3 Supersymmetry and Geometry

In this section, we will begin our journey into the territory of mathematicians. Our strategy is to think about the physics of a particle moving on a manifold. As this section progresses, we will learn that the quantum ground states of this particle encode some precious information about the manifold.

Before we get to supersymmetry, let's set the scene. We consider a massive, non-relativistic particle moving on the manifold M of dimension $\dim(M) = n$. The dynamics of this particle is described by the Lagrangian

$$L = \frac{1}{2}\, g_{ij}(x)\, \dot{x}^i \dot{x}^j \tag{3.1}$$

where x^i are coordinates on the manifold, with $i = 1, \ldots, n$, and $g_{ij}(x)$ is a Riemannian metric on M.

Lagrangians of the form (3.1) are commonplace in physics, both in quantum mechanics and in higher dimensional quantum field theories. They often go by the unhelpful name of a *sigma model*. Sometimes they are called non-linear sigma models to reflect the fact that, unless g_{ij} is constant, the equations of motion will be non-linear. The name "sigma model" is utterly unilluminating; it dates from one of the first such models written down by Gell-Mann and Levy to describe the dynamics of mesons. (Somewhat comically, Gell-Mann and Levy were building on an earlier model that described both pions and an extra meson known as the "sigma". They then wrote down an improved model that described just the mesons but chose to name it after the missing particle. And the name stuck.)

Geometrically, we should think of the sigma model as a map from the worldline of the particle W to the manifold,

$$x(t) : W \mapsto M$$

The manifold M is known as the *target space*. For much of what we do below, the story will be simplest if M is a compact, orientable manifold and we'll assume this to be the case in what follows.

Strictly speaking, the metric $g_{ij}(x)$ in the Lagrangian should be viewed as the pull back of the metric from M to W. As we saw in earlier courses covering differential geometry, strictly speaking the sigma model only describes the particle in a patch of the manifold M that is covered by the coordinates x^i. One might think that to understand more subtle topological issues, we should be willing to consider overlapping patches. Perhaps surprisingly, it will turn out that this is not necessary, at least in these lectures.

We now ask: what does the particle described by (3.1) know about the manifold M, and what kind of mathematics might it encapsulate? To get a sense for this, we could first think about the Lagrangian (3.1) as describing a classical particle. In this case the equations of motion are the geodesic equations

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0 \tag{3.2}$$

where Γ^i_{jk} is the Levi-Civita connection,

$$\Gamma^i_{jk} = \frac{1}{2}g^{il}\left(\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk}\right)$$

and we're using the notation $\partial_i = \partial/\partial x^i$.

There is certainly a lot of interesting physics in the geodesic equation. But it's challenging to extract any interesting mathematical statements about the manifold M from knowledge of these geodesics. In particular, at any given time, the particle knows only about its immediate surrounding, yet any point looks much the same as any other locally. This means that the state of the particle cannot know anything about the global properties of the manifold. To extract any such information, we would need to know about the entire history of the particle.

This can be contrasted with the situation in quantum mechanics. Now the wavefunction spreads over the manifold M, which suggests that the state of the particle may well know about some of the manifold's quirks. In particular, the state of a quantum particle may be sensitive to the topology of M. Ultimately, we will see that this is indeed the case, at least when we consider supersymmetric extension of our theory. But, for now, let's push on can consider the quantum theory associated to the non-supersymmetric Lagrangian (3.1).

To describe the quantum theory, we first need the momentum

$$p_i = \frac{\partial L}{\partial \dot{x}^i} = g_{ij}\dot{x}^j$$

We then impose canonical commutation relations $[x^i, p_j] = i\delta^i_j$ and construct the Hamiltonian

$$H = p_i\dot{x}^i - L = \frac{1}{2}g^{ij}p_ip_j$$

Already here, things are not so straightforward because the metric g_{ij} depends on x^i and these don't commute with p_i. Different choices of ordering give different quantum Hamiltonians and so different theories.

There is no right or wrong choice here. But we can narrow down our options by requiring that the resulting theory has certain desirable properties. Given that we're interested in the geometry of M, it makes sense to search for a Hamiltonian that is covariant with respect to changes of coordinates on M. In other words, to stick as closely as possible to differential geometry. The action of the momentum on the wavefunction is, as usual, $p_i = -i\partial_i$, so the Hamiltonian should be a second order differential operator with terms that involve no more than two derivatives acting on the metric. There is a one-parameter family of such Hamiltonians, labelled by $\alpha \in \mathbb{R}$,

$$H = -\frac{1}{2\sqrt{g}}\frac{\partial}{\partial x^i}\left(\sqrt{g}g^{ij}\frac{\partial}{\partial x^j}\right) + \alpha R \tag{3.3}$$

where $g = \det g_{ij}$ and R is the Ricci scalar. The first term in this expression is the Laplacian, acting on functions, and can also be written more simply using the covariant derivative,

$$H = -\frac{1}{2}g^{ij}\nabla_i\nabla_j + \alpha R \tag{3.4}$$

We should also decide what Hilbert space we want our operators to act on. The obvious choice is to take the wavefunctions $\psi(x)$ as functions over M, with the norm given by

$$||\psi||^2 = \int d^nx\sqrt{g}\;|\psi(x)|^2 \tag{3.5}$$

Note, in particular, that the inner product includes the factor of $\sqrt{g}$ in the measure, as is appropriate in the geometric context.

Now we have our Hamiltonian (3.4) describing a quantum particle roaming around on a manifold M. What do we do with it? As physicists, our natural inclination is to find the spectrum of the Hamiltonian. We would typically expect that the particle has a unique ground state, with an infinite tower of excited states. This prompts two interesting questions: first, is it possible to calculate this spectrum? Second, what can we do with this information?

Both of these questions are interesting, although neither is easy. In general, it is a difficult problem to determine the spectrum of the Hamiltonian (3.4). Which properties of the manifold can be reconstructed from this spectrum is reminiscent of the famous question "can you hear the shape of a drum?". Mathematicians have spent much time on this question. It is known, for example, that two manifolds may have the same spectrum even though they are not isometric. The first examples are 16-dimensional tori, but subsequent examples have been found in any dimension $n \geq 2$. In fact, it's known that two manifolds may share the same spectrum even if they have different topology (e.g. their fundamental group may be different). All of which is to say that the problem of a quantum particle moving on a manifold M is certainly interesting, but thinking as a physicist provides no particular advantage. We will now see that this situation changes (for the better!) when we introduce supersymmetry.

3.1 The Supersymmetric Sigma Model

There is a beautiful generalisation of the sigma model Lagrangian (3.1) that admits supersymmetry. In addition to the n coordinates x^i, we also introduce n complex Grassmann variables ψ^i, and then consider the action

$$S = \int dt\ \frac{1}{2}\, g_{ij}(x)\, \dot{x}^i \dot{x}^j + i g_{ij} \psi^{\dagger\, i} \nabla_t \psi^j - \frac{1}{4} R_{ijkl} \psi^i \psi^j \psi^{\dagger\, k} \psi^{\dagger\, l} \qquad (3.6)$$

The index i on ψ^i is telling us that the fermions live in the tangent space (strictly the tangent bundle) of M. This is highlighted by the appearance of the covariant derivative, pulled back to the worldline, in the fermion kinetic term

$$\nabla_t \psi^i = \frac{d\psi^i}{dt} + \Gamma^i_{\ jk} \frac{dx^j}{dt} \psi^k$$

As the particle moves on M, the fermions rotate due to this extra term. Finally, note that the four fermion term contracts with the Riemann tensor R_{ijkl}. This is the first suggestion that there might be some pretty geometry lurking in this theory. It's sometimes useful to note that the last term can also be written as

$$\frac{1}{4} R_{ijkl} \psi^i \psi^j \psi^{\dagger\, k} \psi^{\dagger\, l} = \frac{1}{2} R_{ijkl} \psi^i \psi^{\dagger\, j} \psi^k \psi^{\dagger\, l}$$

The equivalence of these two expressions follows from the Riemann tensor identity $R_{i[jkl]} = 0$.

The action (3.6) is invariant under $N = 2$ supersymmetry, given by the following supersymmetry transformations,

$$\begin{aligned} \delta x^i &= \epsilon^\dagger \psi^i - \epsilon \psi^{\dagger\, i} \\ \delta \psi^i &= \epsilon(-i\dot{x}^i + \Gamma^i_{jk} \psi^{\dagger\, j} \psi^k) \\ \delta \psi^{\dagger\, i} &= \epsilon^\dagger(+i\dot{x}^i + \Gamma^i_{jk} \psi^{\dagger\, j} \psi^k) \end{aligned} \qquad (3.7)$$

The associated supercharges are:

$$Q = g_{ij} \dot{x}^i \psi^{\dagger\, j} \quad \text{and} \quad Q^\dagger = g_{ij} \dot{x}^i \psi^j \qquad (3.8)$$

Note that, in contrast to Section 1, we have taken the supercharge Q to depend on ψ rather than $\psi^\dagger$. This is a notational convenience whose advantage we will see as we go along.

How to Show that the Sigma Model is Supersymmetric

Conceptually, it's straightforward to demonstrate the supersymmetricness of the sigma model: you just vary the action, use the transformations (3.7), and show that it vanishes. In practice, you end up with a tsunami of terms. Here's some help to guide you along the way.

First, when implementing the supersymmetry transformation it's useful to set $\epsilon = 0$ and just keep the $\epsilon^\dagger$ terms in the variation. There's no subtlety here: it's just means that we only have to keep track of half the terms in the variation. The other half are then fixed by ultimately requiring that the action and its variation are real. In particular, setting $\epsilon = 0$ means that we have $\delta\psi = 0$ while $\delta\psi^\dagger \neq 0$.

Second, there's a familiar trick, described in the lectures on Quantum Field Theory, that is used to compute the conserved charges associated to any symmetry: we do local variations, instead of global variations. To this end we promote $\epsilon^\dagger \to \epsilon^\dagger(t)$. We will then find the supercharges multiplying the $\dot{\epsilon}^\dagger$ terms in the variation of the action.

Now we can start. Varying the action with $\delta\psi = 0$ but $\delta x, \delta\psi^\dagger \neq 0$ gives

$$
\begin{aligned}
\delta S = \int dt\ g_{ij}\Big(&\dot{x}^i\delta\dot{x}^j + i\delta\psi^{\dagger i}\dot{\psi}^j + i\delta\psi^{\dagger i}\Gamma^j_{kl}\dot{x}^k\psi^l + i\psi^{\dagger i}\delta\Gamma^j_{kl}\dot{x}^k\psi^l + i\psi^{\dagger i}\Gamma^j_{kl}\delta\dot{x}^k\psi^l\Big) \\
&+ \delta g_{ij}\left(\frac{1}{2}\dot{x}^i\dot{x}^j + i\psi^{\dagger i}\dot{\psi}^j + i\Gamma^j_{kl}\dot{x}^k\psi^{\dagger i}\psi^l\right) \\
&- \frac{1}{4}\delta R_{ijkl}\psi^i\psi^j\psi^{\dagger\, k}\psi^{\dagger\, l} - \frac{1}{2}R_{ijkl}\psi^i\psi^j\delta\psi^{\dagger\, k}\psi^{\dagger\, l}
\end{aligned}
$$

where we've used $R_{ijkl} = R_{ij[kl]}$ in the final term. Next it's useful to tame the terms by counting the number of fermions that they contain. There will be terms with 1 fermion, 3 fermions and 5 fermions and if the action is to be invariant, these must individually cancel.

