their relative position.)

As for the quantum factor $A_{n,\text{qu.}}$, there are, in principle, two contributions. Whilst the particle rests on either of the hills (the straight segments in Fig. 3.8), quadratic fluctuations around the classical (i.e. spatially constant) configuration play the same role as the quantum fluctuations considered in the previous section, the only difference being that we are working in a Wick rotated picture. There it was found that quantum fluctuations around a classical configuration which stays for a (real) time t at the bottom of the well, result in a factor $\sqrt{1/\sin(\omega t)}$ (the remaining constants being absorbed into the prefactor K^n). Rotating to imaginary times, $t \to -i\tau$, one can infer that the quantum fluctuation accumulated during the stationary time $\tau_{i+1}-\tau_i$ is given by

$$\sqrt{\frac{1}{\sin(-i\omega(\tau_{i+1}-\tau_i))}} \sim e^{-\omega(\tau_{i+1}-\tau_i)/2},$$

where we have used the fact that, for the dilute configuration, the typical separation times between bounces are much larger than the inverse of the characteristic oscillator scales of each of the minima. (It takes the particle much longer to tunnel through a high barrier than to oscillate in either of the wells of the *real* potential.)

Now, in principle, there are also fluctuations around the 'bouncing' segments of the path. However, due to the fact that a bounce takes a time of $O(\omega^{-1}) \ll \Delta\tau$, where $\Delta\tau$ represents the typical time *between* bounces, one can neglect these contributions (which is to say that they can be absorbed into the prefactor K without explicit calculation). Within this approximation, setting $\tau_0 \equiv 0$, $\tau_{n+1} \equiv \tau$, the overall quantum fluctuation correction is given by

$$A_{n,\text{qu.}}(\tau_1,\ldots,\tau_n) = \prod_{i=0}^{n} e^{-\omega(\tau_{i+1}-\tau_i)/2} = e^{-\omega\tau/2}, \tag{3.39}$$

again independent of the particular spacial configuration $\{\tau_i\}$. Combining (3.38) and (3.39), one finds that

$$\begin{aligned} G(a,\pm a;\tau) &\simeq \sum_{n\ \text{even}\,/\,\text{odd}}^{\infty} K^n e^{-nS_{\text{inst}}/\hbar} e^{-\omega\tau/2} \overbrace{\int_0^{\tau} d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{n-1}} d\tau_n}^{\tau^n/n!} \\ &= e^{-\omega\tau/2} \sum_{n\ \text{even}\,/\,\text{odd}} \frac{1}{n!}\left(\tau K e^{-S_{\text{inst}}/\hbar}\right)^n. \end{aligned} \tag{3.40}$$

Finally, performing the summation, one obtains the transition amplitude

$$G(a,\pm a;\tau) \simeq C e^{-\omega\tau/2} \begin{cases} \cosh\left(\tau K e^{-S_{\text{inst.}}/\hbar}\right) \\ \sinh\left(\tau K e^{-S_{\text{inst.}}/\hbar}\right) \end{cases} \tag{3.41}$$

where C is some factor that depends in a non–exponential way on the transition time.

Before we turn to a discussion of the physical content of this result, let us check the self–consistency of our central working hypothesis — the diluteness of the instanton gas. To this end, consider the representation of G in terms of the partial amplitudes (3.40). To determine the typical number of instantons contributing to the sum, one may make use of the fact that, for a general sum $\sum_n c_n$ of positive quantities $c_n > 0$, the 'typical' value of the summation index can be estimated as $\langle n\rangle \equiv \sum_n c_n n / \sum_n c_n$. With the abbreviation $X \equiv \tau K e^{-S_{\text{inst}}/\hbar}$, the application of this estimate to our current sum yields

$$\langle n\rangle \equiv \frac{\sum_n n X^n/n!}{\sum_n X^n/n!} = X,$$

where we have used the fact that, as long as $\langle n\rangle \gg 1$, the even/odd distinction in the sum is irrelevant. Thus, we can infer that the average instanton density, $\langle n\rangle/\tau = K e^{-S_{\text{inst}}/\hbar}$ is both exponentially small in the instanton action S_{inst}, and independent of τ confirming the validity of our diluteness assumptions above.

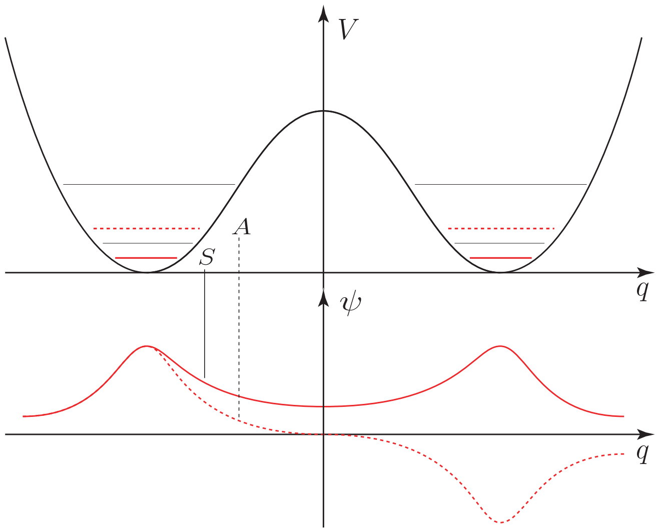


Figure 3.9: Quantum states in the double well: Dashed: Harmonic oscillator states. Solid: Exact eigenstates.

Finally, let us discuss how the form of the transition amplitude (3.41) can be understood in physical terms. To this end, let us reconsider the basic structure of the problem we are dealing with (see Fig. 3.9). While there is no coupling across the barrier, the Hamiltonian has two independent, oscillator–like sets of low lying eigenstates sitting in the two local minima. Allowing for a weak inter–barrier coupling, the oscillator ground states (like all higher states) split into a doublet of a symmetric and an antisymmetric eigenstate, $|S\rangle$ and $|A\rangle$ with energies ϵ_A and ϵ_S, respectively. Focusing on the low energy sector formed by the ground state doublet, we can express the transition amplitudes (3.32) as

$$G(a, \pm a; \tau) \simeq \langle a| \Big(|S\rangle e^{-\epsilon_S \tau/\hbar} \langle S| + |A\rangle e^{-\epsilon_A \tau/\hbar} \langle A| \Big) | \pm a\rangle.$$

Setting $\epsilon_{A/S} = \hbar\omega/2 \pm \Delta\epsilon/2$, where $\Delta\epsilon$ represents the tunnel-splitting, the symmetry properties $|\langle a|S\rangle|^2 = |\langle -a|S\rangle|^2 = C/2$ and $\langle a|A\rangle\langle A|-a\rangle = -|\langle a|A\rangle|^2 = -C/2$ imply that

$$G(a, \pm a; \tau) \simeq \frac{C}{2}\left(e^{-(\hbar\omega-\Delta\epsilon)\tau/2\hbar} \pm e^{-(\hbar\omega+\Delta\epsilon)\tau/2\hbar}\right) = Ce^{-\omega\tau/2}\left\{\begin{array}{l}\cosh(\Delta\epsilon\tau/\hbar)\\ \sinh(\Delta\epsilon\tau/\hbar)\end{array}\right. .$$

Comparing this expression with Eq. (3.41) the interpretation of the instanton calculation becomes clear: At long times, the transition amplitude engages the two lowest states — the symmetric and anti–symmetric combination of the two oscillator ground states. The energy splitting $\Delta\epsilon$ accommodates the energy shift due to the tunneling between the two wells. Remarkably, the effect of tunneling was obtained from a purely classical picture (formulated in imaginary time!). The instanton calculation also produced a prediction for the tunnel splitting of the energies, viz.

$$\Delta\epsilon = \hbar K \exp(-S_{\rm inst}/\hbar),$$

which, up to the prefactor, agrees with the result of a WKB–type analysis of the tunnel process.

Before leaving this section, two general remarks on instantons are in order:

- ▷ In hindsight, was the approximation scheme used above consistent? In particular, terms at second order in $\hbar$ were neglected, while terms non–perturbative in $\hbar$ (the instanton) were kept. Yet, the former typically give rise to a larger correction to the energy than the latter. However, the large perturbative shift effects the energies of the symmetric and antisymmetric state equally. The instanton contribution gives the *leading* correction to the splitting of the levels. It is the latter which is likely to be of more physical significance.

- ▷ Secondly, it may — legitimately — appear as though the development of the machinery above was a bit of an "overkill" for describing a simple tunnelling process. As a matter of fact, the basic result (3.41) could have been obtained in a simpler way by more elementary means (using, for example, the WKB method). Why then did we discuss instantons at such length? One reason is that, even within a purely quantum mechanical framework, the instanton formulation of tunnelling is much stronger than WKB. The latter represents, by and large, an uncontrolled approximation. In general it is hard to tell whether WKB results are accurate or not. In contrast, the instanton approximation to the path integral is controlled by a number of well–defined expansion parameters. For example, by going beyond the semiclassical approximation and/or softening the diluteness assumption, the calculation of the transition amplitudes can, in principle, be driven to arbitrary accuracy.

- ▷ A second, and for our purposes, more important motivation is that instanton techniques are of crucial importance within higher dimensional field theories (here we regard the path integral formulation of quantum mechanics as a 0 space +1 time = 1–dimensional field theory). The reason is that instantons are intrinsically non–perturbative objects, which is to say that instanton solutions to stationary phase equations describe a type of physics that cannot be obtained by a perturbative expansion around a non–instanton sector of the theory. (For example, the bouncing orbits in the example above cannot be incorporated into the analysis by doing a kind of perturbative expansion around a trivial orbit.) This non–perturbative nature of instantons can be understood by topological reasoning:

Relatedly, one of the features of the instanton analysis above was that the *number* of instantons involved was a stable quantity; 'stable' in the sense that by including perturbative fluctuations around the n instanton sector, say, one does not connect with the $n+2$ sector. Although no rigorous proof of this statement has been given, it should be heuristically clear: a trajectory involving n bounces between the hills of the inverted potential cannot be smoothly connected with one of a different number. Suppose for instance we would forcably attempt to interpolate between two paths with different bounce numbers: Inevitably, some of the intermediate configurations would be charged with actions that are far apart from any stationary phase like value. Thus, the different instanton sectors are separated by an energetic barrier that cannot be penetrated by smooth interpolation and, in this sense, they are **topologically distinct**.

▷ Info. **Fluctuation determinant**: Our analysis above provided a method to extract the tunneling rate between the quantum wells to a level of exponential accuracy. However, in some applications, it is useful to compute the exponential prefactor K. Although such a computation follows the general principles outlined above and implemented explicitly for the single well, there are some idiosyncracies in the tunneling system which warrant discussion.

According to the general principles outlined in section 3.2.2, integrating over Gaussian fluctuations around the saddle–point field configurations, the contribution to the transition amplitude from the n–instanton section is given by

$$G_n = J\det\left(-m\partial_\tau^2 + V''(q_{\text{cl},n})\right) e^{-nS_{\text{inst.}}/\hbar}$$

where $q_{\text{cl},n}(\tau)$ represents an n–instanton configuration and J the normalisation. Now, in the zero instanton sector, the evaluation of the functional determinant recovers the familiar harmonic oscillator result, $G(0,0;\tau) = (m\omega/\pi\hbar)^{1/2}\exp[-\omega\tau/2]$. Let us now consider the one instanton sector of the theory. To evaluate the functional determinant, one must consider the spectrum of the operator $-m\partial_\tau^2 + V''(q_{\text{cl},1})$. Differentiating the defining equation for $q_{\text{cl},1}$ (3.34), one may confirm that

$$\left(-m\partial_\tau^2 + V''(q_{\text{cl},1})\right)\partial_\tau q_{\text{cl},1} = 0,$$

i.e. the function $\partial_\tau q_{\text{cl},1}$ presents a zero mode of the operator!. Physically, the origin of the zero mode is elucidated by noting that a translation of the instanton along the time axis, $q_{\text{cl},1}(\tau) \to q_{\text{cl},1}(\tau+\delta\tau)$ should leave the action approximately invariant. However, for small $\delta\tau$, $q_{\text{cl},1}(\tau+\delta\tau) \simeq q_{\text{cl},1}(\tau)+\delta\tau\partial_\tau q_{\text{cl},1}$, i.e. to first order, the addition of the increment function $\partial_\tau q_{\text{cl},1}$ leaves the action invariant, and $\delta\tau$ is a 'zero mode coordinate'.

