1
Introduction

This book covers the methods by which we can use instantons. What is an instanton? A straightforward definition is the following. Given a quantum system, an instanton is a solution of the equations of motion of the corresponding classical system; however, not for ordinary time, but for the analytically continued classical system in imaginary time. This means that we replace t with $-i\tau$ in the classical equations of motion. Such solutions are alternatively called the solutions of the Euclidean equations of motion.

This type of classical solution can be important in the semi-classical limit $\hbar \to 0$. The Feynman path integral, which we will study in its Euclideanized form in great detail in this book, gives the matrix element corresponding to the amplitude for an initial state at $t = t_i$ to be found in a final state at $t = t_f$ as a "path integral"

$$\begin{aligned}\langle \text{final}, t_f | \text{initial}, t_f \rangle &= \langle \text{final}, t_f | e^{-\frac{i}{\hbar}(t_f - t_i)\hat{h}(\hat{q},\hat{p})} | \text{initial}, t_i \rangle \\ &= \int_{\text{initial},t_i}^{\text{final},t_f} \mathcal{D}p\mathcal{D}q\, e^{\frac{i}{\hbar}\int dt (p\dot{q} - h(p,q))}\end{aligned} \tag{1.1}$$

where $\hat{h}(\hat{q},\hat{p})$ is the quantum Hamiltonian and $h(q,p)$ is the corresponding classical Hamiltonian of the dynamical system. The "path integral" and integration measure $\mathcal{D}p\mathcal{D}q$ defines an integration over all classical "paths" which satisfy the boundary conditions corresponding to the initial state at t_i and to the final state at t_f. It is intuitively evident, or certainly from the approximation method of stationary phase, that the dominant contribution, as $\hbar \to 0$, should come from the neighbourhood of the classical path which corresponds to a stationary (critical) point of the exponent, since the contributions from non-stationary points of the exponent become suppressed as the regions around them cause wild, self-annihilating variations of the exponential.

However, the situation can occur where the particle (or quantum system in general) is classically forbidden from entering some parts of the configuration

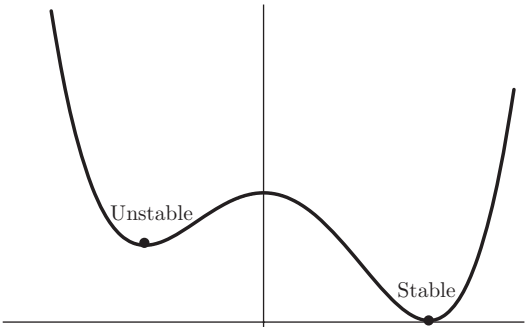

Figure 1.1. A system trapped in the false vacuum will tunnel through the barrier to the state of lower energy

space. In this case we are, generally speaking, considering tunnelling through a barrier, as depicted in Figure 1.1. Classically the particle is not allowed to enter the space where the potential energy is greater than the total energy of the particle. Indeed, if the energy of a particle is given by

$$E = T + V = \frac{\dot{q}^2}{2} + V(q) \tag{1.2}$$

then for a classically fixed energy, regions where $E < V(q)$ require that $T = \frac{\dot{q}^2}{2} < 0$, which means that the kinetic energy has to be negative, and such regions are classically forbidden. Then what takes the role of the dominant contribution in the limit $\hbar \to 0$, since no classical path can interpolate between the initial and final states?

Heuristically such a region is attainable if t becomes imaginary. Indeed, if $t \to -i\tau$ then $\left(\frac{dq}{dt}\right)^2 \to \left(i\frac{dq}{d\tau}\right)^2 = -\left(\frac{dq}{d\tau}\right)^2$, T becomes negative and then perhaps such regions are accessible. Hence it could be interesting to see what happens if we analytically continue to imaginary time, equivalent to continuing from Minkowski spacetime to Euclidean space, which is exactly what we will do. In fact, we will be able to obtain many results of the usual semi-classical WKB (Wentzel, Kramers and Brillouin) approximation [119, 77, 22], using the Euclidean space path integral. The amplitudes that we can calculate, although valid for the small $\hbar$ limit, are not normally attainable in any order in perturbation theory; they behave like $\sim Ke^{-S_0/\hbar}(1 + o(\hbar))$. Such a behaviour actually corresponds to an essential singularity at $\hbar = 0$.