For example, the one-fermion terms come from $\delta x^i = \epsilon^\dagger\psi^i$ in terms that started off with no fermions, and from the first part of the fermion variation $\delta_1\psi^{\dagger\, i} = i\epsilon^\dagger\dot{x}^i$ in terms that started off with two fermions. These are

$$
\begin{aligned}
\delta S\Big|_{1-\text{fermion}} &= \int dt\ g_{ij}\dot{x}^i\delta\dot{x}^j + ig_{ij}\delta_1\psi^{\dagger i}(\dot{\psi}^j + \Gamma^j_{kl}\dot{x}^k\psi^l) + \frac{1}{2}\delta g_{ij}\dot{x}^i\dot{x}^j \\
&= \int dt\ g_{ij}\dot{x}^i(\dot{\epsilon}^\dagger\psi^j + \epsilon^\dagger\dot{\psi}^j) - g_{ij}\epsilon^\dagger\dot{x}^i(\dot{\psi}^j + \Gamma^j_{kl}\dot{x}^k\psi^l) + \frac{1}{2}\partial_l g_{ij}\,\epsilon^\dagger\psi^l\dot{x}^i\dot{x}^j
\end{aligned}
$$

There are two terms with $\epsilon^\dagger\dot{x}\dot{\psi}$ that immediately cancel. We're left with

$$
\delta S\Big|_{1-\text{fermion}} = \int dt\ \epsilon^\dagger\left[\frac{1}{2}\partial_l g_{ij} - g_{ik}\Gamma^k_{ij}\right]\dot{x}^i\dot{x}^j\psi^{\dagger l} + \dot{\epsilon}^\dagger g_{ij}\dot{x}^i\psi^j
$$

The first term vanishes because the metric is covariantly constant, $\nabla g = 0$. Or, in more detail, we use the definition of the Levi-Civita connection, $g_{ik}\Gamma^k_{jl} = \frac{1}{2}\left(\partial_j g_{il} + \partial_l g_{ij} - \partial_i g_{jl}\right)$. But this comes multiplied by $\dot{x}^i\dot{x}^j$ in the variation which means that we get to symmetrise, so $g_{ik}\Gamma^k_{jl}\dot{x}^i\dot{x}^j = \frac{1}{2}\partial_l g_{ij}\dot{x}^i\dot{x}^j$ which, happily, cancels the other term in the variation of the action. We're left with

$$\delta S\Big|_{1-\text{fermion}} = \int dt\ \dot{\epsilon}^\dagger g_{ij}\dot{x}^i\psi^j$$

As explained above, we identify this as the conserved charge arising from the symmetry, $\delta S = \dot{\epsilon}^\dagger Q^\dagger$, giving $Q^\dagger = g_{ij}\dot{x}^i\psi^j$ as advertised in (3.8).

It's simplest to next look at terms with 5 fermions. These come from the δR_{ijkl} term and the $R_{ijkl}\psi^i\delta_2\psi^{\dagger j}\psi^k\psi^{\dagger l}$ term where the we include only the part of the fermion variation that itself has two fermions, $\delta_2\psi^{\dagger j} = \epsilon^\dagger\Gamma^j_{mn}\psi^{\dagger m}\psi^n$. Combined, these terms give

$$\delta S\Big|_{5-\text{fermion}} = \int dt\ \epsilon^\dagger\left[-\frac{1}{4}\partial_m R_{ijkl}\psi^m\psi^i\psi^j\psi^{\dagger\,k}\psi^{\dagger\,l} + \frac{1}{2}R_{ijkl}\Gamma^j_{mn}\psi^i\psi^j\psi^{\dagger m}\psi^n\psi^{\dagger l}\right]$$

After using the fermions to impose anti-symmetry, this term vanishes by virtue of the Bianchi identity $\nabla_{[m}R_{ij]kl} = 0$.

This leaves us with the 3-fermions terms in the variation of the action. They are, of course, everything that we didn't yet consider.

$$\begin{aligned}
\delta S\Big|_{3-\text{fermion}} &= \int dt\ \delta g_{ij}\left(i\psi^{\dagger i}\dot{\psi}^j + i\Gamma^j_{kl}\dot{x}^k\psi^{\dagger i}\psi^l\right) + ig_{ij}\,\delta_2\psi^{\dagger i}(\dot{\psi}^j + \Gamma^j_{kl}\dot{x}^k\psi^l) \\
&\qquad + ig_{ij}\psi^{\dagger i}\left(\delta\Gamma^j_{kl}\dot{x}^k\psi^l + \Gamma^j_{kl}\delta\dot{x}^k\psi^l\right) - \frac{1}{2}R_{ijkl}\psi^i\psi^j\delta_1\psi^{\dagger\,k}\psi^{\dagger\,l} \\
&= \int dt\ \epsilon^\dagger\partial_m g_{ij}\psi^m\left(i\psi^{\dagger i}\dot{\psi}^j + i\Gamma^j_{kl}\dot{x}^k\psi^{\dagger i}\psi^l\right) + ig_{ij}\epsilon^\dagger\Gamma^i_{mn}\psi^{\dagger m}\psi^n(\dot{\psi}^j + \Gamma^j_{kl}\dot{x}^k\psi^l) \\
&\qquad - ig_{ij}\epsilon^\dagger\left[\psi^{\dagger i}\partial_m\Gamma^j_{kl}\psi^m\dot{x}^k\psi^l + \psi^{\dagger i}\Gamma^j_{kl}\dot{\psi}^k\psi^l\right] + \frac{i}{2}\epsilon^\dagger R_{ijkl}\psi^i\psi^j\dot{x}^k\psi^{\dagger l}
\end{aligned}$$

There are two different kinds of terms in this expression. The first take the form $\psi^\dagger\psi\dot{\psi}$. Gathering them together, we find that they come multiplying $\nabla g = 0$. The second take the form $\dot{x}\psi^\dagger\psi\psi$. The first of these involve combinations of the connection that gather together to give $\partial\Gamma + \Gamma^2$. But this is the definition of the Riemann tensor and is cancelled by the final term above. The upshot is that, for a global variation with $\dot{\epsilon}^\dagger = 0$, we have $\delta S = 0$: the action is supersymmetric.

3.1.1 Quantisation: Filling in Forms

Quantising the sigma model needs a little care due to operator ordering issues. The canonical momenta are

$$p_i = \frac{\partial L}{\partial \dot{x}^i} = g_{ij}\left(\dot{x}^i + i\Gamma^i_{kl}\psi^{\dagger\,k}\psi^l\right) \quad \text{and} \quad \frac{\partial L}{\partial \dot{\psi}^i} = ig_{ij}\psi^{\dagger\,i}$$

We have, as always

$$[x^i, p_j] = \delta^i_j \quad \text{and} \quad \{\psi^i, \psi^{\dagger\,j}\} = g^{ij}$$

The tricky commutator is, it turns out, the one between bosons and fermions. This is best described in the terms of the *mechanical momentum* as opposed to the canonical momentum,

$$\pi_i = g_{ij}\dot{x}^j = p_i - ig_{il}\Gamma^l_{jk}\psi^{\dagger\,j}\psi^k$$

The associated commutation relations turn out to be

$$[\pi_i, \psi^j] = i\Gamma^j_{ik}\psi^k \quad , \quad [\pi_i, \psi^{\dagger\,j}] = i\Gamma^j_{ik}\psi^{\dagger\,k} \quad \text{and} \quad [\pi_i, \pi_j] = -R_{ijkl}\psi^{\dagger\,k}\psi^l$$

Let's now look more closely at the Hilbert space of fermions. We quantise the fermions in the usual way: we introduce a state $|0\rangle$ that obeys

$$\psi^i|0\rangle = 0$$

for all $i = 1, \ldots, n$. We then build up the Hilbert space by acting with successive $\psi^{\dagger\,i}$. At the first level we have n states, $\psi^{\dagger\,i}|0\rangle$. At the next level we have $\frac{1}{2}N(N-1)$ states, $\psi^{\dagger\,i}\psi^{\dagger\,j}|0\rangle = -\psi^{\dagger\,j}\psi^{\dagger\,i}|0\rangle$, and so on. The natural anti-symmetry of Grassmann objects means that there are $\binom{n}{p}$ states of the form $(\psi^\dagger)^p|0\rangle$.

As we already advertised in Section 1.4.1, this is a very familiar structure in geometry: it arises for totally anti-symmetric $(0, p)$ tensor fields, also known as *p-forms*. This prompts the identification

$$\begin{aligned}
|0\rangle &\quad\longleftrightarrow\quad 1 \\
\psi^{\dagger\,i}|0\rangle &\quad\longleftrightarrow\quad dx^i \\
\psi^{\dagger\,i}\psi^{\dagger\,j}|0\rangle &\quad\longleftrightarrow\quad dx^i \wedge dx^j \\
&\quad\vdots \\
\psi^{\dagger\,1}\ldots\psi^{\dagger\,n}|0\rangle &\quad\longleftrightarrow\quad dx^1 \wedge \ldots \wedge dx^n
\end{aligned}$$

States in the Hilbert space of supersymmetric quantum mechanics are no longer just functions over the manifold M, but now all forms over the manifold M. States of the kind $f(x)(\psi^\dagger)^p|0\rangle$ correspond to p-forms. We denote the space of p-forms over M as $\Lambda^p(M)$.

This relation between Grassmann variables and forms, identifying $\psi^{\dagger\, i} \;\longleftrightarrow\; dx^i\wedge$ provides the key link between supersymmetry and more interesting aspects of geometry. From this, many lovely geometrical facts follow. For example, we can ask: what is the geometrical interpretation of ψ^i? From the commutation relation $\{\psi^i, \psi^{\dagger\, j}\} = g^{ij}$, it clearly acts as a map $\psi^i : \Lambda^p(M) \mapsto \Lambda^{p-1}(M)$. We can be more explicit and check

$$\begin{aligned}\psi^i\psi^{\dagger\, j}\psi^{\dagger\, k}\ldots\psi^{\dagger\, l}|0\rangle &= \{\psi^i, \psi^{\dagger\, j}\psi^{\dagger\, k}\ldots\psi^{\dagger\, l}\}|0\rangle \\ &= \left[(g^{ij}\psi^{\dagger\, k}\ldots\psi^{\dagger\, l}) - (\psi^{\dagger\, j}g^{ik}\ldots\psi^{\dagger\, l}) + \ldots\right]|0\rangle\end{aligned}$$

But, in the language of forms, this is the action of the *interior product*,

$$\psi^i \;\longleftrightarrow\; g^{ij}\iota_{\partial/\partial x^j}$$

Meanwhile, the inner product between states in the Hilbert space is,

$$\langle\omega|\eta\rangle = \int_M \bar{\omega}\wedge\star\eta \tag{3.9}$$

Where $\bar{\omega}$ is the complex conjugation of ω and $\star$ is the Hodge dual. Note that this is non-vanishing only if ω and η are forms of the same degree p. Furthermore, evaluated on functions $\omega \in \Lambda^0(M)$, it reproduces the norm (3.5).