With this interpretation, it becomes clear how to repair the formula for the fluctuation determinant. While the Gaussian integral over fluctuations is controlled for the non–zero eigenvalues, its execution for the zero mode must be rethought. Indeed, by integrating over the coodinate of the instanton, viz. $\int_0^\tau d\tau_0 = \tau$, one finds that the contribution to the transition amplitude in the one instanton sector is given by

$$J\tau\sqrt{\frac{S_{\text{inst.}}}{2\pi\hbar}}\det{}'\left[-m\partial_\tau^2 + V''(q_{\text{cl},1})\right]^{-1/2} e^{-S_{\text{inst.}}/\hbar}$$

where the prime indicates the exclusion of the zero mode from the determinant, and the factor $\sqrt{S_{\text{inst.}}/2\pi\hbar}$ reflects the Jacobian associated with the change to a new set of integration variables which contains the zero mode coordinate τ as one of its elements.[20] To fix the, as yet, undetermined coupling constant J, we normalize by the fluctuation determinant of the (imaginary time) harmonic oscillator, i.e. we use the fact that (cf. section 3.3.1), for the harmonic oscillator, the return amplitude evaluates to $G(0,0,\tau) = J\det(m(-\partial_\tau^2+\omega^2)/2)^{-1/2} = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-\omega\tau/2}$, where the first/second representation is the imaginary time variant of Eq. (3.30)/Eq.(3.31). Using this result, and noting that the zero mode analysis above generalizes to the n–instanton sector, we find that the pre–exponential constant K used in our analysis of the double well problem above affords the explicit representation

$$K = \omega\sqrt{\frac{S_{\text{inst.}}}{2\pi\hbar}}\left[\frac{m\omega^2\det'\left[-m\partial_\tau^2+V''(q_{\text{cl},1})\right]}{\det\left[-m\partial_\tau^2+m\omega^2\right]}\right]^{-1/2}.$$

Naturally, the instanton determinant depends sensitively on the particular nature of the potential $V(q)$. For the quartic potential $V(q) = m\omega^2(x^2-a^2)^2/8a^2$, it may be confirmed that the

$$\frac{m\omega^2\det'\left[-m\partial_\tau^2+V''(q_{\text{cl},1})\right]}{\det\left[-m\partial_\tau^2+m\omega^2\right]} = \frac{1}{12},$$

while $S_{\text{inst}} = \sqrt{2/3}\,m\omega a^2$. For further details of the calculation, we refer to, e.g., Zinn-Justin.

Escape From a Metastable Minimum: "Bounces"

The instanton gas approximation for the double well system can be easily adapted to explore the problem of quantum mechanical tunneling from a metastable state such as that presented by an unstable nucleus. In particular, suppose one wishes to estimate the "survival probability" of a particle captured in a metastable minimum of a one–dimensional potential such as that shown in Fig. 3.10.

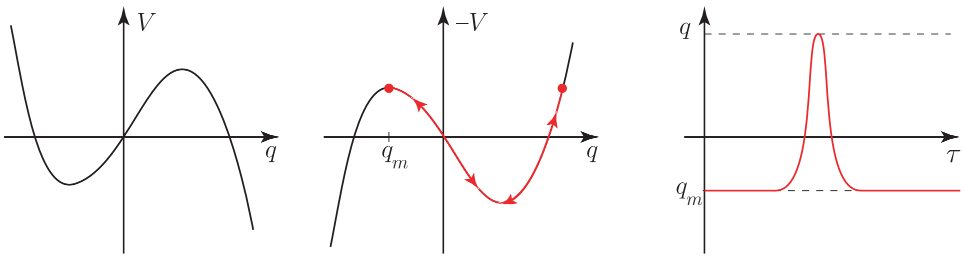


Figure 3.10: Effective potential showing a metastable minimum together with the inverted potential and a sketch of a bounce solution. To obtain the tunnelling rate it is necessary to sum over a dilute gas of bounce trajectories.

According to the path integral scheme, the survival probability, defined by the probability amplitude to remain at the potential minimum q_m, i.e. the propagator $G(q_m, q_m; t)$,

[20] For an explicit calculation of this Jacobian see, e.g., J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, 1993).

can be evaluated by making use of the Euclidean time formulation of the Feynman path integral. As with the double well, in the Euclidean time formalism, the dominant contribution to the transition probability arises from the classical path minimising the action corresponding to the inverted potential (see Fig. 3.10). However, in contrast to the double well potential, the classical solution takes the form of a '**bounce**' (i.e. the particle spends only a short time away from the potential minimum — there is only one metastable minimum of the potential). As with the double well, one can expect mutliple bounce trajectories to present a significant contribution. Summing over all bounce trajectories (note that in this case we have an exponential series — no even/odd parity effect), one obtains the survival probability

$$G(q_m, q_m; \tau) = Ce^{-\omega\tau/2} \exp\left[\tau K e^{-S_{\text{bounce}}/\hbar}\right].$$

Applying an analytic continuation to real time, one finds $G(\theta_m, \theta_m; t) = Ce^{-i\omega t/2}\exp[-\frac{\Gamma}{2}t]$, where the decay rate is given by $\Gamma/2 = |K|e^{-S_{\text{bounce}}/\hbar}$. (Note that on physical grounds we can see that K must be imaginary.[21]

▷ Exercise. Consider a heavy nucleus having a finite rate of α-decay. The nuclear forces can be considered very short-ranged so that the rate of α particle emission is controlled by tunneling under a Coulomb barrier. Taking the effective potential to be spherically symmetric with a deep minimum core of radius r_0 beyond which it decays as $U(r) = 2(Z-1)e^2/r$ where Z is the nuclear charge, find the temperature of the nuclei above which α-decay will be thermally assisted if the energy of the emitted particles is E_0. Estimate the mean energy of the α particles as a function of temperature.

▷ Exercise. A uniform electric field E is applied perpendicular to the surface of a metal with work function W. Assuming that the electrons in the metal describe a Fermi gas of density n, with exponential accuracy, find the tunneling current at zero temperature ("cold emission"). Show that, effectively, only electrons with energy near the Fermi level are tunneling. With the same accuracy, find the current at finite temperature ("hot emission"). What is the most probable energy of tunneling electrons as function of temperature?

3.3.3 †Tunneling of Quantum Fields: 'Fate of the False Vacuum'

▷ Additional Example: Hitherto we have focussed on applications of the Feynman path integral to the quantum mechanics of isolated point–like particles. In this setting, the merit of the path integral scheme over, say, standard perturbative methods or the 'WKB' approach is perhaps not compelling. Therefore, by way of motivation, let us here present an example which builds upon the structures elucidated above and which illustrates the power of the path integral method.

[21] In fact, a more careful analysis shows that this estimate of the decay rate is too large by a factor of 2 (for further details see, e.g., Coleman, *Aspects of Symmetry: Selected Erice Lectures*, CUP.)

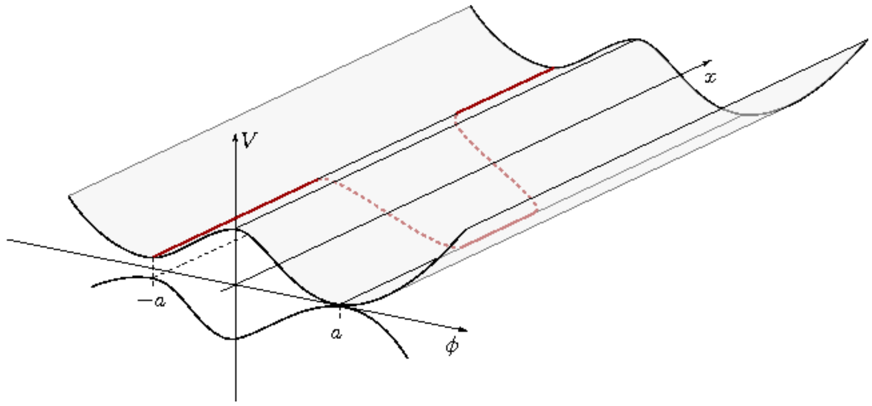


Figure 3.11: Snapshot of a field configuration $\phi(x, t = \text{const.})$ in a potential landscape with two nearly degenerate minima. For further discussion, see the text.

To this end, let us consider a theory involving a continuous classical field which can adopt two homogeneous equilibrium states with different energy densities: To be concrete, one may consider an harmonic chain confined to one or other minimum of an asymmetric quasi one–dimensional 'gutter–like' double well potential (see Fig. 3.11). When quantised, the state of higher energy density becomes unstable through barrier penetration — it is said to be a "false vacuum".[22] Specifically, drawing on our discussion of the harmonic chain in chapter 1, let us consider a quantum system specified by the Hamiltonian density

$$\hat{\mathcal{H}} = \frac{\hat{\pi}^2}{2m} + \frac{k_s a^2}{2}(\partial_x \hat{\phi})^2 + V(\hat{\phi}), \tag{3.42}$$

where $[\hat{\pi}(x), \hat{\phi}(x')] = -i\hbar\delta(x-x')$. Here we have included a potential $V(\phi)$ which, in the present case, assumes the form of a double well. The inclusion of a weak bias $-f\phi$ into $V(\phi)$ identifies a stable and a metastable potential minimum. Previously, we have seen that, in the absence of the confining potential, the quantum string exhibits low–energy collective wave–like excitations — phonons. In the confining potential, these harmonic fluctuations are rendered massive. However, drawing on the quantum mechanical principles established in the single–particle system, one might assume that the string tunnels freely between the two potential minima. To explore the capacity of the system to tunnel, let us suppose that, at some time $t = 0$, the system adopts a field configuration in which the string is located in the (metastable) minimum of the potential at, say, $\phi = -a$. What is the probability that the *entire* string of length L will tunnel across the barrier into the potential minimum at $\phi = a$ in a time t?

▷ INFO. The tunneling of fields between nearly degenerate ground state plays a role in numerous physical contexts. By way of example, consider a **superheated liquid**. In this context, the 'false' vacuum is the liquid state, and the true one the gaseous phase. The role of the field is taken by the local density distribution in the liquid. Thermodynamic fluctuations trigger the continuous appearance of vapor bubbles in the liquid. For bubbles of too small a

[22] For a detailed discussion of the history and ramifications of this idea, we refer on the original insightful paper by Sidney Coleman, *Fate of the false vacuum: semiclassical theory*, Phys. Rev. D **15**, 2929 (1977). Indeed, many of the ideas developed in this work were anticipated in an earlier analysis of metastability in the context of classical field theories by J. S. Langer, *Theory of the condensation point*, Ann. Phys. (N.Y.) **41**, 108 (1967).

diameter, the gain in volume energy is outweighed by the surface energy cost — the bubble will collapse. However, for bubbles beyond a certain critical size, the energy balance is positive. The bubble will grow and, eventually, swallow the entire mass density of the system; the liquid has vaporised or, more formally, the density field has tunneled[23] from the false ground state into the true one.

More speculative (but also potentially more damaging) manifestations of the phenomenon have been suggested in the context of **cosmology**:[22] What if the big bang released our universe not into its true vacuum configuration but into a state separated by a huge barrier from a more favourable sector of the energy landscape. In this case, everything would depend on the tunneling rate: *'If this time scale is of the order of milliseconds, the universe is still hot when the false vacuum decays... if this time is on the order of years, the decay will lead to a sort of secondary big bang with interesting cosmological consequences. If this time is of the order of 10^9 years, we have occasion for anxiety.'* (S. Coleman, *ibid.*).