The importance of being able to do this is manifold. Indeed, it is interesting to be able to reproduce the results that can be obtained by the standard WKB method for quantum mechanics using a technique that seems to have absolutely no connection with that method. Additionally, the methods that we will enunciate here can be generalized rather easily to essentially any quantum system, especially to the case of quantum field theory. Tunnelling phenomena

in quantum field theory are extremely important. The structure of the quantum chromodynamics (QCD) vacuum and its low-energy excitations is intimately connected to tunnelling. Various properties of the phases of quantum field theories are dramatically altered by the existence of tunelling. The decay of the false vacuum and the escape from inflation is also a tunnelling effect that is of paramount importance to cosmology, especially the early universe. The method of instantons lets us study all of these phenomena in one general framework.

1.1 A Note on Notation

We will use the following notation throughout this book:

$$\text{metric} \quad \eta_{\mu\nu} = (1, -1, -1, -1) \tag{1.3}$$

$$\text{Minkowsi time} \quad t \tag{1.4}$$

$$\text{Euclidean time} \quad \tau \tag{1.5}$$

$$\text{Analytic continuation of time} \quad t \to -i\tau \tag{1.6}$$

2

Quantum Mechanics and the Path Integral

2.1 Schrödinger Equation and Probability

Our starting point will be single-particle quantum mechanics as defined by the Schrödinger equation

$$i\hbar\frac{d}{dt}\Psi(x,t)=\hat{h}\left(x,-i\hbar\frac{d}{dx}\right)\Psi(x,t). \tag{2.1}$$

Here $\hat{h}(x,-i\hbar\frac{d}{dx})$ is a self-adjoint operator, the Hamiltonian on the space of wave-functions $\Psi(x,t)$, where x stands for any number of spatial degrees of freedom. The connection to physics of $\Psi(x,t)$ comes from the interpretation of $\Psi(x,t)$ as the amplitude of the probability to find the particle between x and $x+dx$ at time t; hence, the probability density is given by

$$\mathcal{P}[x,x+dx]=\Psi^*(x,t)\Psi(x,t). \tag{2.2}$$

Correspondingly, the probability of finding the particle in a volume V is given by

$$\mathcal{P}[V]=\int_V dx\Psi^*(x,t)\Psi(x,t). \tag{2.3}$$

The state of the system is completely described by the wave function $\Psi(x,t)$. It is the content of a standard course on quantum mechanics to find $\Psi(x,t)$ for a given $\hat{h}(x,-i\hbar\frac{d}{dx})$.

2.2 Position and Momentum Eigenstates

For our purposes, we introduce the set of (improper) states $|x\rangle$ which diagonalize the position operator $\hat{X}$, with

$$\hat{X}|x\rangle=x|x\rangle \tag{2.4}$$

and

$$\int dx|x\rangle\langle x|=\mathbb{I}. \tag{2.5}$$

We are in principle working in d dimensions, but we suppress the explicit dependence on the number of coordinates. The states are improper in the sense that the normalization is

$$\langle x|y\rangle = \delta(x-y), \tag{2.6}$$

where $\delta(x-y)$ is the Dirac delta function. We also introduce the set of (improper) states $|p\rangle$ which diagonalize the momentum operator $\hat{P}$

$$\hat{P}|p\rangle = p|p\rangle \tag{2.7}$$

with

$$\int dp|p\rangle\langle p| = 1 \tag{2.8}$$

but as with the position eigenstates

$$\langle p|p'\rangle = \delta(p-p'), \tag{2.9}$$

where $\delta(p-p')$ is the Dirac delta function in momentum space. The improper states $|x\rangle$ and $|p\rangle$ are not vectors in the Hilbert space of states, they have infinite norm. They actually define vector-valued distributions, linear maps from the space of the square integrable functions of x or p or some suitable set of test functions usually taken to be of compact support, to actual vectors in the Hilbert space,

$$|x\rangle : \quad f(x) \to |f\rangle \sim \int dx f(x)|x\rangle, \tag{2.10}$$

where the $\sim$ should be interpreted as "loosely defined by". For a more rigorous definition, see the book by Reed and Simon [107] or Glimm and Jaffe [55].