The Lagrangian (3.6) has a $U(1)$ symmetry acting on fermions as

$$\psi^i \to e^{i\alpha}\psi^i \quad \text{and} \quad \psi^{\dagger\, i} \to e^{-i\alpha}\psi^{\dagger\, i}$$

The corresponding Noether charge is

$$F = g_{ij}\psi^{\dagger\, i}\psi^j$$

which counts the number of fermionic excitations or, in our new geometrical language, the degree of the form. If we have a state $|\phi\rangle \in \Lambda^p(M)$, then

$$F|\phi\rangle = p|\phi\rangle$$

The fact that F is conserved means that Hamiltonian evolution doesn't mix up forms of different degrees: energy eigenstates lie in a particular $\Lambda^p(M)$. The fermion number F also provides the grading that splits our Hilbert space into bosonic and fermionic pieces: $\mathcal{H} = \mathcal{H}_B \oplus \mathcal{H}_F$. These comprise of even and odd forms respectively.

$$\mathcal{H}_B = \mathbb{C}\otimes\left[\bigoplus_{p\ \text{even}}\Lambda^p(M)\right] \quad \text{and} \quad \mathcal{H}_F = \mathbb{C}\otimes\left[\bigoplus_{p\ \text{odd}}\Lambda^p(M)\right]$$

where the overall factor of $\mathbb{C}$ is there simply because wavefunctions are complex valued in quantum mechanics rather than real.

Finally, we come to the supercharges Q and $Q^\dagger$ themselves. The presence of the momentum operator means that these act as derivatives, while the fermions ensure that they also map $Q : \Lambda^p(M) \mapsto \Lambda^{p+1}(M)$. But there is a very natural object in differential geometry with these properties: it is the exterior derivative

$$Q = i\psi^{\dagger\, i} p_i \;\longleftrightarrow\; dx^i \wedge \frac{\partial}{\partial x^i} = d$$

Similarly, $Q^\dagger : \Lambda^p(M) \mapsto \Lambda^{p-1}(M)$ act as the adjoint operator

$$Q^\dagger = i\psi^i p_i \;\longleftrightarrow\; g^{ij}\iota_{\partial/\partial x^i}\frac{\partial}{\partial x^j} = d^\dagger$$

Acting on p-forms, the adjoint operator can also be written as

$$d^\dagger = (-1)^{n(p+1)+1} \star d \star$$

This adjoint operator annihilates functions $d^\dagger f = 0$ for $f \in \Lambda^0(M)$. This is to be expected since it follows from $\psi^i|0\rangle = 0$. Similarly, the exterior derivative itself annihilates top forms, $d\omega = 0$ for all $\omega \in \Lambda^n(M)$. Before we go on, note that the correspondence $Q \equiv d$ and $Q^\dagger \equiv d^\dagger$ is the reason that we chose to define Q to be the supercharge involving $\psi^\dagger$ rather than, as in Section 1, in terms of ψ.

The identification of the supercharges also gives a geometric meaning to the Hamiltonian. It is

$$H = \frac{1}{2}\{Q, Q^\dagger\} \quad \Rightarrow \quad H = \frac{1}{2}\Delta$$

with

$$\Delta = dd^\dagger + d^\dagger d$$

This is the *Laplacian* operator in differential geometry. It is clear from its definition in terms of d and $d^\dagger$ that it is a prime candidate for a supersymmetric Hamiltonian; in some sense everything that we've done above is just to realise this possibility in terms of Grassmann variables ψ and $\psi^\dagger$.

The Laplacian is positive definite, as befits a supersymmetric Hamiltonian. This follows from the fact that the $\dagger$ in $Q^\dagger$ (or, equivalently $d^\dagger$) means the adjoint operation with respect to the inner product (3.9) so that, for any $\omega \in \Lambda^p(M)$,

$$\langle\omega|\Delta\omega\rangle = \langle\omega|dd^\dagger\omega\rangle + \langle\omega|d^\dagger d\omega\rangle = ||d^\dagger\omega||^2 + ||d\omega||^2 \geq 0 \tag{3.10}$$

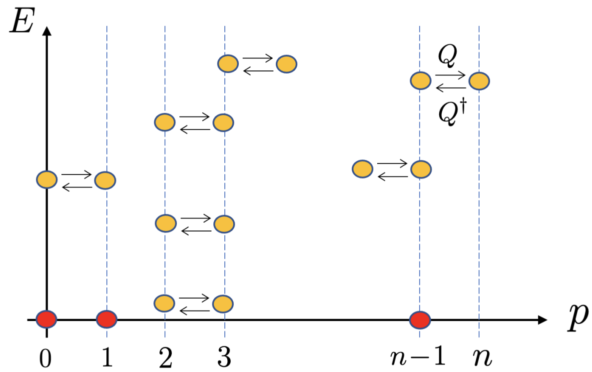


Figure 8. A typical energy spectrum of the quantum mechanics, with pairs of states related by supersymmetry in yellow and the lone ground states, associated to de Rham cohomology, in red.

A short calculation (see, for example, Section 3.1.4 of the lectures on General Relativity) shows that, when acting on function $f \in \Lambda^0(M)$, the Laplacian is given by

$$\Delta f = \frac{1}{\sqrt{g}} \partial_i (\sqrt{g} g^{ij} \partial_j f)$$

in agreement with the Hamiltonian (3.3) for the non-supersymmetric sigma model. Note, however, that in the absence of supersymmetry there was always the option to add the αR term in (3.3) to the Hamiltonian. Supersymmetry removes this ambiguity.

3.1.2 Ground States and de Rham Cohomology

We'll now consider the kind of spectrum that we expect to find. As we saw in Section 1, all states with energy $E \neq 0$ must come in pairs. In particular, if an energy eigenstate state with $E \neq 0$ obeys

$$Q|\alpha\rangle = 0$$

then $|\alpha\rangle$ is Q-*exact*, meaning that it can be written as $|\alpha\rangle = Q|\phi\rangle$ for some $|\phi\rangle$. To see this, we just need to use $QQ^\dagger + Q^\dagger Q = 2E$ to see that

$$|\alpha\rangle = \frac{1}{2E} Q \left(Q^\dagger |\alpha\rangle \right)$$

This tells us that Q and $Q^\dagger$ map us back and forth between the two states related by supersymmetry. In the form language, we see that supersymmetry relates pairs of p and $p + 1$ forms. These are shown as the yellow dots in Figure 8.

However, as we've seen in previous examples, the ground states with $E = 0$ are special since there is no need for these to be paired. In the present context, the ground states arise from forms that obey

$$\Delta\gamma = 0 \quad \Leftrightarrow \quad d\gamma = d^\dagger\gamma = 0$$

Forms of this kind are called *harmonic*. These are depicted as red dots in Figure 8. The space of harmonic p-forms is denoted $\mathrm{Harm}^p(M)$. We learn that the Hilbert space of ground states is

$$\mathcal{H}_{\text{ground}} = \bigoplus_p \mathrm{Harm}^p(M)$$

This discussion also tells us that there are three kinds of states in the Hilbert space: those for which $|\phi\rangle = Q|\alpha\rangle$ or $|\phi\rangle = Q^\dagger|\beta\rangle$, which sit in supersymmetric pairs. And those for which $Q|\phi\rangle = Q^\dagger|\phi\rangle = 0$ which are the supersymmetric ground states. This means that any state $|\phi\rangle \in \mathcal{H}$ has a unique decomposition as

$$|\phi\rangle = Q|\alpha\rangle + Q^\dagger|\beta\rangle + |\omega\rangle \tag{3.11}$$

where $\Delta|\omega\rangle = 0$. In the geometric language, this is equivalent to saying that any form can be written uniquely as

$$\omega = d\alpha + d^\dagger\beta + \gamma \tag{3.12}$$

where ω is harmonic. This is known as the *Hodge decomposition theorem*.

There is an important comment to make here. The Hodge decomposition theorem is not a trivial statement in mathematics. It took Hodge much of the 1930s to prove and, even then, needed corrections from Weyl and Kodaira. Yet the statement about the decomposition of states in the Hilbert space (3.11) follows trivially from the structure of supersymmetric quantum mechanics! What's going on?

Shortly we will "prove" other theorems in geometry where we will make use of the physicist's secret weapon, the path integral. Here, however, the power of physics comes only from our blatant disregard for anything approaching rigour. In geometry, the space of differential forms is *not* a Hilbert space because the inner product (3.9) is not complete. In quantum mechanics, we deal with this by restricting attention to L^2 forms but then one has to worry whether the exterior derivative acts solely within this space. All of these are subtleties that we sweep under the rug in physics, but present the real challenge behind the proof of the Hodge decomposition theorem.

Cohomology

There is another way to view the ground states in terms of *cohomology*. As we've seen, the exterior derivative d (or equivalently the supercharge Q) maps us from $d : \Lambda^p(M) \to \Lambda^{p+1}(M)$. We can depict this in terms of what mathematicians call a *chain complex*

$$\Lambda^0(M) \xrightarrow{d} \Lambda^1(M) \xrightarrow{d} \Lambda^2(M) \xrightarrow{d} \Lambda^3(M) \xrightarrow{d} \ldots$$

Because $d^2 = 0$, the image of one map necessarily lies in the kernel of the next. The idea of cohomology is that it's interesting to look more closely at the difference between the kernel and image.

First some definitions. A form ϕ is said to be *closed* if $d\phi = 0$. We denote the space of all closed p-forms as $Z^p(M)$. Another way to say this is that $Z^p(M)$ is the kernel of the map $d : \Lambda^p(M) \to \Lambda^{p+1}(M)$.

A form ϕ is said to be *exact* if it can be written as $\phi = d\alpha$ for some α. We denote the space of all exact p-forms as $B^p(M)$. Another way to say this is that $B^p(M)$ is image of the map $d : \Lambda^{p-1}(M) \to \Lambda^p(M)$.

As we mentioned above, we necessarily have $B^p(M) \subset Z^p(M)$. The *de Rham cohomology group* is defined to be

$$H^p(M) = Z^p(M)/B^p(M)$$

The quotient here is an equivalence class. Two closed forms ϕ and $\phi' \in Z^p(M)$ are said to be equivalent if $\phi = \phi' + d\alpha$ for some α. We say that ϕ and ϕ' sit in the same equivalence class $[\phi]$. The cohomology group $H^p(M)$ is the set of equivalence classes. In other words, it consists of closed forms mod exact forms.

Finally, we define the *Betti numbers*,

$$b_p = \dim H^p(M)$$

There are a number of interesting things about these Betti numbers. First, this counting of cohomology classes is just another way of counting the ground states in quantum mechanics, and the Betti numbers can equally well be viewed as counting harmonic forms. This follows from...

Claim: There is an isomorphism $H^p(M) \cong \mathrm{Harm}^p(M)$ and so

$$b_p = \dim \mathrm{Harm}^p(M)$$

Proof: The proof follows straightforwardly from the Hodge decomposition (3.12). We'll first show that each harmonic form is associated to an element of $H^p(M)$. Clearly any harmonic form γ is closed, with $d\gamma = 0$. But the unique nature of the Hodge decomposition (3.12) means that γ cannot be written as $\gamma = d(\text{something})$ and so forms the basis of an equivalence class $[\gamma] \in H^p(M)$.