Previously, for the point–particle system, we have seen that the transition probability between the minima of the double well is most easily accessed by exploring the classical field configurations of the Euclidean time action. In the present case, anticipating to some extent our discussion of the quantum field integral in the next chapter, the Euclidean time action associated with the Hamiltonian density (3.42) assumes the form[24]

$$S[\phi] = \int_0^T d\tau \int_0^L dx \left[\frac{m}{2}(\partial_\tau \phi)^2 + \frac{k_s a^2}{2}(\partial_x \phi)^2 + V(\phi) \right] ,$$

where the time integral runs over the interval $[0, T = it]$. Here, for simplicity, let us assume that the string obeys periodic boundary conditions in space, viz. $\phi(x + L, \tau) \equiv \phi(x, \tau)$. To estimate the tunneling amplitude, we will explore the survival probability of the metastable state imposing the boundary conditions $\phi(x, \tau = 0) = \phi(x, \tau = T) = -a$ on the path integral. Once again, when the potential barrier is high, and the time T is long, one may assume that the path integral is dominated by the saddle–point field configuration of the Euclidean action. In this case, varying the action with respect to the field $\phi(x, \tau)$, one obtains the classical equation of motion

$$m\partial_\tau^2 \phi + k_s a^2 \partial_x^2 \phi = \partial_\phi V(\phi) ,$$

which must be solved subject to the boundary conditions above.

Now, motivated by our consideration of the point–particle problem, one might seek a solution in which the string tunnels as a single rigid entity without 'flexing'. However, it is evident from the spatial translational invariance of the system that the instanton action would scale with the system size L. In the infinite system $L \to \infty$, such a trajectory would therefore not contribute significantly to the tunneling amplitude. Instead, one must consider a different type of field configurataion in which the transfer of the chain is by degree: Elements of the string cross the

[23]At this point, readers should no longer be confused regarding the mentioning of 'tunneling' in the context of a classical system: Within the framework of the path integral, the classical partition sum maps onto the path integral of a fictitious quantum system. It is the tunneling of the latter we have in mind.

[24]Those readers who wish to develop a more rigorous formulation of the path integral for the string may either turn to the discussion of the field integral in the next chapter or, alternatively, may satisfy themselves of the validity of the Euclidean action by (re–)discretisating the harmonic chain, presenting the transition amplitude as a series of Feynman path integrals for each element of the string and, finally, taking the continuum limit.

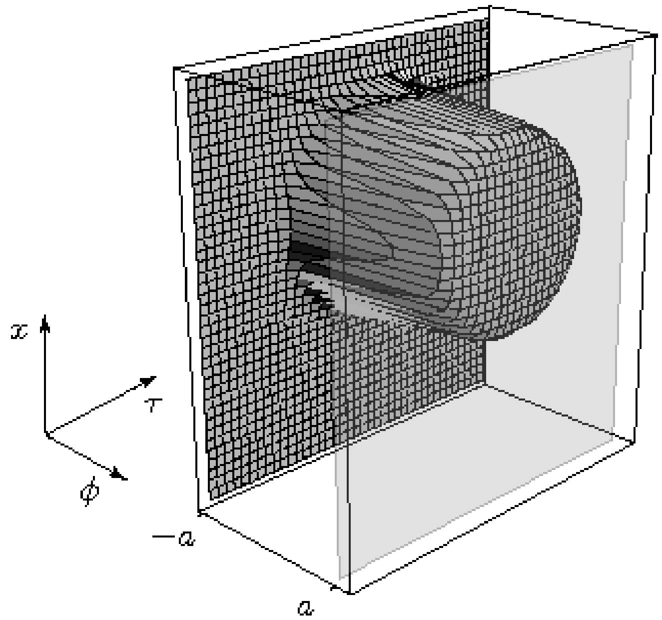


Figure 3.12: On the tunneling between two nearly degenerate vacuum states. As time moves on, a one–dimensional 'world sheet' sweeps through a circular structure in Euclidean space time. This results in the inflation of a 'bubble' of the true vacuum state in real space.

barrier in a consecutive sequence as two outwardly propagating 'domain walls' (see the figure where the emergence of such a double–kink configuration is shown as a function of space and time; notice the spherical shape of the resulting space–time droplet — a consequence of the rotational symmetry of the rescaled problem). Such a field configuration can be motivated from symmetry considerations by noting that, after rescaling $x \mapsto v_s x$ (where $v_s = \sqrt{k_s a^2/m}$ denotes the classical sound wave velocity), the saddle–point equation assumes the isotropic form $m\partial^2\phi = \partial_\phi V(\phi)$, where $\partial^2 = \partial_\tau^2 + \partial_x^2$. Then, setting $r = \sqrt{x^2 + (\tau - T/2)^2}$, and sending $(T, L) \to \infty$, the space–time rotational symmetry suggests a solution of the form $\phi = \phi(r)$ where $\phi(r)$ obeys the radial diffusion equation

$$\partial_r^2\phi + \frac{1}{r}\partial_r\phi = \partial_\phi V\,,$$

with the boundary condition $\lim_{r\to\infty}\phi(r) = -a$. This equation describes the one–dimensional motion of a particle in a potential $-V$ and subject to a strange "friction force" $-r^{-1}\partial_r\phi$ whose strength is inversely proportional to 'time' r.

To understand the profile of the non–trivial bounce solution of the problem, suppose that at time $r = 0$ the particle has been released at rest at a position slightly to the left of the (inverted) potential maximum at a. After rolling through the potential minimum it will climb the potential hill at $-a$. Now, the initial position may be fine tuned such that the viscous damping of the particle compensates for the excess potential energy (which would otherwise make the particle overshoot and disappear to infinity): there exists a solution where the particle starts close to $\phi = a$ and eventually winds up at $\phi = -a$, in accord with the imposed boundary conditions. In general, the analytical solution for the bounce depends sensitively on the form of the confining potential. However, while we assume that the well asymmetry imposed by external potential $-f\phi$ is small, the radial equation may be considerably simplified. In this limit, one may invoke a "thin-wall" approximation in which one assumes that the bounce configuration is described by a domain wall of thickness Δr, at a radius $r_0 \gg \Delta r$ separating an inner region where $\phi(r < r_0) = a$ from the outer region where $\phi(r > r_0) = -a$. In this case, and to lowest order in an expansion in f, the action of the friction force is immaterial, i.e. we may set $m\partial_r^2\phi = \partial_\phi V$ — the very instanton equation formulated earlier for the point–particle system!

Then, when substituted back into S, one finds that the bounce (or kink–like) solution is characterised by the Euclidean action

$$S = v_s \left[2\pi r_0 S_{\text{inst.}} - \pi r_0^2\, 2af\right]$$

where $S_{\text{inst.}}$ denotes the action of the instanton associated with the point–particle system (3.36), and the last term accommodates the effect of the potential bias on the field configuration. Crucially, one may note that the instanton contribution to the action scales with the circumference of the domain wall in the space–time, while that of the potential bias scales with the area of the domain. From this scaling dependence, it is evident that, however small is the external force f, at large enough r_0, the contribution of the second term will always outweigh the first and the string will tunnel from the metastable minimum to the global minimum of the potential. More precisely, the optimal size of domain is found by minimising the action with respect to r_0. In doing so, one finds that $r_0 = S_{\text{inst.}}/2af$. Then, when substituted back into the action, one obtains the tunneling rate

$$\Gamma \sim \exp\left[-\frac{1}{\hbar}\frac{\pi v_s S_{\text{inst.}}^2}{2af}\right] .$$

From this result, one can conclude that, in the absence of an external force f, the tunneling of the string across the barrier is *completely quenched*! In the zero temperature unbiased system, the symmetry of the quantum Hamiltonian is broken: The ground state exhibits a two–fold degeneracy in which the string is confined to one potential minimum or another.

The ramifications of the tunneling amplitude suppression can be traced to the statistical mechanics of the corresponding classical system: As emphasized in section 3.2.1, any Euclidean time path integral of a d–dimensional system can be identified with the statistical mechanics of a classical system $(d+1)$–dimensional problem. In the double well system, the Euclidean time action of the point–particle quantum system is isomorphic to the one–dimensional realisation of the classical Ising ferromagnet, viz.

$$\beta H_{\text{Ising}} = \int_0^L d^d\mathbf{x}\left[\frac{t}{2}m^2 + um^4 + \frac{K}{2}(\nabla m)^2\right] \tag{3.43}$$

Translated into this context, the saddle–point (or mean–field) analysis suggests that the system will exhibit a spontaneous symmetry breaking to an ordered phase ($m \neq 0$) when the parameter t (the reduced temperature) becomes negative. However, drawing on our analysis of the quantum point–particle system, in the thermodynamic limit, we see that fluctuations (non–perturbative in temperature) associated with instanton field configurations of the Hamiltonian $m(x)$ may restore the symmetry of the system and destroy long–range order at any finite temperature $1/\beta$. Whether this happens or not depends on the competition between the energy cost of instanton creation and the entropy gained by integrating over the instanton zero mode coordinates. It turns out that in $d = 1$, the latter wins, i.e. the system is 'disordered' at any finite temperature. In contrast, for $d \geq 2$, the creation of instantons is too costly, i.e. the system will remain in its energetically preferred ground state.

3.3.4 †Tunneling in a Dissipative Environment

▷ ADDITIONAL EXAMPLE: In the condensed matter context it is, of course, infeasible to completely divorce a system from its environment. Indeed, in addition to the dephasing effect of thermal fluctuations, the realization of quantum mechanical phenomena depends sensitively on the strength and nature of the coupling to the external degrees of freedom. For example, the tunneling of an atom from one interstitial site in a crystal to another is likely to be heavily influenced by its coupling to the phonon degrees of freedom that characterise the crystal lattice. By exchanging energy with the phonons, which act in the system as an external bath, a quantum particle can loose its phase coherence and with it, its quantum mechanical character. Beginning with the seminal work of Caldeira and Leggett,[25] there have been numerous theoretical investigations of the effect of an enviroment on the quantum mechanical properties of a system. Such effects are particularly acute in systems where the quantum mechanical degree of freedom is *macroscopic* such as the magnetic flux trapped in a superconducting quantum interference device (SQUID). In the following, we will show that the Feynman path integral provides a natural (and almost unique) setting in which the effects of the environment on a microscopic or macroscopic quantum mechanical degree of freedom can be explored.

Before we begin, let us note that the phenomenon of macroscopic quantum tunneling represents an extensive and still active area of research recently reinvigorated by the burgeoning field of quantum computation. By contrast, our discussion here will be necessarily limited in scope, targeting a particular illustrative application, and highlighting only the guiding principles. For a more thorough and detailed discussion, we refer the reader to one of the many comprehensive reviews.[26]

Caldeira–Leggett Model

Previously, we have discussed the ability of the Feynman path integral to describe quantum mechanical tunneling of a particle q across a potential barrier $V(q)$. In the following, we will invoke the path integral to explore the capacity for quantum mechanical tunneling when the particle is coupled to degrees of freedom of an external environment. Following Calderia and Leggett's original formulation, let us represent the environment by a bath of N quantum harmonic oscillators characterised by a set of frequencies $\{\omega_\alpha\}$,

$$\hat{H}_{\rm bath}[q_\alpha] = \sum_\alpha^N \left[\frac{\hat{p}_\alpha^2}{2m_\alpha} + \frac{m_\alpha}{2}\omega_\alpha^2 q_\alpha^2\right] .$$

For simplicity, let us suppose that in the leading approximation, the coupling of the particle to the degrees of freedom of the bath is linear such that $\hat{H}_{\rm c}[q, q_\alpha] = -\sum_\alpha^N f_\alpha[q] q_\alpha$, where $f_\alpha[q]$ represents some function of the particle coordinate q. Expressed as a Feynman path integral, the survival probability of a particle confined to a metastable minimum at a position $q = a$, and coupled to an external environment, can then be expressed as ($\hbar = 1$)

$$\langle a|e^{-i\hat{H}t/\hbar}|a\rangle = \int_{q(0)=q(t)=a} Dq\ e^{iS_{\rm part.}[q]} \int Dq_\alpha\ e^{iS_{\rm bath}[q_\alpha]+iS_{\rm c}[q,q_\alpha]} ,$$

[25] A. O. Calderia and A. J. Leggett, *Influence of Dissipation on Quantum Tunneling in Macroscopic Systems*, Phys. Rev. Lett. **46**, 211 (1981).