The operators $\hat{X}$ and $\hat{P}$ must satisfy the canonical commutation relation

$$[\hat{X}, \hat{P}] = i\hbar. \tag{2.11}$$

The algebraic relation Equation (2.11) is not adequate to determine $\hat{P}$ completely; there are infinitely many representations of the commutator Equation (2.11) in which $\hat{X}$ is diagonal. Taking the matrix element of Equation (2.11) between position eigenstates gives

$$(x-y)\langle x|\hat{P}|y\rangle = \langle x|[\hat{X}, \hat{P}]|y\rangle = i\hbar\langle x|y\rangle = i\hbar\delta(x-y). \tag{2.12}$$

For the more mathematically inclined, this expression does not make good sense, since the position and momentum operators are unbounded, though self-adjoint operators. They may only act on their respective domains and, correspondingly, the product of two unbounded operators requires proper analysis of the domains and ranges of the operators concerned and similar other difficulties can exist. We leave these subtleties out in what follows, and refer the interested reader to the

book on functional analysis by Reed and Simon [107]. We find the solution for $\langle x|\hat{P}|y\rangle$ as

$$\begin{aligned}\langle x|\hat{P}|y\rangle &= -i\hbar\frac{d}{dx}\delta(x-y)+c\delta(x-y)\\ &= -i\hbar\frac{d}{dx}\langle x|y\rangle + c\delta(x-y),\end{aligned} \tag{2.13}$$

where c is an arbitrary constant, using the property of the δ function that $(x-y)\delta(x-y)\equiv 0$. We will call the x representation the one in which the momentum operator is represented by a simple derivative, *i.e.* $c=0$,

$$\langle x|\hat{P}|y\rangle = -i\hbar\frac{d}{dx}\langle x|y\rangle. \tag{2.14}$$

In this representation,

$$\begin{aligned}\langle x|\hat{P}|p\rangle &= \int dy\langle x|\hat{P}|y\rangle\langle y|p\rangle = \int dy\left(-i\hbar\frac{d}{dx}\langle x|y\rangle\right)\langle y|p\rangle\\ &= -i\hbar\frac{d}{dx}\langle x|p\rangle.\end{aligned} \tag{2.15}$$

Acting to the right directly in the left-hand side of Equation (2.15) gives

$$\langle x|\hat{P}|p\rangle = p\langle x|p\rangle = -i\hbar\frac{d}{dx}\langle x|p\rangle. \tag{2.16}$$

The appropriately normalized solution of the resulting differential equation is

$$\langle x|p\rangle = \frac{1}{(2\pi\hbar)^{\frac{d}{2}}}e^{i\frac{p\cdot x}{\hbar}}, \tag{2.17}$$

where d is the number of spatial dimensions.

2.3 Energy Eigenstates and Semi-Classical States

We can write the eigenstate of the Hamiltonian in the form $|\Psi_E\rangle$,

$$\hat{h}(\hat{X},\hat{P})|\Psi_E\rangle = E|\Psi_E\rangle, \tag{2.18}$$

where $\hat{h}(\hat{X},\hat{P})$ is defined such that

$$\langle x|\hat{h}(\hat{X},\hat{P})|f\rangle = \hat{h}\left(x,-i\hbar\frac{d}{dx}\right)\langle x|f\rangle \tag{2.19}$$

for any vector $|f\rangle$ in the Hilbert space. Then

$$\langle x|\hat{h}(\hat{X},\hat{P})|\Psi_E\rangle = \hat{h}\left(x,-i\hbar\frac{d}{dx}\right)\langle x|\Psi_E\rangle = E\langle x|\Psi_E\rangle, \tag{2.20}$$

which implies the energy eigenfunctions are given by

$$\Psi_E(x) = \langle x|\Psi_E\rangle. \tag{2.21}$$

Correspondingly,

$$|\Psi_E\rangle = \int dx|x\rangle\langle x|\Psi_E\rangle = \int dx\Psi_E(x)|x\rangle \tag{2.22}$$