Next we must go the other way and show that each equivalence class of $[\omega] \in H^p(M)$ is associated to a harmonic form. We decompose $\omega = d\alpha + d^\dagger\beta + \gamma$. By the definition of $[\omega] \in H^p(M)$, we must have $d\omega = 0$ and so, using the inner product (3.9), we have

$$0 = \langle d\omega|\beta\rangle = \langle \omega|d^\dagger\beta\rangle = \langle d\alpha + d^\dagger\beta + \gamma|d^\dagger\beta\rangle = \langle d^\dagger\beta|d^\dagger\beta\rangle$$

where, in the final step, we integrated by parts and used the facts that $dd\alpha = 0$ and $d\gamma = 0$. The upshot is that $d^\dagger\beta = 0$ and any element of the equivalence class $[\omega] \in H^p(M)$ takes the form $\omega = d\alpha + \gamma$. Any other member of the same equivalence class $\omega' \in [\omega]$ can be written as $\omega' = d\eta + \gamma$ and is associated to the same harmonic form γ. $\square$

There's an analogy here with gauge symmetry that is worth highlighting. In Maxwell theory, the gauge potentials A and $A + d\alpha$ are physically equivalent as they are related by a gauge transformation. If we want to pick a representative of this equivalence class then we need gauge fixing condition that picks out one particular choice of A. For cohomology, the equivalence class $[\omega]$ relates $\omega \sim \omega + d\alpha$. A representative of this class can be picked by the "gauge fixing condition" $d^\dagger\omega = 0$. This then picks out the harmonic forms as special.

Any manifold M with dimension $\dim(M) = n$ always has $b_0 = 1$ and $b_n = 1$. The zero forms are just functions over the manifold, and any constant function over M is clearly harmonic, but cannot be written as $d(\text{somethng})$ as there are no $p = -1$ forms. Similarly, the volume form $\text{Vol} = \star 1$ provides the harmonic top form.

Other Betti numbers come in pairs with $b_p = b_{n-p}$, a relationship that follows from Poincaré duality. It turns out that all these higher Betti numbers are non-vanishing only if the manifold M has some interesting topology. To explain this, we need to remove the co in cohomology.

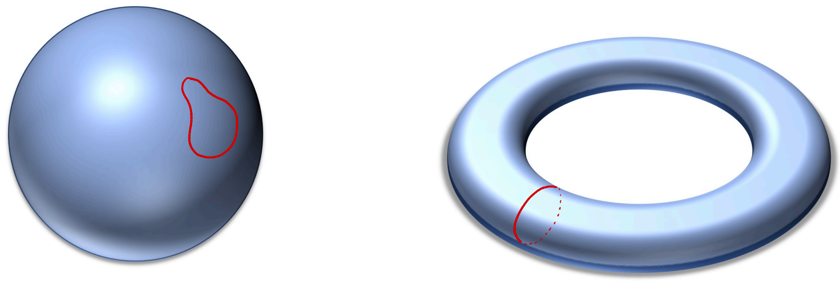


Figure 9. The red lines depict a topologically trivial submanifold C on the left, and a topologically non-trivial sub-manifold C on the right.

Homology

Here we give a brief overview of how the de Rham cohomology, and associated harmonic forms, contain information about the topology of the manifold M.

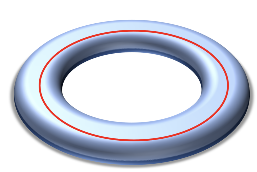
Consider a submanifold $C \subset M$. We'll take this to be a closed submanifold, meaning that it has no boundary

$$\partial C = 0$$

An interesting question is whether C itself can be thought of as the boundary of another manifold, meaning $C = \partial D$. This is a question of topology.

We can see this in two simple examples shown in Figure 9. There we depict two manifolds of dimension two: the sphere $M = \mathbf{S}^2$ and the torus $M = \mathbf{T}^2$. On each we've drawn a one-dimensional submanifold C as a red line. For $C \subset \mathbf{S}^2$, this submanifold is the boundary of a disc $C = \partial D$. For $C \subset \mathbf{T}^2$ there is no such bounding manifold D. This reflects the fact that there is interesting topology in the torus, but not in the sphere.

Indeed, there are actually two different topologically non-trivial submanifolds of the torus: in addition to the circle C shown in Figure 9, there is also the circle C' that winds in the way shown on the right.



The algebraic structure of these topologically non-trivial submanifolds is identical to those of forms. In particular, a boundary of a boundary is always vanishing, which we write as $\partial^2 = 0$. This, obviously, is the strikingly reminiscent of the exterior derivative relation $d^2 = 0$. We use this to define homology groups using ∂ analogous to the cohomology groups that we defined previously using d. The *homology group* $H_p(M)$ is the equivalence class of closed p-dimensional submanifolds that are not themselves the boundary of a $(p+1)$-dimensional manifold. In particular, two submanifolds C_1 and C_2 lie in the same cohomology class if one can be smoothly deformed into the other. In terms of equations, this mean that difference is a boundary,

$$C_1 \sim C_1 \;\text{ if and only if }\; C_1 - C_2 = \partial D$$

The relationship between homology and cohomology is more than just an analogy. The spaces $H_p(M)$ and $H^p(M)$ are dual to each other, and hence isomorphic. This statement, known as *de Rham's theorem*, is not straightforward to prove but it's easy to get some intuition for how it works. Given a closed submanifold $C \subset M$ and a form ω on M we can define a map to the real numbers, given by

$$(C, \omega) = \int_C \omega$$

Strictly speaking, the integral only makes sense if $\dim C = p$ and ω is a p-form. If the form ω has a degree different than $\dim C$ then the pairing is simply said to be zero. In what follows, we will sometimes refer to such a closed submanifold C as a *cycle*.

This pairing has some lovely properties that follow from Stokes' theorem. First, the answer depends only on the equivalence class $[\omega] \in H^p(M)$. To see this, note that

$$(C, \omega + d\alpha) = \int_C (\omega + d\alpha) = \int_C \omega + \int_C d\alpha$$

but the total derivative $\int_C d\alpha = 0$ because $\partial C = 0$.

Conversely, if we consider two submanifolds C_1 and C_2 that can be smoothly deformed into each other, so $C_1 - C_2 = \partial D$, then integrating any closed form ω gives

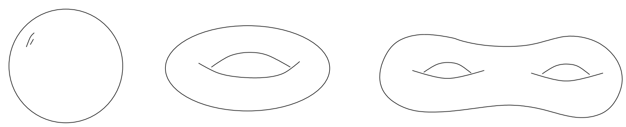
$$\int_{C_1} \omega - \int_{C_2} \omega = \int_{\partial D} \omega = \int_D d\omega = 0$$

We see that the answer only depends on the equivalence class $[C] \in H_p(M)$.

The upshot of these arguments is that the ground states of the supersymmetric sigma model (3.6) are determined by the topology of the target space M. Heuristically, the quantum particle can minimise its energy by spreading its wavefunction over topologically non-trivial submanifolds of M.

Here are some simple examples. The sphere $\mathbf{S}^n$ is boring: its only Betti numbers are $b_0 = 1$ corresponding to constant functions and $b_n = 1$ corresponding to the top form. The torus $\mathbf{T}^n$ is more interesting: it has Betti numbers $b_p = \binom{n}{p}$.

In $n = 2$ dimensions, closed manifolds are known as Riemann surfaces and are labelled by their genus g which counts the number of holes. Here are some examples of manifolds with genus $g = 0$, $g = 1$ and $g = 2$ respectively



The Betti numbers are $b_0 = b_2 = 1$ and $b_1 = 2g$. Each extra hole introduces two new topologically non-trivial 1-manifolds that encircle the hole in different ways.

Finally, I should mention in any logical presentation, homology precedes cohomology. Our physics approach has lead us to introduce these in an inverted order.

3.1.3 The Witten Index and the Chern-Gauss-Bonnet Theorem

In Section 1, we learned that there is something special about the Witten index in supersymmetric quantum mechanics. Recall that this is defined by $\mathrm{Tr}\,(-1)^F e^{-\beta H}$ and counts then number of supersymmetric ground states, up to a sign.

For our supersymmetric sigma model, the Witten index is just the alternating sum of Betti numbers

$$\mathrm{Tr}\,(-1)^F e^{-\beta H} = \chi(M) := \sum_p (-1)^p b_p \tag{3.13}$$

This is perhaps the most famous topological invariant in mathematics: it is known as the *Euler character* of the manifold.

Again, some examples. The sphere $\mathbf{S}^n$ has Euler character

$$\chi(\mathbf{S}^n) = 1 + (-1)^n$$

so is either $\chi(\mathbf{S}^n) = 2$ for n even or $\chi(\mathbf{S}^n) = 0$ for n odd. The torus $\mathbf{T}^n$ always has $\chi(\mathbf{T}^n) = 0$. The 2d Riemann surface of genus g has $\chi(M) = 2 - 2g$.

We can see from our discussion of quantum mechanics why this is a topological invariant. We know that the Witten index is robust against any small change of the parameters in the quantum mechanics. In the present case, that means that if we vary the metric g_{ij}, at least within reason so that we avoid singularities, then the Witten index should remain unchanged. But that means that object $\chi(M)$ defined in (3.13) must be independent of the the choice of metric: it can be depend only on cruder aspects of M, specifically its topology.

Finally, note that the sigma models provide many other examples in which the Witten index vanishes but there are, nonetheless, ground states with $E = 0$. For example, the sigma model on $\mathbf{S}^3$ (or, indeed, any odd dimensional sphere) has $\chi(\mathbf{S}^3) = 0$ but there are two ground states, one the constant function corresponding $b_0 = 1$ and the other the volume form corresponding to $b_3 = 1$. These ground states are also protected by topology, this time by the cohomology rather than the cruder Euler character.

The Path Integral Again

As we saw in Section 2, there is a straightforward description of the Witten index in terms of the path integral. We simply need to calculate

$$\mathcal{I} = \mathrm{Tr}\,(-1)^F e^{-\beta H} = \int \mathcal{D}x\mathcal{D}\psi^\dagger\mathcal{D}\psi\ e^{-S_E[x,\psi,\psi^\dagger]}$$

where Euclidean time τ has period β and both x and ψ are assigned periodic boundary conditions. The Euclidean action is

$$S_E = \oint d\tau\ \frac{1}{2}\,g_{ij}(x)\,\dot{x}^i\dot{x}^j + g_{ij}\psi^{\dagger\,i}\nabla_\tau\psi^i + \frac{1}{4}R_{ijkl}\psi^i\psi^j\psi^{\dagger\,k}\psi^{\dagger\,l}$$

with $\nabla_\tau\psi^i = \dot{\psi}^i + \Gamma^i_{jk}\dot{x}^j\psi^k$. We know that the Witten index is independent of β. We will use this to compute the path integral in the limit $\beta \to 0$. The key idea is that, in this limit, any non-trivial excitations around the Euclidean circle costs an increasing amount of action and so we can restrict ourselves to constant configurations, where the path integral reduces to a normal integral.