[26] See, e.g., A. J. Leggett *et al.*, *Dynamics of the dissipative two–state system*, Rev. Mod. Phys **48**, 357 (1976), and the text by Weiss [**?**].

where $\hat{H} = \hat{H}_{\text{part.}} + \hat{H}_{\text{bath}} + \hat{H}_{\text{c}}$ denotes the total Hamiltonian of the system,

$$S_{\text{part.}}[q] = \int_0^t dt \left[\frac{m}{2}\dot{q}^2 - V(q)\right], \qquad S_{\text{bath}}[q_\alpha] = \int_0^t dt \sum_\alpha \frac{m_\alpha}{2}\left[\dot{q}_\alpha^2 - \omega_\alpha^2 q_\alpha^2\right],$$

denote, respectively, the action of the particle and bath, while

$$S_{\text{coupling}}[q, q_\alpha] = -\int_0^t dt \sum_\alpha f_\alpha[q] q_\alpha - \int dt \sum_a \frac{f_\alpha[q]^2}{2m_a\omega_a^2},$$

represents their coupling.[27] Here we assume that the functional integral over $q_\alpha(t)$ is taken over all field configurations of the bath while, as before, the path integral on $q(t)$ is subject to the boundary conditions $q(0) = q(t) = a$.

To reveal the effect of the bath on the capacity for tunneling of the particle, it is instructive to integrate out fluctuations q_α and thereby obtain an effective action for q. Fortunately, being Gaussian in the coordinates q_α, the integration can performed straightforwardly. Although not crucial, since we are dealing with quantum mechanical tunneling, it is useful to transfer to the Euclidean time representation. Taking the boundary conditions on the fields $q_\alpha(\tau)$ to be periodic on the interval $[0, T^{-1} \equiv \beta]$, it may be confirmed that the Gaussian functional integral over q_α induces a time non–local interaction of the particle (exercise) $\langle a|e^{-i\hat{H}t/\hbar}|a\rangle = \int Dq\, e^{-S_{\text{eff}}[q]}$ where a constant of integration has been absorbed into the measure and

$$S_{\text{eff}}[q] = S_{\text{part.}}[q] + \frac{1}{2T}\sum_{\omega_n,\alpha} \frac{\omega_n^2 f_\alpha[q(\omega_n)] f_\alpha[q(-\omega_n)]}{m_\alpha\omega_\alpha^2(\omega_\alpha^2 + \omega_n^2)}.$$

Here, the sum $\sum_{\omega_n}$ runs over the discrete set of Fourier frequencies $\omega_n = 2\pi n/\beta$ with n integer.[28] By integrating out the bath degrees of freedom, the particle action acquires an induced contribution. To explore its effect on dissipation and tunneling, it is necessary to specialise our discussion to a particular form of coupling.

In the particular case that the coupling to the bath is linear, viz. $f_\alpha[q(\tau)] = c_\alpha q(\tau)$, the effective action assumes the form (exercise)

$$\boxed{S_{\text{eff}}[q] = S_{\text{part.}}[q] - T\int_0^\beta d\tau\, d\tau'\, K(\tau - \tau') q(\tau) q(\tau')}$$

where $K(\tau) = \int_0^\infty \frac{d\omega}{\pi} J(\omega) D_\omega(\tau)$, $J(\omega) = \frac{\pi}{2}\sum_\alpha \frac{c_\alpha^2}{m_\alpha\omega_\alpha}\delta(\omega - \omega_\alpha)$, and

$$D_\omega(\tau) = -\sum_{\omega_n} \frac{2\omega_n^2}{\omega(\omega^2 + \omega_n^2)} e^{i\omega_n\tau},$$

resembles the Green function of a boson with energy $\hbar\omega$. Physically, the non–locality of the action is easily understood: By exchanging fluctuations with the external bath, a particle can

[27]The second term in the coupling action has been added to keep the effect of the environment minimally invasive (purely dissipative). If it would not be present, the coupling to the oscillator degrees of freedom would effectively *shift* the extremum of the particle potential, i.e. change its potential landscape. Exercise: substitute the solutions of the Euler–Lagrange equations $\delta_{q_\alpha} S[q, q_\alpha] = 0$ — computed for a fixed realization of q — into the action to obtain the said shift.

[28]More preceisely, anticipating our discussion of the Matsubara frequency representation, we have defined the Fourier decomposition on the Euclidean time interval T, viz. $q(\tau) = \sum_m q_m e^{i\omega_m\tau}$, $q_m = T\int_0^\beta d\tau q(\tau) e^{-i\omega_m\tau}$, where $\omega_m = 2\pi m/\beta$ with m integer.

affect a self–interaction, retarded in time. Taken as a whole, the particle and the bath maintain quantum phase coherence. However, when projected onto the particle degree of freedom, the total energy of the system appears to fluctuate and the phase coherence of the particle transport is diminished. To explore the properties of the dissipative action, it is helpful to separate the non–local interaction according to the identity $q(\tau)q(\tau') = [q^2(\tau) + q^2(\tau')]/2 - [q(\tau) - q(\tau')]^2/2$. The former squared contribution presents an innocuous renormalisation of the potential $V(q)$ and, applying equally to the classically allowed motion as well as quantum tunneling, presents an unobservable perturbation. Therefore, we will suppose that its effect has been absorbed into a redefinition of the particle potential $V(q)$. By contrast, the remaining contribution is always positive.

The particular form of the "spectral function" $J(\omega)$ may be obtained either from an *a priori* knowledge of the microscopic interactions of the bath, or phenomenologically, it can be inferred from the structure of the classical damped equations of motion. For example, for a system subject to an "ohmic" dissipation (where, in real time, the classical equations of motion obtain a dissipative term $-\eta\dot{q}$ with a "friction coefficient" η), one has $J(\omega) = \eta|\omega|$ for all frequencies smaller than some characteristic cut–off (at the scale of the inverse Drude relaxation time of the environment). By contrast, for a defect in a three–dimensional crystal, interaction with acoustic phonons present a frequency dependence of ω^3 or ω^5 depending on whether ω is below or above the Debye frequency.

▷ INFO. Consider, for example, the coupling of a particle to a continuum of bosonic modes whose spectral density $J(\omega) = \frac{\eta}{8}\omega$ grows linearly with frequency. In this case,

$$K(\omega_n) = -\frac{\eta\omega_n^2}{8\pi}\int_0^\infty d\omega \frac{1}{\omega^2 + \omega_n^2} = -\frac{\eta}{4}|\omega_n|.$$

describes **Ohmic dissipation** of the particle. Fourier transforming this expression we obtain

$$K(\tau) = -\frac{\pi T\eta}{4}\frac{1}{\sin^2(\pi T\tau)} \overset{\tau \ll T^{-1}}{\simeq} -\frac{\eta}{4\pi T}\frac{1}{\tau^2}, \tag{3.44}$$

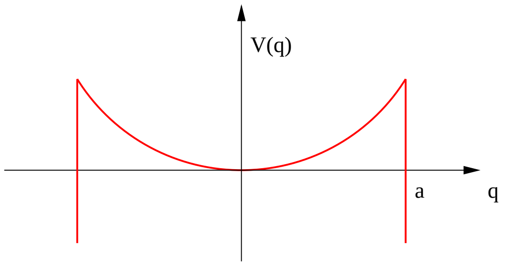
i.e. a strongly time non–local 'self–interaction' of the particle.

Disssipative Quantum Tunneling

Returning to the particular problem at hand, previously, we have seen that the tunneling rate of a particle from a metastable potential minimum can be inferred from the extremal field configurations of the Euclidean action: the bounce trajectory. To explore the effect of the dissipative coupling, it is necessary to understand how it revises the structure of the bounce solution. Now, in general, the non–local character of the interaction inhibits access to an exact solution of the classical equation of motion. In such cases, the effect of the dissipative coupling can be explored perturbatively or with the assistance of the renormalisation group (see the discussion in section **??**). However, by tailoring our choice potential $V(\phi)$, we can gain some intuition about the more general situation.

To this end, let us consider a particle of mass m confined in a metastable minimum by a (semi–infinite) *harmonic* potential trap (see figure),

$$V(q) = \begin{cases} m\omega_c^2 q^2/2 & 0 < q \le a\,, \\ -\infty & q > a\,. \end{cases}$$



Further, let us assume that the environment imparts an ohmic dissipation with a damping or viscosity η. To keep our discussion general, let us consider the combined impact of dissipation and temperature on the rate of tunneling from the potential trap. To do so, following Langer[29] it is natural to investigate the "quasi–equilibrium" quantum partition function $\mathcal{Z}$ of the combined system. In this case, the tunneling rate appears as an imaginary contribution to the free energy $F = -T \ln \mathcal{Z}$, viz. $\Gamma = -2\text{Im}\, F$.

Drawing on the path integral, the quantum partition function of the system can be presented as a functional integral $\mathcal{Z} = \int_{q(\beta)=q(0)} Dq e^{-S_{\text{eff}}}$ where, as we have seen above, for ohmic coupling, the Euclidean action assumes the form

$$S_{\text{eff}}[q] = \int_0^\beta d\tau \left(\frac{m}{2}\dot{q}^2 + V(q)\right) + \frac{\eta}{4\pi}\int_0^\beta d\tau\, d\tau' \left(\frac{q(\tau) - q(\tau')}{\tau - \tau'}\right)^2 .$$

Once again, to estimate the tunneling rate, we will suppose that the barrier is high and the temperature is low so that the path integral is dominated by stationary configurations of the action. In this case, one may identify three distinct solutions: In the first place, the particle may remain at $q = 0$ poised precariously on the maximum of the inverted harmonic potential. Contributions from this solution and the associated harmonic fluctuations reproduce terms in the quantum partition function associated with states of the *closed* harmonic potential trap. Secondly, there exists a singular solution in which the particle remains at the minimum of the inverted potential, i.e. perched on the potential barrier. The latter presents a negligible contribution to the quantum partition function and can be neglected. Finally, there exists a bounce solution in which the particle injected at a position q inside the well accelerates down the inverted potential gradient, is reflected from the potential barrier, and returns to the initial position q in a time β. While, in the limit $\beta \to \infty$, the path integral singles out the boundary condition $q(0) = q(\beta) \to 0$, at finite β, the boundary condition will depart from 0 in a manner that depends non–trivially on the temperature. It is this general bounce solution which governs the decay rate.

Since, in the inverted potential, the classical bounce trajectory stays within the interval over which the potential is quadratic, a variation of the Euclidean action with respect to $q(\tau)$ obtains the classical equation of motion

$$-m\ddot{q} + m\omega_c^2 q + \frac{\eta}{\pi}\int_0^\beta d\tau' \frac{q(\tau) - q(\tau')}{(\tau - \tau')^2} = A\delta(\tau - \beta/2)\,,$$

where the term on the right hand side of the equation imparts an impulse which changes discontinuously the velocity of the particle, while the coefficient A is chosen to ensure symmetry of the bounce solution on the Euclidean time interval. Turning to the Fourier representation, the solution of the saddle–point equation then assumes the form

$$q_n = ATe^{-i\omega_n\beta/2}g(\omega_n), \qquad g(\omega_n) \equiv [m(\omega_n^2 + \omega_c^2) + \eta|\omega_m|]^{-1}. \tag{3.45}$$

[29] J. S. Langer, Ben: . . .