and

$$\hat{h}(x, -i\hbar\frac{d}{dx})\Psi_E(x) = E\Psi_E(x). \tag{2.23}$$

A particle described by $\Psi_E(x)$ is most likely to be found in the region where $\Psi_E(x)$ is peaked. The time-dependent solution of the Schrödinger equation for static Hamiltonians is given by $\Psi_E(x,t) = \Psi_E(x)e^{-\frac{i}{\hbar}Et}$, and the most general state of the system is a linear superposition

$$\Psi(x,t) = \sum_E A_E\Psi_E(x)e^{-\frac{i}{\hbar}Et} \tag{2.24}$$

with

$$\sum_E A_E^* A_E = 1. \tag{2.25}$$

Suppose the Hamiltonian can be modified by adjusting the potential, say, such that $\Psi_E(x)$ approaches a delta function:

$$\Psi_E(x) \to \delta(x - x_0). \tag{2.26}$$

We would then say that a particle in the energy level E is localized at the point x_0. But in the limit of Equation (2.26) we clearly have

$$|\Psi_E\rangle \to |x_0\rangle \tag{2.27}$$

from Equation (2.22). Thus the states $|x\rangle$ describe particles localized at the spatial point x. This is conceptually important for the semi-classical limit. Semi-classically we think of particles as localized at points in the configuration space. Thus the states $|x\rangle$ and their generalizations are useful in the description of quantum systems in the semi-classical limit.

2.4 Time Evolution and Transition Amplitudes

Given a particle in a state $|\Psi;0\rangle = |\Psi\rangle$ at $t = 0$, the Schrödinger equation, Equation (2.1), governs the time evolution of the state. The state at $t = T$ is given by

$$|\Psi;T\rangle = e^{-i\frac{T}{\hbar}\hat{h}(\hat{X},\hat{P})}|\Psi\rangle, \tag{2.28}$$

which satisfies the Schrödinger equation. The exponential of a self-adjoint operator, which accurs on the right-hand side of Equation (2.28), is rigorously defined via the spectral representation [107]. The probability amplitude for finding the particle in a state $|\Phi\rangle$ at $t = T$ is then given by

$$\langle\Phi|\Psi;T\rangle = \langle\Phi|e^{-i\frac{T\hat{h}(\hat{X},\hat{P})}{\hbar}}|\Psi\rangle. \tag{2.29}$$

We could derive an expression for this matrix element in terms of a "path integral". Such an integral would be defined as an integral over the space of all classical paths starting from the initial state and ending at the final state, and we would find that the function that we would integrate is the exponential of $-i$ times the classical action for each path. This is the standard Feynman path integral [45, 46], which was actually suggested by Dirac [40].

2.5 The Euclidean Path Integral

Rather than the matrix element Equation (2.29), we are more interested in a path-integral representation of the matrix element

$$\langle \Phi | e^{-\frac{\beta \hat{h}(\hat{X},\hat{P})}{\hbar}} | \Psi \rangle, \tag{2.30}$$

where β can be thought of as imaginary time

$$T \to -i\beta. \tag{2.31}$$

The derivation of the path-integral representation of Equation (2.30) is more rigorous than that for Equation (2.29); however, the derivation which follows can be almost identically taken over to the case of real time. This can be completed by the reader. It is the matrix element of Equation (2.30) that will interest us in future chapters.

First of all, due to the linearity of quantum mechanics, it is sufficient to consider the matrix element

$$\langle y | e^{-\frac{\beta}{\hbar}\hat{h}(\hat{X},\hat{P})} | x \rangle. \tag{2.32}$$

To obtain Equation (2.30) we just integrate over x and y with appropriate smearing functions as in Equation (2.10). Now we write

$$e^{-\frac{\beta \hat{h}(\hat{X},\hat{P})}{\hbar}} = \underbrace{e^{-\frac{\epsilon \hat{h}(\hat{X},\hat{P})}{\hbar}} \cdot e^{-\frac{\epsilon \hat{h}(\hat{X},\hat{P})}{\hbar}} \cdots e^{-\frac{\epsilon \hat{h}(\hat{X},\hat{P})}{\hbar}}}_{N+1 \text{ factors}}, \tag{2.33}$$

where we mean $N+1$ factors on the right-hand side and $(N+1)\epsilon = \beta$. Next we insert complete sets of position eigenstates