Putting these words into formulae, we first rescale the time coordinate to work with $\tau' = \tau/\beta$ so the new time coordinate has period $\tau' \in [0, 1)$. We also rescale $\psi \to \beta^{-1/4}\psi$, leaving us with the Euclidean action

$$S_E = \oint_0^1 d\tau'\ \frac{1}{2\beta}\,g_{ij}(x)\,\dot{x}^i\dot{x}^j + \frac{1}{\sqrt{\beta}}g_{ij}\psi^{\dagger\,i}\nabla_\tau\psi^i + \frac{1}{4}R_{ijkl}\psi^i\psi^j\psi^{\dagger\,k}\psi^{\dagger\,l}$$

where we now see explicitly that in the limit $\beta \to 0$, the modes with $\dot{x}$ and $\dot{\psi}$ non-zero are heavily suppressed. The path integral then reduces to the ordinary integral

$$\mathrm{Tr}\,(-1)^F e^{-\beta H} = \frac{1}{(2\pi)^{n/2}} \int d^n x \; \frac{1}{\sqrt{g}} \int d^n\psi \, d^n\psi^\dagger \exp\left(-\frac{1}{4} R_{ijkl}\psi^i\psi^j\psi^{\dagger\, k}\psi^{\dagger\, l}\right)$$

As in previous examples, we have to saturate the Grassmann integration. But this time, there's clear way to do it. We simply expand out the exponential until we find the right number of fermions.

Since the fermions always come in groups of four, if n is odd the integral necessarily vanishes. We learn that

$$\chi(M) = 0 \;\; \text{if } \dim M = \text{odd}$$

This simple result also follows from the relation $b_p = b_{n-p}$. However, if n is even then the term with $n/2$ powers of the Riemann tensor will saturate the integral.

We start with $n = 2$. In this case, we pull down just a single copy of the Riemann tensor. After doing the Grassmann integrations, we find

$$\mathrm{Tr}\,(-1)^F e^{-\beta H} = \frac{1}{4\pi} \int d^2x \sqrt{g}\; R$$

This is the well known Gauss-Bonnet expression for the Euler character of a Riemann surface.

In general, the Grassmann integrations leave us with $n/2$ copies of the Riemann tensor, contracted with epsilon symbols

$$\mathrm{Tr}\,(-1)^F e^{-\beta H} = \frac{1}{(4\pi)^{n/2}(n/2)!} \int d^n x \frac{1}{\sqrt{g}}\; \epsilon^{i_1 \ldots i_n} \epsilon^{j_1 \ldots j_n} R_{i_1 i_2 j_1 j_2} \ldots R_{i_{n-1} i_n j_{n-1} j_n}$$

This is the generalisation of the Gauss-Bonnet theorem, first proven by Chern in 1944. The contraction of the epsilon symbols results in an expression known as the *Euler density*. The slightly unusual looking $1/\sqrt{g}$ should be thought of as $\sqrt{g} \times \frac{1}{\sqrt{g}} \times \frac{1}{\sqrt{g}}$ with the $\frac{1}{\sqrt{g}}$ factors combining with the epsilon symbols to give tensor densities.

As an example, for $n = 4$ dimensional manifolds the expansion of the Euler density gives

$$\chi(M) = \frac{1}{8\pi^2} \int_M d^4x \sqrt{g}\;\; \left(R_{ijkl}R^{ijkl} - 4R_{ij}R^{ij} + R^2\right)$$

The magic of the Chern-Gauss-Bonnet theorem is that a global topological object, $\chi(M)$, is described in terms of an integral of local data, the Euler density. The magic of supersymmetric quantum mechanics is that it gives a straightforward derivation of this result, with the only real complication the combinatoric factors that arise from Grassmann integration. This is first example where a deep mathematical result can derived in a different way using the path integral. It won't be the last.

3.2 Morse Theory

Our goal in this section is to understand some basic ideas of Morse theory, viewed through the lens of supersymmetric quantum mechanics.

We stick with our $N = 2$ supersymmetric sigma model (3.6), describing a particle moving on a manifold M. The novelty is that we now also include a potential $h(x)$ over the manifold. The resulting supersymmetric theory is a combination of the sigma model and the kind of theories we considered in Section 1.4.1,

$$L = \frac{1}{2} g_{ij}\dot{x}^i\dot{x}^j + ig_{ij}\psi^{\dagger\, i}\nabla_t\psi^j - \frac{1}{4}R_{ijkl}\psi^i\psi^j\psi^{\dagger\, k}\psi^{\dagger\, l} - \frac{1}{2}g^{ij}\frac{\partial h}{\partial x^i}\frac{\partial h}{\partial x^j} - (\nabla_i\partial_j h)\psi^{\dagger\, i}\psi^j \tag{3.14}$$

Note that the final, fermionic term has the opposite sign from that of Section 1.3; this is purely a choice of convention and, as we will see shortly, will bring us in line with definitions used in mathematics. This action is invariant under the supersymmetry transformations

$$\begin{aligned} \delta x &= \epsilon^\dagger\psi - \epsilon\psi^\dagger \\ \delta\psi^i &= \epsilon\left(-i\dot{x}^i + \Gamma^i_{jk}\psi^{\dagger\, j}\psi^k - g^{ij}\frac{\partial h}{\partial x^j}\right) \\ \delta\psi^{\dagger\, i} &= \epsilon^\dagger\left(+i\dot{x}^i + \Gamma^i_{jk}\psi^{\dagger\, j}\psi^k - g^{ij}\frac{\partial h}{\partial x^j}\right) \end{aligned} \tag{3.15}$$

These are a combination of the transformations (1.19) for our original quantum mechanics with a potential and (3.7) for the supersymmetric sigma model.

In the absence of the potential, we know that the ground states of the supersymmetric quantum mechanics spread over cycles of M. However, when we add a the potential h, the wavefunctions get squeezed and, as the potential gets larger, the wavefunctions are increasingly localised at the minima of the potential. We know that the Witten index can't change. But, more strongly, the total number of $E = 0$ ground states doesn't change either and, even in the presence of the potential, is given by the Betti numbers of the manifold.

To see this statement, first note that the supercharges associated to (3.15) are given by

$$Q = \left(g_{ij}\dot{x}^i + i\frac{\partial h}{\partial x^j} \right) \psi^{\dagger\, j} \quad \text{and} \quad Q^\dagger = \left(g_{ij}\dot{x}^i - i\frac{\partial h}{\partial x^j} \right) \psi^j$$

Translated into the geometric language, we have

$$Q \;\longleftrightarrow\; d + dx^i \wedge \partial_i h = d + dh\wedge = e^{-h} d e^{h} \tag{3.16}$$

Similarly

$$Q^\dagger \;\longleftrightarrow\; (d + dh\wedge)^\dagger = e^{h} d^\dagger e^{-h}$$

We saw in Section 3.1.2 that the ground states are determined by the cohomology of Q. But the cohomology when $h \neq 0$ is isomorphic to the cohomology when $h = 0$. We simply take the wavefunctions in the latter case and multiply them by e^{-h}. Indeed, this is the form of the wavefunctions (1.12) that we found back Section 1.2 when considering a particle on a line.

The fact that the number of supersymmetric ground states is independent of h means that something interesting must be going on. Because if we crank up h to be very large, the ground states are localised around the minima of the potential $V = |\partial_i h|^2$. This means that there must be some relationship between these minima and the topology of the manifold. This relationship goes under the name of *Morse theory*.

The minima lie at critical points of h which we will label $x = X$. They obey

$$\frac{\partial h}{\partial x^i}(X) = 0 \quad \text{for all } i = 1, \ldots, n$$

The function h is said to be a *Morse function* if it has the property that the critical points $x = X$ are isolated and non-degenerate. From now on, we'll assume that this is the case.

Consider the situation where we scale the Morse function $h(x) \to \zeta h(x)$, and subsequently send $\zeta \to \infty$. In this limit, the physics is entirely dominated by the critical points of the potential and, at the semi-classical level, the ground state wavefunction is localised at the critical point $x = X$. That's not to say that all critical points are necessarily true $E = 0$ ground states; there may well be tunnelling of the kind that we discussed in Section 2.3 that lifts putative ground states in pairs. But the true ground states must be contained within the set of critical points.

We also need to figure out what's going on with fermions. This is the same calculation that we already met in Section 1.4.1. There, we learned that we should look at the eigenvalues of the Hessian $\partial_i\partial_j h$,

$$(\partial_i\partial_j h)\, e_k^j = \lambda_k\, e_k^j$$

where e_k^j are the eigenvectors and λ_k the eigenvalues, with $k = 1, \ldots, n$. (The index k labels the eigenvectors and eigenvalues and shouldn't be summed over. Note also that we flipped the sign of h in the action (3.14) relative to our discussion in Section 1.4.1, and that shows up as a change of minus sign in this equation relative to (1.27).)

For each negative eigenvalue $\lambda_k < 0$, the final term in (3.14) tells us that we can lower the energy by exciting the corresponding collection of fermions $e_k^j \psi^{\dagger j}$. Meanwhile, for each positive eigenvalue $\lambda_k > 0$, we're better off in the unexcited state.

We define the *Morse index*, $\mu(X)$, to be

$$\mu(X) = \text{The number of negative eigenvalues of } \partial_i\partial_j h(X)$$

We learn that the semi-classical ground state sits in the sector with $\mu(X)$ fermions excited. In other words, the semi-classical ground state at the critical point $x = X$ is a p-form with $p = \mu(X)$.

Already we learn something striking. We can compute the Witten index by simply summing over the critical points X, just as we did in (1.24). The novelty is that we know that, for our supersymmetric sigma model, the Witten index tells us the Euler character of the manifold M. This means that we can compute the Euler character of M from the critical points of a function over M,

$$\chi(M) = \sum_X (-1)^{\mu(X)}$$

In fact, we can say more than this. The total number of critical points may well be more than the total number of $E = 0$ ground states, since states can be lifted in pairs. But the number of critical points can never be smaller than the number of ground states! Suppose that there are m_p critical points X with Morse index $p = \mu(X)$. This can be no less than the number of ground states associated to p-forms, so

$$m_p \geq b_p \tag{3.17}$$

with b_p the Betti number. This is known as the *weak Morse inequalities*.

Nice as this is, it's possible to do better. We can, in fact, recover the original Betti numbers b_p from an understanding of the critical points and the relationships between them. In the rest of this section we explain how.

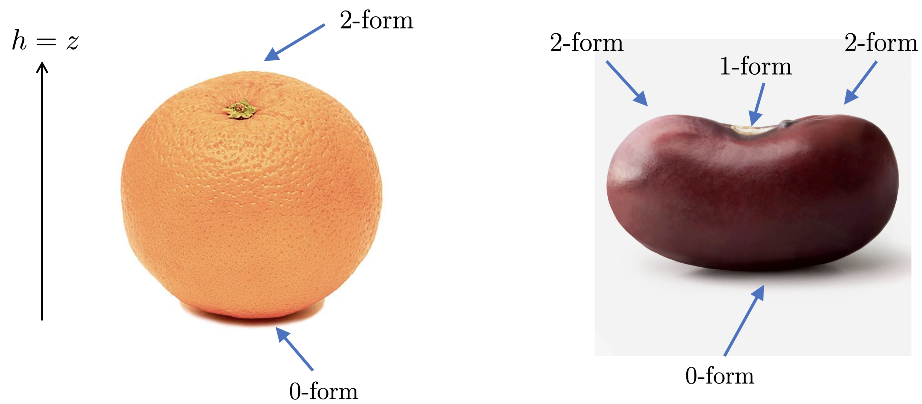


Figure 10. Two shapes, both topologically $\mathbf{S}^2$ with the Morse function given by the height. On the left there are two critical points, on the right there are four. My wife thought it important to point out that these are not drawn to scale.

A Simple Example: The Two Sphere

To illustrate these ideas, we can look at the case of $\mathbf{S}^2$. We know that the Betti numbers are $b_0 = b_2 = 1$ and $b_1 = 0$.