Imposing the condition that $q(\tau = \beta/2) = a$, one finds that $A = a/f$ where $f \equiv T\sum_n g(\omega_n)$. Finally, the action of the bounce is given by

$$S_{\rm bounce} = \frac{1}{2T}\sum_n (m(\omega_n^2 + \omega_c^2) + \eta|\omega_m|)|q_n|^2 = \frac{a^2}{2f}. \tag{3.46}$$

(a) To make sense of these expressions, as a point of reference, let us first determine the **zero temperature tunneling rate in the absence of dissipation**, viz. $\eta \to 0$ and $\beta \to \infty$. In this case, the (Matsubara) frequency summation translates to the continuous integral, $f = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} g(\omega) = (2m\omega_c)^{-1}$. Using this result, the bounce action (3.46) takes the form $S_{\rm bounce} = m\omega_c a^2$. As one would expect, the tunnelling rate $\Gamma \sim e^{-S_{\rm bounce}}$ is controlled by the ratio of the potential barrier height $m\omega_c^2 a^2/2$ to the attempt frequency ω_c. Also notice that the bounce trajectory is given by

$$q(\tau) = \frac{a}{f}\int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\; e^{i\omega(\tau-\beta/2)} g(\omega) = a\, e^{-\omega_c|\tau-\beta/2|},$$

i.e. as expected from our discussion in section 3.3.2, the particle spends only a time $1/\omega_c$ in the under barrier region.

(b) Now, restricting attention to the zero temperature limit, let us consider the **influence of dissipation** on the nature of the bounce solution and the capacity for tunneling. Focussing on the limit in which the dynamics of the particle is overdamped, $\eta \gg m\omega_c$, $f = \int_{-\infty}^{\infty} g(\omega) \simeq \frac{2}{\pi\eta}\ln(\eta/m\omega_c)$, which implies $S_{\rm bounce} = \frac{\pi\eta a^2}{4\ln[\eta/(m\omega_c)]}$. In particular, this result shows that, in the limit $\eta \to \infty$, the coupling of the particle to the ohmic bath leads to an *exponential* suppression of the tunneling rate while only a weak dependence on the jump frequency persists. Physically, this result is easy to rationalise: Under–barrier tunneling is a feature of the quantum mechanical system. By transferring energy to and from the external bath, the phase coherence of the particle is lost. At zero temperature, the tunneling rate becomes suppressed and the particle confined.

(c) Let us now consider the **influence of temperature on the tunneling rate when the dissipative coupling is inactive** $\eta \to 0$. In this case, the discrete frequency summation takes the form[30] $f = T\sum_n g(\omega_n) = \frac{\coth(\beta\omega_c/2)}{2\omega_c m}$. Using this result, one obtains the action $S_{\rm bounce} = m\omega_c a^2\tanh(\beta\omega_c/2)$. In the low temperature limit $\beta \to \infty$, $S_{\rm bounce} = m\omega_c a^2$ as discussed above. At high temperatures $\beta \to 0$, as expected, one recovers a classical activated dependence of the escape rate, viz. $S \simeq \beta m\omega_c^2 a^2/2$.

(d) Finally, let us briefly remark on the **interplay of thermal activation with ohmic dissipation**. Applying the the Euler-Maclaurin formula $\sum_{m=0}^{\infty} f(m) = \int_0^\infty dx\, f(x) + \frac{f(0)}{2} - \frac{f'(0)}{12} + \ldots$ to relate discrete sums over Matsubara frequencies to their zero temperature integral limits, one finds that $S_{\rm bounce}(T) - S_{\rm bounce}(T=0) \propto \eta T^2$. This shows that, in the dissipative regime, an increase in temperature diminishes the tunneling rate with a scale proportion to the damping.

This concludes our cursory discussion of the application of the Feynman path integral to dissipative quantum tunneling. As mentioned above, our brief survey was able only to touch upon the broad field of research. Those interested in learning more about the field of macroscopic

[30]For details on how to implement the discrete frequency summation, see the info block on p 137 below.

quantum tunneling are referred to the wider literature. To close this chapter, we turn now to our penultimate application of the path integral — quantum mechanical spin.

3.3.5 †Path Integral for Spin

▷ ADDITIONAL EXAMPLE: The quantum mechanics of a spin 1/2–particle is a standard example in introductory courses. Indeed, there is hardly any other system whose quantum mechanics is as easy to formulate. Given that, it is perhaps surprising that for a long time the spin problem defied all attempts to cast it in path integral form: Feynman, the architect of the path integral, did not succeed in incorporating spin into the new formalism. It took several decades to fill this gap (for a review of the early history up to 1980, see Schulman's text [20]), and a fully satisfactory formulation of the subject was obtained no earlier than 1988. (The present exposition follows closely the lines of the review by Michael Stone, *Supersymmetry and the Quantum Mechanics of Spin*, Nucl. Phys. **B 314**, 557 (1989).)

Why then is it so difficult to find a path integral of spin? In hindsight it turns out that the spin path integral is in fact no more complex than any other path integral, it merely appears to be a bit unfamiliar. The reason is that, on the one hand, the integrand of the path integral is essentially the exponentiated *classical* action whilst, on the other, the **classical mechanics of spin** is a subject that is not standard in introductory or even advanced courses. In other words, the path integral approach must, by necessity, lead to an unusual object. The fact that the *classical* mechanics of spin is hardly ever mentioned is not only related to the common view that spin is something 'fundamentally quantum' but also to the fact that the mechanics of a classical spin (see below) cannot be expressed within the standard formulation of Hamiltonian mechanics, i.e. there is no formulation in terms of a set of globally defined coordinates and equally many global momenta. It is therefore inevitable that one must resort to the (less widely applied) symplectic formulation of Hamiltonian mechanics.[31] However, as we will see below, the classical mechanics of spin can nevertheless be quite easily understood physically.

Besides attempting to elucidate the connections beween quantum and classical mechanics of spin, there is yet an other motivation for discussing the spin path integral. Pretending that we have forgotten essential quantum mechanics, we will formulate the path integral ignoring the fact that spin quantum numbers are half integer or integer. The quantization of spin will then be derived in hindsight, by way of a *geometric consideration*. In other words, the path integral formulation demonstrates how quantum mechanical results can be obtained by geometric rather than standard algebraic reasoning. Finally, the path integral of spin will serve as a basic platform on which our analysis of higher dimensional spin systems below will be based.

A reminder of finite–dimensional SU(2)–representation theory

In order to formulate the spin path integral, it is neccessary to recapitulate some facts regarding the role of SU(2) in quantum mechanics. The special unitary group in two dimensions, SU(2), is defined as $SU(2) = \{g \in \mathrm{Mat}(2\times 2, \mathbb{C}) | g^\dagger g = \mathbf{1}_2, \det g = 1\}$, where $\mathbf{1}_2$ is the two–dimensional unit matrix. Counting independent components one finds that the group has three free real

[31] Within this formulation, the phase space is regarded as a differential manifold with a symplectic structure (cf. Arnold's text on classical mechanics [**?**]). (In the case of spin, this manifold is the two—sphere S^2.)

parameters or, equivalently, that its Lie algebra, su(2), is three dimensional. As we have seen, the basis vectors of the algebra — the group generators — $\hat{S}^i$, $i = x, y, z$ satisfy the closure relation $[\hat{S}^i, \hat{S}^j] = i\epsilon_{ijk}\hat{S}^k$, where ϵ_{ijk} is the familiar fully antisymmetric tensor. An alternative, and often more useful basis representation of su(2) is given by the spin **raising and lowering operators**, $\hat{S}^\pm = (\hat{S}^x \pm i\hat{S}^y)/2$. Again, as we have seen earlier, the algebra $\{\hat{S}^+, \hat{S}^-, \hat{S}^z\}$ is defined by the commutation relations $[\hat{S}^+, \hat{S}^-] = 2\hat{S}^z$, $[\hat{S}^z, \hat{S}^\pm] = \pm 2\hat{S}^\pm$.

Each group element can be uniquely parametrized in terms of the exponentiated algebra. For example, in the **Euler angle representation**,[32] the group is represented as

$$\mathrm{SU}(2) = \left\{ g(\phi, \theta, \psi) = e^{-i\phi\hat{S}_3} e^{-i\theta\hat{S}_2} e^{-i\psi\hat{S}_3} \middle| \phi, \psi \in [0, 2\pi], \theta \in [0, \pi] \right\}.$$

The Hilbert space $\mathcal{H}_S$ of a quantum spin represents an irreducible representation space of SU(2). Within the spaces $\mathcal{H}_S$, SU(2) acts in terms of representation matrices (which will be denoted by g) and the matrix representations of its generators $\hat{S}_i$. The index S is the so–called weight of the representation (physically: the total spin).[33] Within each $\mathcal{H}_S$, there is a distinguished state, a state of highest weight $|\uparrow\rangle$, which is defined as the (normalized) eigenstate of $\hat{S}^z$ with maximum eigenvalue, S (physically: a spin state polarised in the 3–direction, often denoted as $|S, S_z = S\rangle$, where m is the azimuthal quantum number). Owing to the irreducibility of the representation, each (normalized) state of the Hilbert space $\mathcal{H}_S$ can be obtained by applying the Euler–angle–parameterized elements of the representation to the maximum weight state.

Being a compact group, SU(2) can be integrated over; i.e. it makes sense to define objects like $\int_{\mathrm{SU}(2)} dg f(g)$, where f is some function of g and dg is a realization of a group measure.[34] Among the variety of measures that can be defined in principle, the (unique) **Haar measure** plays a distinguished role. It has the convenient property that it is invariant under left and right multiplication of g by fixed group elements; i.e.

$$\forall h \in \mathrm{SU}(2): \qquad \int dg f(gh) = \int dg f(hg) = \int dg f(g),$$

where, for notational simplicity, we have omitted the subscript in $\int_{\mathrm{SU}(2)}$.

Construction of the path integral

With this background, we are now in a position to formulate the Feynman path integral for quantum mechanical spin. To be specific, let us consider a particle of spin S subject to the Hamiltonian

$$\hat{H} = \mathbf{B} \cdot \hat{\mathbf{S}},$$

32

Leonhard Euler 1707–1783: Swiss mathematician and physicist, one of the founders of pure mathematics. He not only made decisive and formative contributions to the subjects of geometry, calculus, mechanics, and number theory but also developed methods for solving problems in observational astronomy and demonstrated useful applications of mathematics in technology and public affairs.

33 The index S is defined in terms of the eigenvalues of the Casimir operator (physically: the total angular momentum operator) $\hat{\mathbf{S}}^2 \equiv \sum_i \hat{\mathbf{S}}_i^2$ according to the relation $\forall |s\rangle \in \mathcal{H}_S: \ \hat{\mathbf{S}}^2|s\rangle = S(S+1)|s\rangle$.

34 To define group measures in a mathematically clean way, one makes use of the fact that (as a Lie group) SU(2) is a 3–dimensional differentiable manifold. Group measures can then be defined in terms of the associated volume form (see the primer in differential geometry on page **??** below).

where $\mathbf{B}$ is a magnetic field and $\hat{\mathbf{S}} \equiv (\hat{S}_1, \hat{S}_2, \hat{S}_3)$ is a vector of spin operators in the spin–S representation. Our aim is to calculate the imaginary time path integral representation of the quantum partition function $\mathcal{Z} \equiv \mathrm{tr}\, e^{-\beta \hat{H}}$. In constructing the path integral we will follow the general strategy outlined at the end of section 3.2.3, i.e. the first step is to represent $\mathcal{Z}$ as $\mathcal{Z} = \mathrm{tr}\,(e^{-\epsilon \hat{H}})^N$, where $\epsilon = \beta/N$. Next, we have — the most important step in the construction — to insert a suitably chosen resolution of identity between each of the factors $e^{-\epsilon \hat{H}}$. A representation that will lead us directly to the final form of the path integral is specified by

$$\mathrm{id.} = C \int dg\, |g\rangle\langle g| \tag{3.47}$$

where 'id.' represents the unit operator in $\mathcal{H}_S$, $\int dg$ is a group integral over the Haar measure, C is some constant and $|g\rangle \equiv g|\uparrow\rangle$ is the state obtained by letting the representation matrix g act on the maximum weight state $|\uparrow\rangle$ (cf. the summary of the SU(2) representation theory above).