$$\int dz_i |z_i\rangle\langle z_i| = \mathbb{I}, \tag{2.34}$$

where $\mathbb{I}$ is the identity operator. Between the evolution operators appearing on the right-hand side of Equation (2.33), there will be N such insertions, *i.e.* $i : 1 \to N$. Consider one of the matrix elements

$$\langle z_i | e^{-\frac{\epsilon \hat{h}(\hat{X},\hat{P})}{\hbar}} | z_{i-1} \rangle \tag{2.35}$$

between position eigenstates $|z_i\rangle$and $|z_{i-1}\rangle$ for Hamiltonians of the form

$$\hat{h}(\hat{X},\hat{P}) = \frac{\hat{P}^2}{2} + V(\hat{X}). \tag{2.36}$$

Then

$$\begin{aligned}\langle z_i|e^{-\frac{\epsilon\hat{h}(\hat{X},\hat{P})}{\hbar}}|z_{i-1}\rangle &= \langle z_i|1-\frac{\epsilon}{\hbar}\left(\hat{P}^2/2+V(\hat{X})\right)|z_{i-1}\rangle+\mathrm{o}(\epsilon^2)\\ &= \int dp_i\langle z_i|1-\frac{\epsilon}{\hbar}\left(p_i^2/2+V(z_{i-1})\right)|p_i\rangle\langle p_i|z_{i-1}\rangle+\mathrm{o}(\epsilon^2)\\ &= \int dp_i\left(1-\frac{\epsilon}{\hbar}\left(p_i^2/2+V(z_{i-1})\right)\right)\langle z_i|p_i\rangle\langle p_i|z_{i-1}\rangle+\mathrm{o}(\epsilon^2)\\ &= \int\frac{dp_i}{(2\pi\hbar)^d}e^{-\frac{\epsilon}{\hbar}\left(\frac{p_i^2}{2}+V(z_{i-1})-ip_i\frac{(z_i-z_{i-1})}{\epsilon}\right)+\mathrm{o}(\epsilon^2)}\\ &= \left(\int\frac{dp_i}{(2\pi\hbar)^d}e^{-\frac{\epsilon}{\hbar}\left(\frac{p_i^2}{2}-i\frac{p_i(z_i-z_{i-1})}{\epsilon}-\frac{1}{2}\left(\frac{z_i-z_{i-1}}{\epsilon}\right)^2\right)}\right)\times\\ &\quad\times e^{-\frac{\epsilon}{\hbar}\left(\frac{1}{2}\left(\frac{z_i-z_{i-1}}{\epsilon}\right)^2+V(z_{i-1})\right)+\mathrm{o}(\epsilon^2)},\end{aligned}\tag{2.37}$$

where in the second step, we have inserted a complete set of momentum eigenstates after letting $V(\hat{X})$ act on the position eigenstate $|z_{i-1}\rangle$. The first factor in the last equality is just a (shifted) Gaussian integral, and can be easily evaluated to give

$$\mathcal{N}_\epsilon=\int\frac{dp_i}{(2\pi\hbar)^d}e^{\frac{-\epsilon}{2\hbar}(p_i-i\frac{(z_i-z_{i-1})}{\epsilon})^2}=\left(\frac{1}{\sqrt{2\pi\hbar\epsilon}}\right)^d.\tag{2.38}$$

Now we use Equations (2.37) and (2.38) in Equation (2.33), inserting an independent complete set of position eigenstates between each of the factors to yield

$$\begin{aligned}\langle y|e^{-\frac{\beta}{\hbar}\hat{h}(\hat{X},\hat{P})}|x\rangle &= \int\frac{dz_1\cdots dz_N}{(2\pi\hbar\epsilon)^{\frac{Nd}{2}}}\prod_{i=1}^{N+1}e^{-\frac{\epsilon}{\hbar}\left(\frac{1}{2}\left(\frac{z_i-z_{i-1}}{\epsilon}\right)^2+V(z_{i-1})\right)+\mathrm{o}(\epsilon^2)}\\ &= \int\frac{dz_1\cdots dz_N}{(2\pi\hbar\epsilon)^{\frac{Nd}{2}}}e^{-\frac{\epsilon}{\hbar}\sum_{i=1}^{N+1}\left(\frac{1}{2}\left(\frac{z_i-z_{i-1}}{\epsilon}\right)^2+V(z_{i-1})\right)+\mathrm{o}(\epsilon^2)},\end{aligned}\tag{2.39}$$