Suppose that we embed $\mathbf{S}^2$ with its round metric in $\mathbb{R}^3$. Then we can consider the height function

$$h = z$$

This is shown in the left-hand side of Figure 10. Clearly there are two critical points of the height function: at the bottom of the sphere where it is a minimum and at the top of the sphere where it is a maximum. The Morse index is $\mu = 0$ and $\mu = 2$ respectively, so from the discussion above we know that these ground states are associated to 0-forms and 2-forms. We also know that ground states localised around these minima must be exact $E = 0$ states.

Now we deform the system. We could change the Morse function h but, for illustrative purposes, it is simplest if we instead change the metric on the sphere. We'll turn it into the bean shape shown in the right-hand side of Figure 10, keeping the same height function $h = z$. This time there are four critical points, one with $\mu = 0$, two at the top with $\mu = 2$, and the saddle point in the middle with $\mu = 1$.

Note that the Euler character hasn't changed,

$$\chi(\mathbf{S}^2) = \sum_X (-1)^{\mu(X)} = 1 + (-1) + 1 + 1 = 2$$

Moreover, the weak Morse inequalities hold, with $m_0 = b_0 = 1$ and $1 = m_1 > b_1 = 0$ and $2 = m_2 > b_2 = 1$.

For the bean shaped metric, we know that two of four semi-classical ground states must be lifted to have $E > 0$. Clearly, it should be the 1-form and some combination of the two 2-forms that gets lifted. Our goal now is to understand how this works, both in the case of the kidney bean and more generally. We will see that much of the technology that we will need has already been covered in the supersymmetric instanton calculation of Section 2.3

3.2.1 Instantons Again

Suppose that our Morse function has r critical points at $x = X_a$ with $a = 1, \ldots, r$, such that

$$\frac{\partial h}{\partial x^i}(X_a) = 0$$

The weak Morse inequalities (3.17) tell us that $r \geq \sum_p b_p$, the total number of supersymmetric ground states (counted without sign). If $r = \sum_p b_p$ then the Morse inequalities are saturated, $m_p = b_p$, and we're done: as we crank up the strength of the potential, the ground state wavefunctions morph smoothly from being spread over cycles of the manifold, to being localised at the critical points. This is the situation depicted by the orange in the previous example.

However if $r > \sum_p b_p$, like in the example of the kidney bean, then there are more critical points than genuine ground states and we have some work to do. Some of the semi-classical ground states associated to critical points must be lifted.

The exact energy eigenstate localised around $x = X_a$ will be denoted as $|\Psi_a\rangle$. Some of these states will persist as zero energy states when all quantum corrections are taken into account. Meanwhile others will be lifted but, as we saw in Section 2.3, will remain as low lying states, with energies of order $e^{-S_{\text{inst}}}$. Our goal is to understand this spectrum.

To this end, we will compute the matrix elements

$$\langle \Psi_a | \, Q \, | \Psi_b \rangle$$

Any state with $E = 0$ must be annihilated by Q so, in general, we expect this matrix to have rank $r - \sum_p b_p$, with the zero eigenvectors the true quantum ground states and the remainder those that are lifted.

As is Section 2.3, it's simpler to compute the related set of matrix elements

$$\langle \Psi_a | \, Q \, | \Psi_b \rangle \approx \frac{\langle \Psi_a | \, [Q, h] \, | \Psi_b \rangle}{h(X_b) - h(X_a)}$$

where, after Wick rotation, the commutator wih the supercharge (3.16) gives

$$[Q, h(x)] = \frac{\partial h}{\partial x^i} \psi^{\dagger \, i}$$

The fact that we have just a single fermion $\psi^\dagger$ in the matrix element means that we'll get non-vanishing contributions if the state $|\Psi_a\rangle$ has one additional fermion excited than $|\Psi_b\rangle$. Or, said differently, if the Morse indices differ by one:

$$\mu(X_a) - \mu(X_b) = 1 \tag{3.18}$$

The difference $\Delta\mu = \mu(X_a) - \mu(X_b)$ is called the *relative Morse index*.

The Instanton Equations

It's clear that we are now back in the realm of the quantum tunnelling calculations that we performed in Sections 2.2 and 2.3. To start, we can study the instantons in a sigma model with potential. Focussing just on the bosonic fields for now, the action (3.14) in Euclidean time is

$$S_E = \int d\tau \; \frac{1}{2} g_{ij} \dot{x}^i \dot{x}^j + \frac{1}{2} g^{ij} \frac{\partial h}{\partial x^i} \frac{\partial h}{\partial x^j}$$

where now $\dot{x} = dx/d\tau$. We can write this by completing the square as

$$S_E = \int d\tau \; \frac{1}{2} g_{ik} \left(\frac{dx^i}{d\tau} \mp g^{ij} \frac{\partial h}{\partial x^j} \right) \left(\frac{dx^k}{d\tau} \mp g^{kl} \frac{\partial h}{\partial x^l} \right) \pm \frac{dx^i}{d\tau} \frac{\partial h}{\partial x^i}$$

For the class of configurations that interpolate from $x(\tau) = X_b$ at time $\tau \to -\infty$ to $x(\tau) = X^a$ at time $\tau \to +\infty$, the action is minimised for configurations that obey the instanton equations

$$\frac{dx^i}{d\tau} = \pm g^{ij} \frac{\partial h}{\partial x^j}$$

Solutions to the equation with the $+$ sign are instantons; those with the $-$ sign are anti-instantons. The action is given by

$$S_{\text{inst}} = \pm(h(X_a) - h(X_b))$$

Since we want the action to be positive definite, we should pick the instanton solution when $h(X_a) > h(X_b)$ and the anti-instanton when $h(X_a) < h(X_b)$.

From our previous calculation in Section 2.3, we know that the fermion zero modes play a crucial role in supersymmetric instanton calculations. So our next question: how many fermi zero modes does our instanton have? To answer this, we look at the linearised fermion equation of motion. Here "linearised" means that we drop the Riemann tensor term in (3.14), and the connection term in ∇_t. In Euclidean time, the linearised equations are

$$D\psi^i := \frac{d\psi^i}{d\tau} + g^{ij}\nabla_j\partial_k h\,\psi^k = 0 \tag{3.19}$$

and

$$D^\dagger\psi^{\dagger\,i} := -\frac{d\psi^{\dagger\,i}}{d\tau} + g^{ij}\nabla_j\partial_k h\,\psi^{\dagger\,k} = 0$$

We want to know how many solutions each of these equations have in the background of an instanton. In fact, we really just want to know the difference between the number of solutions to these equations. This is because if both D and $D^\dagger$ have zero modes then they will most likely be lifted by the non-linear terms in the action. And, indeed, generically, this will happen. However if there are unpaired zero modes of, say $D^\dagger$, then these must be saturated in some other way in the path integral. This prompts our interest in the index

$$\mathcal{I}(D) = \dim \operatorname{Ker} D - \dim \operatorname{Ker} D^\dagger$$

where $\operatorname{Ker} D$ is the kernel of D, the space of solutions to $D\psi^i = 0$. Clearly, the index counts the number of unpaired zero modes of the instanton.

Furthermore, because our matrix element $\langle \Psi_a|\,[Q,h]\,|\Psi_b\rangle$ contains just a single fermion $\psi^\dagger$, we're only going to get non-zero contributions from instantons which have one more zero mode for $\psi^\dagger$ than for ψ, namely

$$\mathcal{I}(D) = -1 \quad \text{or, equivalently} \quad \mathcal{I}(D^\dagger) = +1$$

We're always guaranteed to get some fermi zero modes from acting with broken supersymmetry. If we start from an instanton configuration, and originally set all fermions to zero, then acting with the supersymmetry transformations (3.15) (and remembering to Wick rotate to Eulidean time $\tau = it$), we have

$$\delta\psi^i = \epsilon\left(\frac{dx^i}{d\tau} - g^{ij}\frac{\partial h}{\partial x^j}\right)$$
$$\delta\psi^{\dagger\, i} = -\epsilon^\dagger\left(\frac{dx^i}{d\tau} + g^{ij}\frac{\partial h}{\partial x^j}\right)$$

Because we want a $\psi^\dagger$ zero mode, rather than a ψ zero mode, we should focus on configurations that obey

$$\frac{dx^i}{d\tau} = g^{ij}\frac{\partial h}{\partial x^j} \tag{3.20}$$

That is, we should focus on instantons rather than anti-instantons. This means that we should look at configurations that start, at $\tau \to -\infty$ at X_b and end up at $\tau \to +\infty$ at X_a, with $h(X_a) > h(X_b)$. If we think of $h(x)$ as a height function, these are trajectories that go up, rather than down.

Instantons and the Relative Morse Index

We've now played the "Grassmann integration" card twice: once in (3.18) to argue that we should get contributions only between vacua that have relative Morse index 1, and again above to argue that we should only get contributions from instantons with $\mathcal{I}(D^\dagger) = 1$. Clearly we need these two different arguments to coincide. Happily they do because of the following result:

Claim: The index of $D^\dagger$ is equal to the relative Morse index

$$\mathcal{I}(D^\dagger) = \mu(X_a) - \mu(X_b)$$

Proof(ish): Here we give a sketch of the proof of this statement. The operator D, defined in (3.19), acts on an n-dimensional space of fermions ψ^i and takes the form

$$D^\dagger = \frac{d}{d\tau} - \text{Hess}[h]$$

where Hess$[h]$ is the $n \times n$ Hessian matrix

$$\text{Hess}[h]^i{}_j = g^{ik}\nabla_k\partial_j h$$

When evaluated at the critical points $x = X_a$, this coincides with the Hessian that we previously used to define the Morse index. But the equation above provides an

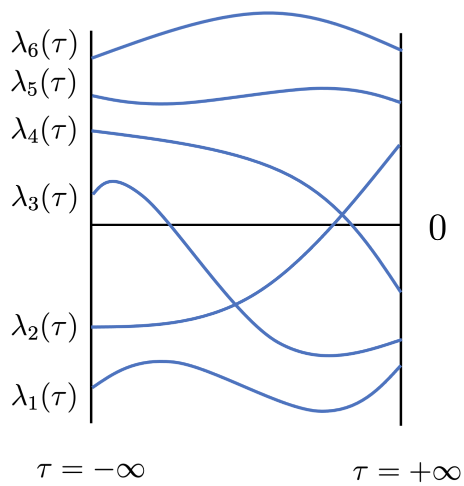


Figure 11. An example of the spectral flow of eigenvalues between two vacua, albeit one in which creative licence has trumped mathematical precision. The level crossings shown above can occur but they are not generic. In a more typical trajectory, the order of eigenvalues remains the same under spectral flow.

extension of the definition of the Hessian to each point along the instanton trajectory $x(\tau)$. As we move along this trajectory, the eigenvalues and orthonormal eigenvectors will smoothly evolve,

$$\mathrm{Hess}[h(\tau)]\,\mathbf{e}_k(\tau) = \lambda_k(\tau)\mathbf{e}_k(\tau) \tag{3.21}$$

There is no sum over $k = 1, \ldots, n$ in this equation which labels the eigenvectors and eigenvalues. (The eigenvectors $\mathbf{e}_k$ have an additional $i = 1, \ldots, n$ index which is the vector index and is suppressed in the equation above.)