Of course it remains to be verified that the integral (3.47) is indeed proportional to the unit operator. That this is so follows from **Schur's lemma** which states that if, and only if, an operator $\hat{A}$ commutes with all representation matrices of an irreducible group representation (in our case the gs acting in the Hilbert space $\mathcal{H}_S$), $\hat{A}$ is either zero or proportional to the unit matrix. That the group above integral fulfils the global commutativity criterion follows from the properties of the Haar measure: $\forall h \in \mathrm{SU}(2)$,

$$h\int dg|g\rangle\langle g| = \int dg|hg\rangle\langle g| \stackrel{\mathrm{Haar}}{=} \int dg|hh^{-1}g\rangle\langle h^{-1}g| = \int dg|g\rangle\langle g|h.$$

Thus, $\int dg|g\rangle\langle g|$ is, indeed, proportional to the unit operator. The proportionality constant appearing in (3.47) will not be of any concern to us — apart from the fact that it is non–zero.[35]

Substituting the resolution of identity into the time–sliced partition function and making use of the fact that

$$\begin{aligned}\langle g_{i+1}|e^{-\epsilon \mathbf{B}\cdot\hat{\mathbf{S}}}|g_i\rangle &\simeq \langle g_{i+1}|g_i\rangle - \epsilon\langle g_{i+1}|\mathbf{B}\cdot\hat{\mathbf{S}}|g_i\rangle \stackrel{\langle g_i|g_i\rangle=1}{=} 1 - \langle g_i|g_i\rangle + \langle g_{i+1}|g_i\rangle - \epsilon\langle g_{i+1}|\mathbf{B}\cdot\hat{\mathbf{S}}|g_i\rangle \\ &\simeq \exp\Big(\langle g_{i+1}|g_i\rangle - \langle g_i|g_i\rangle - \epsilon\langle g_{i+1}|\mathbf{B}\cdot\hat{\mathbf{S}}|g_i\rangle\Big),\end{aligned}$$

one obtains

$$\mathcal{Z} = \lim_{N\to\infty}\int_{g_N=g_0}\prod_{i=0}^{N} dg_i \exp\left[-\epsilon\sum_{i=0}^{N-1}\left(-\frac{\langle g_{i+1}|g_i\rangle - \langle g_i|g_i\rangle}{\epsilon} + \langle g_{i+1}|\mathbf{B}\cdot\hat{\mathbf{S}}|g_i\rangle\right)\right].$$

Taking the limit $N\to\infty$, the latter can be cast in path integral form,

$$\boxed{\mathcal{Z} = \int Dg \exp\left[-\int_0^\beta d\tau\left(-\langle\partial_\tau g|g\rangle + \langle g|\mathbf{B}\cdot\hat{\mathbf{S}}|g\rangle\right)\right]} \tag{3.48}$$

where the $\mathcal{H}_S$–valued function $|g(\tau)\rangle$ is the continuum limit of $|g_i\rangle$. Eq. (3.48) is our final, albeit somewhat over–compact, representation of the path integral. In order to give this expression some physical interpretation, we need to examine more thoroughly the meaning of the states $|g\rangle$.

[35] Actually, the constant C can be straightforwardly computed by taking the trace of (3.47) which leads to C =(dimension of the representation space)/(volume of the group).

In the literature, the states $|g\rangle$ expressed in the Euler–angle representation

$$\boxed{|\widetilde{g}(\phi,\theta,\psi)\rangle \equiv e^{-i\phi\hat{S}_3}e^{-i\theta\hat{S}_2}e^{-i\psi\hat{S}_3}|\uparrow\rangle}$$

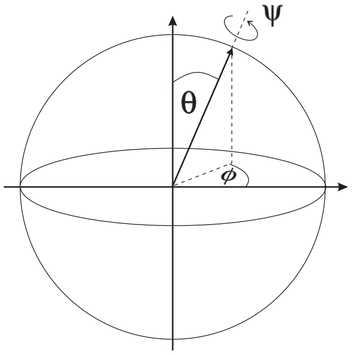
are referred to as **spin coherent states**. Before discussing the origin of this terminology, it is useful to explore the algebraic structure of these states. First, note that the maximum weight state $|\uparrow\rangle$ is, by definition, an eigenstate of $\hat{S}_3$ with maximum eigenvalue S. Thus, $|\widetilde{g}(\phi,\theta,\psi)\rangle \equiv e^{-i\phi\hat{S}_3}e^{-i\theta\hat{S}_2}|\uparrow\rangle e^{-i\psi S}$ and the angle ψ enters the coherent state merely as a phase or gauge factor. By contrast, the two remaining angles θ and ϕ act through true rotations. Now, the angular variables $\phi\in[0,2\pi[$ and $\theta\in[0,\pi[$ define a standard representation of the two–sphere. In view of the fact that (up to normalization factors) the states $|g(\phi,\theta,\psi)\rangle$ cover the entire Hilbert space $\mathcal{H}_S$, we are led to suspect that the latter bears structural similarity with a sphere.[36] To substantiate this view, let us compute the expectation values

$$n_i \equiv \langle\tilde{g}(\phi,\theta,\psi)|\hat{S}_i|\tilde{g}(\phi,\theta,\psi)\rangle, \qquad i=1,2,3. \tag{3.49}$$

To this end, we first derive an auxiliary identity which will spare us much of the trouble that will arise in expanding the exponentials appearing in the definition of $|\widetilde{g}\rangle$. Making use of the general identity $(i\neq j)$

$$e^{-i\phi\hat{S}_i}\hat{S}_j e^{i\phi\hat{S}_i} = e^{-i\phi[\hat{S}_i,\ ]}\hat{S}_j = \hat{S}_j\cos\phi + \epsilon_{ijk}\hat{S}_k\sin\phi, \tag{3.50}$$

where the last equality follows from the fact that $\cos x$ $(\sin x)$ contain x in even (odd) orders and $[\hat{S}_j,\ ]^2\hat{S}_i = \hat{S}_i$, it is a straightforward matter to obtain (exercise) $\mathbf{n} = S(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, i.e. $\mathbf{n}$ is the product of S and a unit vector parameterized in terms of spherical coordinates. This is the key to understanding the terminology 'spin coherent states': The states $|\widetilde{g}(\phi,\theta,\psi)\rangle$ represent the closest approximation of a classical angular momentum one can form out of spin operators (see the figure).



Let us now see what happens if we employ the Euler angle representation in formulating the path integral. A first and important observation is that the path integral is gauge invariant — in the sense that it does not depend on the U(1)–phase, ψ. As for the $\mathbf{B}$–dependent part of the action, the gauge invariance is manifest: Eq. (3.49) implies that

$$S_B[\phi,\theta] \equiv \int_0^\beta d\tau\langle\tilde{g}|\mathbf{B}\cdot\hat{\mathbf{S}}|\tilde{g}\rangle = \int_0^\beta d\tau\langle g|\mathbf{B}\cdot\hat{\mathbf{S}}|g\rangle = S\int_0^\beta d\tau\,\mathbf{n}\cdot\mathbf{B} = SB\int_0^\beta d\tau\cos\theta.$$

[36] There is a group theoretical identity behind this observation, viz. the isomorphism $\mathrm{SU}(2)\simeq S^2\times\mathrm{U}(1)$, where U(1) is the 'gauge' subgroup contained in SU(2).

Here, we have introduced the gauge–independent part $|g\rangle$ of the state vector by setting $|\widetilde{g}\rangle \equiv |g\rangle \exp(-iS\psi)$ or,equivalently, $|g(\phi,\theta)\rangle \equiv e^{-i\phi\hat{S}_3}e^{-i\theta\hat{S}_2}|\uparrow\rangle$. Substituting this representation into the first term of the action of (3.48), one obtains

$$\begin{aligned} S_{\text{top}}[\phi,\theta] &\equiv -\int_0^\beta d\tau \langle\partial_\tau \tilde{g}|\tilde{g}\rangle = -\int_0^\beta d\tau \langle \partial_\tau e^{-iS\psi} g| g e^{-iS\psi}\rangle \\ &= -\int_0^\beta d\tau \left(\langle\partial_\tau g|g\rangle - iS\,\partial_\tau\psi\langle g|g\rangle\right) = -\int_0^\beta d\tau \langle\partial_\tau g|g\rangle, \end{aligned} \tag{3.51}$$

where the last equality holds because $\langle g|g\rangle = 1$ is constant and ψ is periodic in β. As an important intermediate result we have found that the path integral is overall gauge invariant or, equivalently, that the path integral is one over paths living on the two–sphere (rather than the entire group manifold SU(2)). This finding is reasuring in the sense that a degree of freedom living on a sphere comes close to what one might intuitively expect to be the classical counterpart of a quantum particle with conserved angular momentum.

Let us now proceed by exploring the action of the path integral. Using the auxiliary identity (3.50) it is a straightforward matter to show that

$$S_{\text{top}}[\phi,\theta] = -\int_0^\beta d\tau\langle\partial_\tau g|g\rangle = -iS\int_0^\beta d\tau\,\partial_\tau\phi\cos\theta = iS\int_0^\beta d\tau\,\partial_\tau\phi(1-\cos\theta). \tag{3.52}$$

Combining this with the **B**–dependent term discussed above, one obtains

$$\boxed{S[\theta,\phi] = S_B[\phi,\theta] + S_{\text{top}}[\phi,\theta] = S\int_0^\beta d\tau\left[B\cos\theta + i(1-\cos\theta)\partial_\tau\phi\right]} \tag{3.53}$$

for the action of the path integral for spin.

▷ Exercise. Derive the Euler–Lagrange equations asasociated with this action. Show that they are equivalent to the **Bloch equations** $i\partial_\tau \mathbf{n} = \mathbf{B}\times\mathbf{n}$ of a spin with expectation value $\langle\mathbf{S}\rangle = S\mathbf{n}$ subject to a magnetic field. Here, $\mathbf{n}(\phi,\theta)\in S^2$ is the unit vector defined by the two angles ϕ,θ.

Analysis of the action

To formulate the second term in the action (3.53) in a more suggestive way, we note that the velocity of the point $\mathbf{n}$ moving on the unit sphere is given by $\dot{\mathbf{n}} = \dot\theta\hat{\mathbf{e}}_\theta + \dot\phi\sin\theta\,\hat{\mathbf{e}}_\phi$, where $(\hat{\mathbf{e}}_r,\hat{\mathbf{e}}_\theta,\hat{\mathbf{e}}_\phi)$ form a spherical orthonormal system. We can thus rewrite Eq. (3.52) as

$$S_{\text{top}}[\phi,\theta] = iS\int_0^\beta d\tau\,\dot{\mathbf{n}}\cdot\mathbf{A} = iS\oint_\gamma d\mathbf{n}\cdot\mathbf{A}, \tag{3.54}$$

where

$$\mathbf{A} = \frac{1-\cos\theta}{\sin\theta}\hat{\mathbf{e}}_\phi. \tag{3.55}$$

Notice that, in spite of its compact appearance, Eq. (3.54) does not represent a coordinate invariant formulation of the action S_{top}. (The field $\mathbf{A}(\phi,\theta)$ explcitly depends on the coordinates (ϕ,θ).) In fact, the action S_{top} *cannot* be expressed in a coordinate invariant manner, for reasons deeply rooted in the topology of the two–sphere.