where we define $z_0=x$ and $z_{N+1}=y$. Equation (2.39) is actually as far as one can go rigorously. It expresses the matrix element as a path integral over piecewise straight (N pieces), continuous paths weighted by the exponential of a discretized approximation to the negative Euclidean action. In the limit $N\to\infty$, the $\mathrm{o}(\epsilon^2)$ terms are expected to be negligible. Additionally, in the limit that the path becomes differentiable, which is actually almost never the case,

$$\epsilon\sum_{i=1}^{N}\left(\frac{1}{2}\left(\frac{z_i-z_{i-1}}{\epsilon}\right)^2+V(z_{i-1})\right)\to\int d\tau\left(V(z(\tau))+\frac{1}{2}\dot{z}(\tau))^2\right),\tag{2.40}$$

where $\tau\in[0,\beta]$ parametrizes the path such that $z(0)=x$ and $z(\beta)=y$. Hence the matrix element Equation (2.32) can be formally written as the integral over

classical paths,

$$\begin{aligned}\langle y|e^{-\frac{\beta}{\hbar}\hat{h}(\hat{X},\hat{P})}|x\rangle &= \mathcal{N}\int \mathcal{D}z(\tau)e^{-\frac{1}{\hbar}\int_0^\beta d\tau\left(\frac{1}{2}(\dot{z}(\tau))^2+V(z(\tau))\right)}\\ &= \mathcal{N}\int \mathcal{D}z(\tau)e^{-\frac{S_E[z(\tau)]}{\hbar}},\end{aligned} \tag{2.41}$$

where $S_E[z(\tau)]$ is the classical Euclidean action for each path $z(\tau)$, which starts at x and ends at y. $\mathcal{D}z(\tau)$ is the formal integration measure over the space of all such paths and $\mathcal{N}$ is a formally infinite or ill-defined normalization constant, the limit of $\frac{1}{(2\pi\hbar\epsilon)^{\frac{Nd}{2}}}$ as $N\to\infty$.

There exists a celebrated measure defined on the space of paths, the so-called Wiener measure [121], which was defined in the rigorous study of Brownian motion. One can use it to define the Euclidean path integral rigorously and unambiguously, certainly for quantum mechanics, but also in many instances for quantum field theory. We are not interested in these mathematical details, and we will use and manipulate the path integral as if it were an ordinary integral. We will have to define what we mean by this measure and normalization more carefully, later. The measure actually only makes sense, in any rigorous way, for the discretized version Equation (2.39) including the limit $N\to\infty$; however, strictly speaking the path integral for smooth paths, Equation (2.41), is just a formal expression.

We will record here the corresponding formula in Minkowski time:

$$\begin{aligned}\langle y|e^{-\frac{iT}{\hbar}\hat{h}(\hat{X},\hat{P})}|x\rangle &= \mathcal{N}\int \mathcal{D}z(t)e^{\frac{i}{\hbar}\int_0^T dt\left(\frac{1}{2}(\dot{z}(t))^2-V(z(t))\right)}\\ &= \mathcal{N}\int \mathcal{D}z(t)e^{\frac{i}{\hbar}S_M[z(t)]}.\end{aligned} \tag{2.42}$$

This formula can be proved formally by following each of the steps that we have done for the case of the Euclidean path integral; we leave the details to the reader. However, the Gaussian integral that we encountered at Equation (2.38) becomes

$$\mathcal{N}_\epsilon = \int \frac{dp_i}{(2\pi\hbar)^d}e^{\frac{-i\epsilon}{2\hbar}\left(p_i-i\frac{(z_i-z_{i-1})}{\epsilon}\right)^2}. \tag{2.43}$$

This expression is ill-defined, but it only contributes to an irrelevant normalization constant. We can make it well-defined by adding a small negative imaginary part to the Hamiltonian, which then yields

$$\mathcal{N}_\epsilon = \left(\sqrt{\frac{i}{2\pi\hbar\epsilon}}\right)^d. \tag{2.44}$$