We can now follow the n eigenvalues λ_i as we move from one critical point to another. This is known as *spectral flow*, and an example is shown in Figure 11. The number of negative eigenvalues at $\tau = -\infty$ is the Morse index $\mu(X_a)$; the number of negative eigenvalues at $\mu = +\infty$ is $\mu(X_b)$.

Of particular interest are those eigenvalues which start negative and end up positive, or vice versa. The difference between those that cross in one direction, and those that cross in the other, is the relative Morse index $\mu(X_a) - \mu(X_b)$. The example shown in the Figure 11 has one more negative eigenvalue at the end than at the beginning which means that the corresponding instanton interpolates between two value with relative Morse index $\mu(X_a) - \mu(X_b) = +1$.

To solve the Dirac equation $D^\dagger\boldsymbol{\psi} = 0$, we simply expand the fermions in terms of the eigenvectors, writing $\psi^i(\tau) = \sum_k c_k(\tau)e^i_{\ k}(\tau)$. We insert this ansatz into the Dirac equation and, using the orthogonality of eigenvectors $g_{ij}e^i_{\ k}e^j_{\ l} = \delta_{kl}$, we have

$$\left(\frac{d}{d\tau} - \lambda_k(\tau)\right) c_k(\tau) = -\sum_l g_{ij}e^i_{\ k}\dot{e}^j_{\ l}c_l \tag{3.22}$$

with no sum over k. First, let's suppose that we can ignore the term on the right-hand side. Then the equation has a straightforward solution,

$$c_k(\tau) = A_k \exp\left(+\int d\tau'\ \lambda_k(\tau')\right)$$

But this is a normalisable solution to the Dirac equation only if $\lambda(\tau) < 0$ for $\tau \to +\infty$ and $\lambda(\tau) > 0$ for $\tau \to -\infty$. That is, we get a solution for every eigenvalue that flips from positive to negative. Meanwhile, the same analysis shows that every eigenvalue that goes the other way, from negative to positive, gives a solution to $D\boldsymbol{\psi} = 0$. This is precisely what we wanted to show, namely

$$\mathcal{I}(D^\dagger) = \mu(X_a) - \mu(X_b)$$

That leaves us with the question of why it's legal to ignore the term on the right-hand side of (3.22). This is where we get to the "ish" part of proofish. The term captures how the eigenvectors twist as we move along the instanton trajectory due to the Levi-Civita connection. (In (3.22), this takes the form of a Berry connection.) This connection doesn't introduce any further topology into the game and it is a true fact that it doesn't change the index, albeit not a fact that I will demonstrate here. In acknowledgement of this slipshod approach, I'll replace the traditional QED box used at the end of a proof with something more wonky.

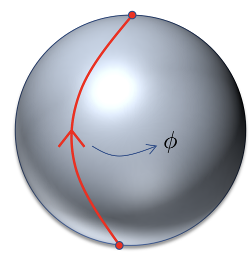
It turns out that for background configurations with $\mathcal{I}(D^\dagger) > 0$ we generically have $\mathrm{Ker}\, D = 0$, so that $\mathcal{I}(D^\dagger) = \dim \mathrm{Ker}\, D^\dagger$. (For example, the situation shown in Figure 11 is not generic.) We will assume that this is the case moving forwards.

The calculation above also tells us about the bosonic collective coordinates of an instanton. Suppose that we find a solution $x(\tau)$ to the instanton equation (3.20). To see if this solution has any collective coordinates, we can look at variations $x(\tau)+\delta x(\tau)$ and see if $\delta x(\tau)$ satisfies the linearised instanton equation,

$$\frac{d}{d\tau}\delta x^i - g^{ij}\nabla_j\partial_k h\ \delta x^k = 0$$

But this coincides with the Dirac equation $D^\dagger \delta \mathbf{x} = 0$ whose solutions we've just counted. The upshot is that the number of bosonic collective coordinates is equal to the number of fermi zero modes, and both are counted by the relative Morse index. A slicker way of saying this is to note that bosonic and fermionic zero modes are related by the unbroken supersymmetry $Q^\dagger$ in the background of an instanton.

For our purposes, we want to consider instantons that interpolate between critical points with relative Morse index 1. Here the sole bosonic collective coordinate is the obvious one: the time τ_1 at which the instanton does its business of interpolating from one critical point to the other. This is the collective coordinate that we met previously in Section 2.2.



Although not of immediate utility, we can also get a feel for where the other bosonic collective coordinates may come from when $\Delta\mu > 1$. Consider the height Morse function on the round sphere $\mathbf{S}^2$. We know that there are two critical points at the south and north pole with Morse index 0 and 2 respectively. Correspondingly, the instanton that interpolates from the south to the north pole has two collective coordinates: one is the time τ_1 at which the instanton makes the jump, the other is the angle ϕ of the trajectory as shown in the figure. In this example, the second collective coordinate is obvious because it arises due to a symmetry. But the arguments above tell us that, perhaps surprisingly, this second collective coordinate persists even when we deform the sphere, or potential, so that there's no longer a rotational symmetry.

Completing the Instanton Computation

The rest of our instanton computation proceeds in exactly the same manner as that of Section 2.3. Our final answer is the obvious generalisation (2.41): for vacua $|\Psi_a\rangle$ and $|\Psi_b\rangle$, whose Morse index differs by 1, we have

$$\langle \Psi_a | \, Q \, | \Psi_b \rangle = \frac{e^{-S_{\rm inst}}}{\sqrt{2\pi}} \sum_\gamma n_\gamma \tag{3.23}$$

Here the sum is over all distinct instantons γ and $n_\gamma = \pm 1$ is a sign that comes from computing the determinants

$$n_\gamma = \frac{\det D}{\sqrt{\det D^\dagger D}} = \pm 1$$

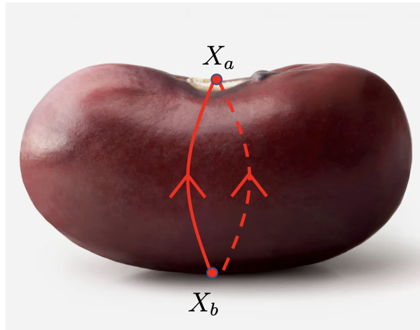


Figure 12. Two different instanton trajectories interpolating between X_b with $\mu(X_b) = 0$ and X_a with $\mu(X_a) = 1$.

We met this sign before in Section 2.3 where we confessed that it is a little tricky to fix. Now it is time to make good on our promise of explaining where it comes from.

Let's start in the vacuum $|\Psi_a\rangle$ localised at the end of the instanton trajectory at X_a. There are $\mu(X_a)$ negative eigenvalues of the Hessian and their eigenvectors span a $\mu(X_a)$-dimensional space that we call V_a. The ground state $|\Psi_a\rangle$ is associated to a $\mu(X_a)$-form and this induces an orientation on V_a.

The tangent to the instanton trajectory at X_a lies in the space V_a of negative eigenvectors. Let us call this tangent vector v. Generically, v will coincide with the eigenvector with largest negative eigenvalue λ_k (or smallest $|\lambda_k|$) since this is usually the unique direction for which the eigenvalue flips sign by the time we reach X_b at the bottom. We denote the subspace of V_a that is orthogonal to v as $\tilde{V}_a$. There is a natural orientation on $\tilde{V}_a$ that comes from taking the interior product $\iota_v \Psi_a$.

Now we propagate the space $\tilde{V}_a$ along the instanton trajectory γ. We can do this, for example, by following the eigenvectors $\mathbf{e}_k(\tau)$ of (3.21) corresponding to those eigenvalues that remain negative along the entire journey.

By the time we reach the end of the trajectory, the orientation on $\tilde{V}_a$ that we started with gives an orientation on V_b, the space of negative eigenvectors of the Hessian at X_b. But there is a different way to define an orientation on V_b, which is that induced by the ground state $|\Psi_b\rangle$ or, more precisely, the corresponding $\mu(X_b)$-form. The question is: do these two ways of defining an orientation coincide? If they do, we take $n_\gamma = 1$. If they do not, we take $n_\gamma = -1$.

For example, consider again the deformed bean-shaped sphere shown in Figure 12. There are two instanton trajectories that interpolate from the minimum X_b at the bottom to the saddle point X_a in the middle. At X_a, the tangent vectors to the two different instanton trajectories point in different directions, and that means that each instanton trajectory induces opposite orientations $\iota_v \Psi_a$ on $\tilde{V}_a$. Correspondingly, one instanton will have $n_\gamma = +1$ and the other $n_\gamma = -1$, and the two cancel out in (3.23). This same argument explains why the ground states are not lifted for the double well on a circle that we discussed in Section 2.3.4.

3.2.2 The Morse-Witten Complex

Let's recap. A Morse function gives us a collection of critical points. There are m_p critical points X with Morse index $p = \mu(X)$ and, associated to each, there is an energy eigenstate $|\Psi_a\rangle$ and an associated p-form. These can be thought of as a basis for an m_p-dimensional space that we will call C^p.

Not all the energy eigenstates $|\Psi_a\rangle$ have vanishing energy or, equivalently, not all of them are annihilated by Q. But, as we saw in Section 2.3, they all have energy that is, at most, of order $e^{-S_{\text{inst}}}$. The tunnelling calculation that we've just done shows that,

$$\langle\Psi_a|\, Q\, |\Psi_b\rangle = \frac{e^{-S_{\text{inst}}}}{\sqrt{2\pi}} \sum_\gamma n_\gamma \quad \text{whenever } \mu(X_a) - \mu(X_b) = 1$$

If we think about Q as acting within this space of states, we can insert an "almost resolution" of the identity $\mathbb{1} \approx \sum_a |\Psi_a\rangle\langle\Psi_a|$ to get

$$\begin{aligned} Q|\Psi_b\rangle &= \sum_a \langle\Psi_a|\, Q\, |\Psi_b\rangle\, |\Psi_a\rangle \\ &= \sum_{a\,:\,\mu_a = \mu_b + 1} \sum_\gamma \frac{n_\gamma}{\sqrt{2\pi}}\, e^{-(h(X_a) - h(X_b))}\, |\Psi_a\rangle \end{aligned}$$

Here the "almost resolution" of $\mathbb{1}$ is because we've neglected all higher energy states. But their overlap with the low lying states $|\Psi_a\rangle$ is exponentially suppressed and can be ignored. This means that we're left with an expression for the action of Q among the critical points. Neither the factor of $\sqrt{2\pi}$, nor the instanton action, are important for our present purposes and can be absorbed into the normalisation of the states. We then have

$$Q|\Psi_b\rangle = \sum_{a\,:\,\mu_a = \mu_b + 1} \sum_\gamma n_\gamma\, |\Psi_a\rangle$$

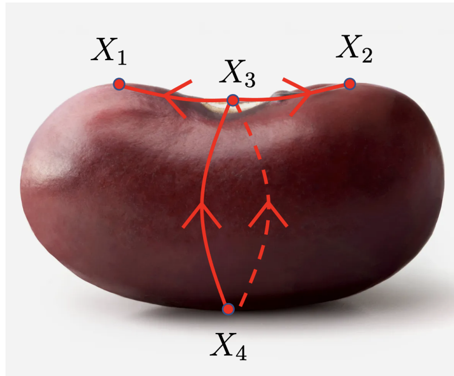


Figure 13. The four different instanton trajectories on the bean.