A second observation is that (3.54) can be read as the (Euclidean time) action of a particle of charge S moving under the influence of a vector potential $\mathbf{A}$ (cf., for example, Ref. [?].) Using standard formulae of vector calculus (cf. Ref. [?]) one finds $\mathbf{B}_\mathrm{m} \equiv \nabla \times \mathbf{A} = S\mathbf{e}_r$, i.e. our particle moves in a radial magnetic field of constant strength S. Put differently, the particle moves in the field of a magnetic 'charge' of strength 4π centered on the origin of the sphere.

▷ Info. If you find this statement difficult to reconcile with the Maxwell equation $\nabla \cdot \mathbf{B} = 0 \leftrightarrow \int_S \mathbf{B} \cdot dS$ for any closed surface S, notice that $\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}) = 0$ holds only if $\mathbf{A}$ is non–singular. However, the vector potential (3.55) is manifestly singular along the line $(r, \theta = \pi)$ through the south pole of the sphere. The physical picture behind this singularity is as follows: Imagine an infinitely thin solenoid running from $r = \infty$ through the south pole of the sphere to its center. Assuming that the solenoid contains a magnetic flux 4π, the center of the sphere becomes a source of magnetic flux, the so–called **Dirac monopole**. This picture is consistent with the presence of a field $\mathbf{B} = \mathbf{e}_r$. It also explains the singularity of $\mathbf{A}$ along the string. (Of course, the solenoidal construction does not lead to the prediction of a genuine monopole potential: Somewhere, at $r = \infty$, our auxiliary magnetic coil has to end, and this is where the flux lines emanating from the point $r = 0$ terminate.) The postulate of a flux line at the singularity of $\mathbf{A}$ merely helps to reconcile the presence of a radial magnetic field with the principles of electrodynamics. However, as far as our present discussion goes, this extra structure is not essential, i.e. we may simply interpret $r = 0$ as the position of a magnetic 'charge'.

———————————————

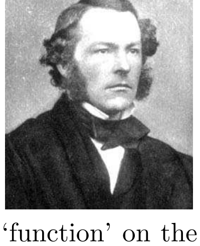
To explore the consequences of this phenomenon, we apply Stokes' theorem[37]

$$S_\mathrm{top}[\mathbf{n}] = iS \oint_\gamma \mathbf{n} \cdot \mathbf{A} = iS \oint_{A_{\gamma,n}} dS \cdot (\nabla \times \mathbf{A}) = iS \oint_{A_{\gamma,n}} dS \cdot \mathbf{e}_r = iSA_{\gamma,n}. \tag{3.56}$$

Here, $A_{\gamma,n}$ is the domain on the two–sphere which (a) has the curve γ as its boundary, and (b) contains the north pole (see the figure). The integral produces the area of this surface which we again denote by $A_{\gamma,n}$. Curiously, the action S_top is but a measure of the area bounded by the curve $\gamma : \tau \mapsto \mathbf{n}(\tau)$. However, simple as it is, this result should raise some suspicion: By assigning a designated role to the *northern* hemisphere of the sphere some symmetry breaking, not present in the original problem, has been introduced. Indeed, we might have defined our action by $S_\mathrm{top}[\phi, \theta] = iS \oint_\gamma d\mathbf{n} \cdot \mathbf{A}'$ where $\mathbf{A}' = -\frac{1+\cos\theta}{\sin\theta}\hat{e}_\phi = \mathbf{A} - 2\nabla\phi$ differs from $\mathbf{A}$ only by a gauge transformation.[38] The newly defined vector potential is non–singular in the *southern* hemisphere, so that application of Stokes' theorem leads to the conclusion $S_\mathrm{top}[\mathbf{n}] = -iS \int_{A_{\gamma,s}} dS \cdot \mathbf{B}_m =$

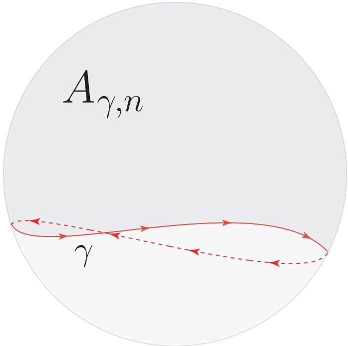
37

George Gabriel Stokes 1819–1903: As Lucasian Professor of Mathematics at Cambridge Stokes etablished the science of hydrodynamics with his law of viscosity (1851), describing the velocity of a small sphere through a viscous fluid. Furthermore, he investigated the wave theory of light, named and explained the phenomenon of fluorescence, and theorised an explanation of the Fraunhofer lines in the solar spectrum.



[38]You may, with some justification, feel uneasy about the fact that ϕ is not a true 'function' on the sphere (or, alternatively, about the fact that $\int d\mathbf{n} \cdot \nabla\phi = \phi(\beta) - \phi(0)$ may be a non–vanishing multiple of 2π). We will return to the discussion of this ambiguity shortly. (Notice that a similarly hazardous manipulation is performed in the last equality of Eq. (3.52).)

$-iSA_{\gamma,s}$. Here, $A_{\gamma,s}$ is the area of a surface bounded by γ but covering the *south* pole of the sphere. The absolute minus sign is due to the outward orientation of the surface $A_{\gamma,s}$.



One has to concede that the result obtained for the action S_{top} depends on the chosen gauge of the monopole vector potential! The difference between the northern and the southern variant of our analysis is given by

$$iS\int_{A_{\gamma,n}} dS\cdot\mathbf{B}_m + iS\int_{A_{\gamma,s}} dS\cdot\mathbf{B}_m = iS\int_{S^2} dS\cdot\mathbf{e}_r = 4\pi iS,$$

where we have made use of the fact that $A_{\gamma,n}\cup A_{\gamma,s}=S^2$ is the full sphere. At first sight, it looks as if our analysis has led us to a gauge dependent, and therefore pathological result. Let's recall, however, that physical quantities are determined by the *exponentiated* action $\exp(iS[\mathbf{n}])$ and not by the action itself. Now, S is either integer or half integer which implies the factor $\exp(4\pi iS)=1$ is irrelevant. In the operator representation of the theory, spin quantization follows from the representation theory of the algebra su(2). It is a 'non–local' feature, in the sense that the action of the spin operators on all eigenstates has to be considered to fix the dimensionality $2S+1$ of $\mathcal{H}_S$. In hindsight, it is thus not too surprising that the same information is encapsulated in a 'global' condition (gauge invariance) imposed on the action of the path integral.

Summarizing, we have found that the classical dynamics of a spin is that of a massless point particle on a sphere coupled to a monopole field $\mathbf{B}_m$. We have seen that the vector potential of the latter cannot be globally continuous on the full sphere. More generally, the phase space S^2 cannot be represented in terms of a global system of 'coordinates and momenta' which places it outside the scope of traditional treatments of classical mechanics. This probably explains the failure of early attempts to describe the spin in terms of a path integral or, equivalently, in terms of a Hamiltonian action.

In chapter **??** we will use the path action (3.53) as a building block for our construction of the field theory of higher dimensional spin systems. However, before concluding this section, let us make some more remarks on the curious properties of the monopole action S_{top}: Contrary to all other Euclidean actions encountered thus far, the action (3.54) is imaginary. In fact, it will stay imaginary upon Wick rotation $\tau\to it$ back to real times. More generally, S_{top} is invariant under the rescaling $\tau\to c\tau$, and invariant even under arbitrary reparameterizations $\tau\to g(\tau)\equiv\tau'$. This invariance is a hallmark of a **topological term**. Loosely speaking (see chapter **??** for a deeper discussion), a topological term is a contribution to the action of a field theory that depends on the global geometry of a field configuration rather than on its local structure. In contrast, 'conventional' operators in field theoretical actions measure the energy cost of dynamical or spatial field fluctuations. In doing to they must relate to a specific spatio–temporal reference frame, i.e. they cannot be invariant under reparamaterisation.

Summarizing our results, we have found that:

1. The classical action of a spin is one of a massless particle (there is no standard kinetic energy term in (3.48)) moving on a unit sphere. The particle carries a magnetic moment of magnitude S. It is coupled to (a) a conventional magnetic field via its magnetic moment, and (b) to a monopole field via its orbital motion. Note that we have come, finally, to a position which hints at the difficulties plaguing attempts to formulate a classical mechanics of spin. The vector potential of a monopole, $\mathbf{A}$, cannot be globally defined on the entire sphere. The underlying physical reason is that, by the very nature of the monopole (flux going radially outwards everywhere), the associated vector potential must be singular at one point of the surface.[39] As a consequence, the classical phase space of the system, the sphere, cannot be covered by a global choice of coordinate system. (Unlike most standard problems of classical mechanics there is no system of globally defined 'p's and 'q's.) This fact largely spoils a description within the standard — coordinate oriented — formulation of Hamiltonian mechanics (cf. the discussion in the article by Stone).

2. Terms akin to the monopole contribution to the spin action appear quite frequently within path integral formulations of systems with non–trivial topology (like the two–sphere above). Depending on the particular context under consideration, one distinguishes between **Wess–Zumino–Witten (WZW) terms,**[40] **θ–terms, Chern–Simons terms** and a few other terms of topological origin. What makes these contributions generally important is that the value taken by these terms depends *only* on the topology of a field configuration but not on structural details.

As a final application of the path integral, we turn now to the consideration of problems in which the dynamics of the classical system is, itself, non–trivial.

3.3.6 †Trace Formulae and Quantum Chaos

▷ Additional Example: Introductory courses on classical mechanics usually convey the impression that dynamical systems behave in a regular and, at least in principle, mathematically predictable way. However, experience shows that the majority of dynamical processes in nature do not conform with this picture: Partly, or even fully chaotic motion (i.e. motion that

[39] To better understand this point, consider the integral of $\mathbf{A}$ along an infinitesimal closed curve γ on the sphere. *If* $\mathbf{A}$ were globally continuous, we would have two choices to transform the integral into a surface integral over $\mathbf{B}$; an integral over the 'large' or the 'small' surface area bounded by γ. The monopole nature of $\mathbf{B}$ would demand that both integrals are proportional to the respective area of the integration domain which, by assumption, are different $\rightsquigarrow$ contradiction. The resolution of this paradox is that $\mathbf{A}$ must be discontinuous at one point on the sphere, i.e. we cannot globally set $\mathbf{B} = \nabla \times \mathbf{A}$ and the choice of the integration area is prescribed by the condition that it must not encompass the singular point.

[40]

Edward Witten 1951–: 1990 Fields Medal for his work in superstring theory. He made significant contributions to Morse theory, supersymmetry, and knot theory. Additionally, he explored the relationship between quantum field theory and the differential topology of manifolds of two and three dimensions.

depends in a singular and, thereby, in an essentially unpredictable way on initial conditions) is the rule rather than the exception. In view of the drastic differences in the observable behaviour of classically integrable and chaotic systems, an obvious question arises: In what way does the *quantum* phenomenology of chaotic systems differ from that associated with integrable dynamics? This question defines the field of **quantum chaos**.

Understanding signatures of classically chaotic motion in quantum mechanics is an issue not only of conceptual, but also of great practical relevance impinging on areas such as quantum electron transport in condensed matter systems: The inevitable presence of impurities and imperfections in any macroscopic solid renders the long–time dynamics of electronic charge carriers chaotic. Relying on a loose interpretation of the Heisenberg principle, $\Delta t \sim \hbar/\Delta E$, i.e. the relation between *long*–time dynamical behaviour and *small* scale structures in energy, one would expect that signatures of chaotic quantum dynamics are especially important in the low–energy response in which one is usually interested. This expectation has been confirmed for innumerable observables related to low temperature electronic transport in solid state systems.

Disordered conducting media represent but one example of a wide class of dynamical systems with long–time chaotic dynamics. Indeed, recent experimental advances have made it possible to realize a plethora of effectively *non–disordered* chaotic dynamical systems in condensed matter devices. For example, employing modern semiconductor device technology, it has become possible to manufacture small two–dimensional conducting systems, of a size $\mathcal{O}(< 1\mu\,\mathrm{m})$ and of almost any geometric shape. Here, the number of imperfections can be reduced to a negligible minimum, i.e. electrons propagate ballistically along straight trajectories, as in a billiard. The smallness of the devices further implies that the ratio between Fermi wavelength and system size is of $\mathcal{O}(10^{-1} - 10^{-3})$, i.e. while semiclassical concepts will surely be applicable, the wave aspects of quantum propagation remain visible. In recent years, the experimental and theoretical study of electron transport in such **quantum billiards** has emerged as a field in its own right.

How then *can* signatures of chaotic dynamics in quantum systems be sought? The most fundamental characteristic of a quantum system is its spectrum. Although not a direct observable, it determines the majority of properties accessible to measurement. On the other hand, it is clear that the manifestations of chaos we are looking for must relate back to the classical dynamical properties of the system. The question then is, *how can a link between classical mechanics and quantum spectra be drawn?* This problem is tailor made for analysis by path integral techniques.

Semiclassical Approximation to the Density of States

The close connection between the path integral and classical mechanics should be evident from the previous sections. However, to address the problem raised above, we still need to understand how the path integral can be employed to analyse the spectrum of a quantum system. The latter are described by the (single–particle) **density of states**

$$\rho(\epsilon) = \mathrm{tr}\,\delta(\epsilon - \hat{H}) = \sum_a \delta(\epsilon - \epsilon_a)\,, \tag{3.57}$$

where $\{\epsilon_a\}$ represents the complete set of energy levels. To compute the sum, one commonly employs a trick based on the **Dirac identity**,

$$\lim_{\delta\searrow 0} \frac{1}{x + i\delta} = -i\pi\delta(x) + \mathcal{P}\frac{1}{x}\,, \tag{3.58}$$

where $\mathcal{P}(1/x)$ denotes for the principal part of $1/x$. Taking the imaginary part of (3.58), Eq. (3.57) can be represented as $\rho(\epsilon) = -\frac{1}{\pi}\,\mathrm{Im}\sum_a \frac{1}{\epsilon^+ - \epsilon_a} = -\frac{1}{\pi}\,\mathrm{Im}\,\mathrm{tr}\,(\frac{1}{\epsilon^+ - \hat{H}})$, where $\epsilon^+ \equiv \epsilon + i\delta$

and the limit $\lim_{\delta\searrow 0}$ is implicit. Using the identity $1/x^+ = -i\int_0^t dt\, e^{ix^+t}$, and representing the trace $\mathrm{tr}\,\hat{A} = \int dq\langle q|\hat{A}|q\rangle$ as a real space integral,

$$\rho(\epsilon) = \frac{1}{\pi\hbar}\int_0^\infty dt\,\mathrm{Re}\,\mathrm{tr}(e^{i(\epsilon^+-\hat{H})t/\hbar}) = \frac{1}{\pi}\mathrm{Re}\int_0^\infty dt\, e^{i\epsilon^+t/\hbar}\int dq\langle q|e^{-i\hat{H}t/\hbar}|q\rangle\,, \tag{3.59}$$

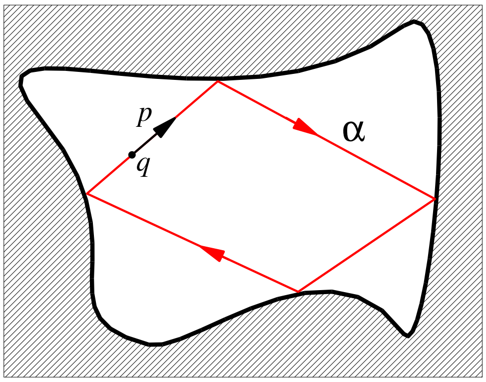
we have made the connection between the density of states and the quantum propagation amplitude explicit.

Without going into full mathematical detail (see, for example, Ref. [**?**] for a modern discourse) we now outline how this integral is evaluated by path integral techniques within the semiclassical approximation. Although, for brevity, some of the more tricky steps of the calculation are swept unter the carpet, the sketch will be accurate enough to make manifest some aesthetic connections between the spectral theory of chaotic quantum systems and classically chaotic dynamics. (For a more formal and thorough discussion, we refer to Gutzwiller and Haake.)

Making use of the semiclassical approximation (3.28) established earlier, when substituted into Eq. (3.59), one obtains $\rho(\epsilon) \simeq \frac{1}{\pi}\mathrm{Re}\int_0^\infty dt\, e^{i\epsilon^+t/\hbar}\int dq\, A[q_{\mathrm{cl}}]e^{\frac{i}{\hbar}S[q_{\mathrm{cl}}]}$, where, following our discussion in section 3.2.2, we have defined $A[q_{\mathrm{cl}}] \equiv \det\left(\frac{i}{2\pi\hbar}\frac{\partial^2 S[q_{\mathrm{cl}}]}{\partial q(0)\partial q(t)}\right)^{1/2}$ and q_{cl} represents a closed classical path that begins at q at time zero and ends at the same coordinate at time t. Again relying on the semiclassical condition $S[q_{\mathrm{cl}}] \gg \hbar$, the integrals over q and t can be performed in a stationary phase approximation. Beginning with the time integral, and noticing that $\partial_t S[q_{\mathrm{cl}}] = -\epsilon_{q_{\mathrm{cl}}}$ is the (conserved) energy of the path q_{cl}, we obtain the saddle point condition $\epsilon \stackrel{!}{=} \epsilon_{q_{\mathrm{cl}}}$ and

$$\rho(\epsilon) \simeq \frac{1}{\pi}\mathrm{Re}\int dq\, A[q_{\mathrm{cl},\epsilon}]e^{\frac{i}{\hbar}S[q_{\mathrm{cl},\epsilon}]}\,,$$

where the symbol $q_{\mathrm{cl},\epsilon}$ indicates that only paths $q\to q$ of energy ϵ are taken into account, and the contribution coming from the quadratic integration around the saddle point has been absorbed into a redefinition of $A[q_{\mathrm{cl},\epsilon}]$.



Turning to the q–integration, making use of the fact that $\partial_{q_i}S[q_{\mathrm{cl}}] = -p_i$, $\partial_{q_f}S[q_{\mathrm{cl}}] = p_f$, where $q_{i,f}$ are the initial and final coordinate of a path q_{cl}, and $p_{i,f}$ are the initial and final momentum, the stationary phase condition assumes the form $0 \stackrel{!}{=} d_q S[q_{\mathrm{cl},\epsilon}] = \left(\partial_{q_i}+\partial_{q_f}\right)S[q_{\mathrm{cl},\epsilon}]|_{q_i=q_f=q} = p_f - p_i$, i.e. the stationarity of the integrand under the q–integration requires the initial and final momentum of the path $q_{\mathrm{cl},\epsilon}$ be identical. We thus find that the paths contributing to the integrated transition amplitude are not only periodic in coordinate space but even in phase space. Such paths are called **periodic orbits** —'periodic' because the path comes back to its initial phase space coordinate after a certain revolution time. As such, the orbit will be traversed repeatedly as time goes by (see the figure, where a periodic orbit α with initial coordinates $x = (p, q)$ is shown).

According to our analysis above, each coordinate point q lying on a periodic orbit is a stationary phase point of the q–integral. The stationary phase approximation of the integral can thus be formulated as

$$\rho(\epsilon) \simeq \frac{1}{\pi}\,\mathrm{Re}\int dq\, A[q_{\mathrm{cl},\epsilon}]e^{\frac{i}{\hbar}S[q_{\mathrm{cl},\epsilon}]} \simeq \sum_{n=1}^{\infty}\sum_{\alpha}\int_{\alpha} dq\, A_\alpha e^{\frac{i}{\hbar}nS_\alpha},$$

where $\sum_\alpha$ stands for a sum over all periodic orbits (of energy ϵ) and S_α is the action corresponding to one traversal of the orbit (all at fixed energy ϵ). The index n accounts for the fact that, due to its periodicty, the orbit can be traversed repeatedly, with total action nS_α. Furthermore, $\int_\alpha dq$ is an integral over all coordinates lying on the orbit and we have again absorbed a contribution coming from the quadratic integration around the stationary phase points in the pre–exponential amplitude A_α.

Finally, noting that $\int_\alpha dq \propto T_\alpha$, where T_α is the period of one traversal of the orbit α (at energy ϵ), we arrive at the result

$$\boxed{\rho(\epsilon) \simeq \frac{1}{\pi}\,\mathrm{Re}\sum_{n=1}^{\infty}\sum_{\alpha} T_\alpha\, A_\alpha e^{\frac{i}{\hbar}nS_\alpha}} \tag{3.60}$$

This is (a simplified[41]) representation of the famous **Gutzwiller trace formula**. The result is actually quite remarkable: The density of states, an observable of quantum mechanical signifi-

[41]Had we carefully kept track of all determinants arising from the stationary phase integrals, the prefactor A_α would have read

$$A_\alpha = \frac{1}{\hbar}\frac{e^{i\frac{\pi}{2}\nu_\alpha}}{|\det M_\alpha^r - 1|^{\frac{1}{2}}},$$

where ν_α is known as the **Maslov index** (an integer valued factor associated with the singular points on the orbit, i.e. the classical turning points). The meaning of this object can be understood, e.g., by applying the path integral to the problem of a quantum particle in a box. To correctly reproduce the spectrum, the contribution of each path must be weighted by $(-)^n = \exp(i\pi n)$, where n is the number of its turning points in the box potential), and M_α represents the **Monodromy matrix**. To understand the meaning of this object, notice that a phase space point $\bar{x}$ on a periodic orbit can be interpreted as a fixed point of the *classical* time evolution operator $U(T_\alpha)$: $U(T_\alpha, \bar{x}) = \bar{x}$, which is just to say that the orbit is periodic. As with any other smooth mapping, U can be linearized in the vicinity of its fixed points, $U(T_\alpha, \bar{x} + y) = \bar{x} + M_\alpha y$, where the linear operator M_α is the monodromy matrix. Evidently, M_α determines the stability of the orbit under small distortions, which makes it plausible that it appears as a controlling prefactor of the stationary phase approximation to the density of states.

▷ Exercise. Making use of the Feynman path integral, show that the propagator for a particle of mass m confined by a square well potential of infinite strength is given by

$$G(q_F, q_I; t) = \sqrt{\frac{m}{2\pi i\hbar t}}\sum_{n=-\infty}^{\infty}\left\{\exp\left[\frac{im(q_F - q_I + 2na)^2}{2\hbar t}\right] - \exp\left[\frac{im(q_F + q_I + 2na)^2}{2\hbar t}\right]\right\}.$$

cance, has been expressed entirely in terms of classical quantities.

3.4 Summary

In this chapter we have introduced the path integral formulation of quantum mechanics, an approach independent of, yet (modulo certain mathematical imponderabilities related to continuum functional integration) equivalent to the standard route of canonical operator quantization. While a few precious exactly solvable quantum problems (e.g. the evolution of a free particle, the harmonic oscillator, and, perhaps intriguingly, quantum mechanical spin) are more efficiently formulated by the standard approach, a spectrum of unique features make the path integral an indispensible tool of modern quantum mechanics: The path integral approach is highly intuitive, powerful in the treatment of non-perturbative problems, and tailor–made to formulation of semiclassical limits. Perhaps most importantly, we have seen that it provides a unifying link whereby quantum problems can be related to classical statistical mechanics. Indeed, we have found that the path integral of a quantum point particle is, in many respects, equivalent to the partition function of a classical one–dimensional continuum system. We have hinted at a generalization of this prinicple, i.e. an equivalence priniciple relating d–dimensional quantum *field* theory to $d+1$–dimensional statistical mechanics. However, before exploring this bridge further, we first need to generalize the concept of path integration to problems involving quantum fields. This will be the subject of the next chapter.