This is a map $Q : C^p \to C^{p+1}$. More abstractly, it can be viewed as a map between spaces of critical points. And, importantly, it satisfies $Q^2 = 0$. This means that we can define a chain complex (strictly a cochain complex), known as the *Morse-Witten complex*, or sometimes the *Morse-Smale-Witten complex*

$$0 \longrightarrow C^0 \xrightarrow{Q} C^1 \xrightarrow{Q} \ldots \xrightarrow{Q} C^n \xrightarrow{Q} 0$$

The cohomology of Q describes the $E = 0$ ground states of the system or, equivalently, the Betti numbers.

As an example we can look once again at the bean shaped manifold shown in Figure 13. We've already seen that the two instantons taking us from X_4 to X_3 cancel out, leaving us with

$$Q|\Psi_4\rangle = 0$$

This means that $|\Psi_4\rangle$ is a true ground state of the system. There are also two instanton trajectories emanating from X_3, one to each of the peaks at X_1 and X_2. These trajectories have different orientations, meaning

$$Q|\Psi_3\rangle = |\Psi_1\rangle - |\Psi_2\rangle$$

Finally, we have $Q|\Psi_1\rangle = Q|\Psi_2\rangle = 0$ as both states are top forms. The true ground states lies in Q-cohomology and there are two of them: the 0-form $|\Psi_0\rangle$ and the 2-form $|\Psi_1\rangle + |\Psi_2\rangle$. This reproduces the cohomology of $\mathbf{S}^2$.

3.3 The Atiyah-Singer Index Theorem

We now turn to a second application of supersymmetric quantum mechanics. We will study a version of supersymmetric quantum mechanics that yields the *Atiyah-Singer index theorem*. Before introducing the physics, we first explain what problem the index theorem addresses.

In n dimensions, where n is even, a Dirac spinor χ has $2^{n/2}$ components. In flat space, the free Dirac equation reads

$$\not{\partial}\chi = \gamma^a \partial_a \chi = 0 \tag{3.24}$$

Here the gamma matrices obey the usual Clifford algebra

$$\{\gamma^a, \gamma^b\} = 2\delta^{ab} \qquad a, b = 1, \ldots, n \tag{3.25}$$

The only solutions to (3.24) are constant spinors. That's a bit boring and, in $\mathbb{R}^n$, more than a bit non-normalisable. Things get more interesting when the fermion lives on a curved manifold M. To describe this situation, we first introduce vielbeins

$$g_{ij} e_a^{\ i} e_b^{\ j} = \delta_{ab}$$

The i, j indices are raised and lowered using the metric g_{ij} while the tangent space indices a, b are raised and lowered using δ_{ab}. (See the lectures on General Relativity for more details.) The Dirac equation then takes the form

$$\not{D}\chi = \gamma^a e_a^{\ i} D_i \chi = 0 \tag{3.26}$$

with the covariant derivative given by

$$D_i = \partial_i + \frac{1}{2}(\omega_i)^{bc} S_{bc} \tag{3.27}$$

Here S_{ab} are the generators of $SO(n)$ (strictly $\mathrm{Spin}(n)$) in the spinor representation

$$S_{ab} = \frac{1}{4}[\gamma_a, \gamma_b]$$

Meanwhile $(\omega_i)^{ab}$ is the *spin connection*, defined by

$$(\omega_i)^a_{\ b} = \Gamma^a_{cb} e^c_{\ i} = e^a_{\ j} \nabla_i e_b^{\ j}$$

We can then ask: how many solutions there are to the Dirac equation (3.26)? This is where the Atiyah-Singer index theorem comes in. It relates the number of solutions to the Dirac equation to the topology of the underlying manifold. The purpose of this section is to give a physics derivation of the index theorem from supersymmetric quantum mechanics.

3.3.1 The $N = 1$ Sigma Model

Our quantum mechanics of choice has half the supersymmetry of the models that we've considered until now in this section. That is, we will have $N = 1$ supersymmetry with a single real supercharge Q. We met some simple theories of this kind already in Section 1.4.3.

With $N = 1$ supersymmetry, the sigma model action (3.6) is replaced by something that, at first glance, appears much simpler,

$$L = \int dt\ \frac{1}{2}\, g_{ij}(x)\, \dot{x}^i \dot{x}^j + \frac{i}{2} g_{ij} \psi^i \nabla_t \psi^j \tag{3.28}$$

The key difference is that the Grassmann variables are now Majorana modes

$$\psi^{\dagger\, i} = \psi^i$$

We no longer have the Riemann tensor interaction term, but the Levi-Civita connection still shows up, as before, in the kinetic term for the fermions,

$$\nabla_t \psi^i = \frac{d\psi^i}{dt} + \Gamma^i_{\ jk} \frac{dx^j}{dt} \psi^k \tag{3.29}$$

Although the action (3.28) has fewer interaction terms, it also has less symmetry. In particular, because the fermions are real we no longer have the $U(1)$ symmetry that rotated the phase of the fermions. For our $N = 2$ sigma model (3.6), this symmetry ensured that the energy eigenstates had a fixed number, p of excited fermions. Now that we no longer have this symmetry, we expect the energy eigenstates to involve a mixture of different fermions. The only protection we have comes from the $(-1)^F$ symmetry that categorises states into $\mathcal{H}_B$ with an even number of fermions and $\mathcal{H}_F$ with an odd number.

We've already seen in Section 1.4.3 what emerges when we quantise the fermions so we will be brief here. The canonical anti-commutation relations are

$$\{\psi^i, \psi^j\} = g^{ij}$$

which is closely related to the Clifford algebra (3.25): the relationship between the fermions and gamma matrices involves a vielbein to accommodate the presence of the metric: $\psi^i = e_a^{\ i} \gamma^a$. This is telling us that, upon quantisation, the fermions will give $2^{n/2}$ states which can be viewed as a Dirac spinor χ living on the manifold M. While quantisation of the $N = 2$ sigma model (3.6) gave us p-forms over the manifold, now we have a spinor.

The action (3.28) is invariant under the supersymmetry transformations

$$\delta x^i = \epsilon\psi^i \quad \text{and} \quad \delta\psi^i = -\epsilon\dot{x}^i$$

with the associated supercharge

$$Q = \frac{1}{2} g_{ij}\psi^i \dot{x}^i$$

The ground states of the quantum mechanics once again obey $Q|\chi\rangle = 0$. We can ask how this equation translates into the geometric language. The answer is clear: the fermion ψ^i in the supercharge is replaced by a gamma matrix, while the mechanical momentum $\dot{x}$ is replaced by the appropriate covariant derivative, so that $Q = i\,\not{D}$. The upshot is that ground states of the quantum mechanics are given by solutions to the Dirac equation

$$\not{D}\chi = 0 \tag{3.30}$$

where the covariant derivative is (3.27), as appropriate for a spinor on a curved manifold M. The Hamiltonian is then $H = Q^2 = -\not{D}^2$.

We can now see why this quantum mechanics is of interest. The ground states are specified by solutions to the Dirac equation which is exactly what we want to count. Moreover, we know how to count ground states in supersymmetric quantum mechanics, at least up to sign: we use the Witten index.

To get an expression for the index, we first need to figure out which states sit in $\mathcal{H}_B$ and which in $\mathcal{H}_F$. As we already saw in Section 1.4.3, this has a particularly nice interpretation in terms of the spinor. Because we are an even dimension n, the Dirac spinor decomposes into two Weyl spinors which are eigenspinors of

$$\hat{\gamma} = i^{n/2}\gamma^1 \dots \gamma^n$$

This obeys $\hat{\gamma}^2 = 1$ and $\{\hat{\gamma}, \gamma^i\} = 0$, and is the generalisation of the "γ^5" matrix in four dimensions. This is the operator that determines whether a given state lies in $\mathcal{H}_B$ or $\mathcal{H}_F$ via the identification

$$\hat{\gamma} = (-1)^F$$

We can always pick a basis of gamma matrices that are block off-diagonal, so that we have

$$\hat{\gamma} = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{pmatrix}$$

where each entry is a $2^{n/2-1}$ dimensional matrix. The Dirac operator then takes the form

$$\not{D} = \begin{pmatrix} 0 & \mathcal{D}^\dagger \\ \mathcal{D} & 0 \end{pmatrix} \tag{3.31}$$

where $\mathcal{D} : \mathcal{H}_B \to \mathcal{H}_F$ and $\mathcal{D}^\dagger : \mathcal{H}_F \to \mathcal{H}_B$. The Witten index coincides with the index of the operator $\mathcal{D}$,

$$\mathrm{Tr}\,(-1)^F e^{-\beta H} = \mathcal{I}(\mathcal{D}) := \dim \mathrm{Ker}\, \mathcal{D} - \dim \mathrm{Ker}\, \mathcal{D}^\dagger \tag{3.32}$$

This is also the quantity of relevance to the Atiyah-Singer index theorem. Our next task is to compute it. But, by now, our strategy for this should be clear: we turn to the path integral.

3.3.2 The Path Integral Again

The same argument that we used in Section 3.1.3 when deriving the Chern-Gauss-Bonnet theorem tells us that the path integral localises on constant configurations $\dot{x}^i = \dot{\psi}^i = 0$. We'll pick a constant configuration and expand

$$x^i(\tau) = x^i_0 + \delta x^i(\tau) \quad \text{and} \quad \psi^i(\tau) = \psi^i_0 + \delta\psi^i(\tau)$$

We then compute the path integral by performing a Gaussian integration over the fluctuations δx and $\delta\psi$. Our life is made easier if we work in normal coordinates in which

$$g_{ij}(x) = \delta_{ij} - \frac{1}{3} R_{ijkl}(x_0)\delta x^k \delta x^l$$

and

$$\delta\Gamma^i{}_{jk}(x) = \partial_l \Gamma^i{}_{jk}(x^0)\,\delta x^l = -\frac{1}{3}\left(R^i{}_{jkl}(x_0) + R^i{}_{kjl}(x_0)\right)\delta x^l$$

To quadratic order, the Euclidean action then becomes

$$S_E = \frac{1}{2}\int d\tau\ \delta_{ij}\left(-\delta x^i \frac{d^2}{d\tau^2}\delta x^j + \delta\psi^i \frac{d}{d\tau}\psi^j\right) - \frac{1}{2} R_{ijkl}\,\psi^i_0\psi^j_0\,\delta x^k\,\frac{d\delta x^l}{d\tau}$$

Performing the Gaussian integral periodic boundary conditions, we have

$$Z = \int \mathcal{D}\delta x \mathcal{D}\delta\psi\ e^{-S_E} = \sqrt{\frac{\det{}'(\delta_{ij}\partial_\tau)}{\det{}'(-\delta_{ij}\,\partial^2_\tau + \Omega_{ij}\,\partial_\tau)}} = \sqrt{\frac{1}{\det{}'(-\delta^i_j\,\partial_\tau + \Omega^i{}_j)}}$$

The fermionic determinant in the numerator now sits under square root, reflecting the fact that the fermions are real. (It could be better thought of as a Pfaffian.) Both determinants have had their zero modes truncated since these correspond to the integrals over x_0 and ψ_0, both of which we will do explicitly below.

After the small cancellation seen above, we're left with the task of computing the determinant operator involving the matrix

$$\Omega^i_{\ j} := \delta^{ip} R_{pjkl}(x_0)\, \psi_0^k \psi_0^l \tag{3.33}$$

This should be thought of as an $n \times n$ matrix that depends both on the point x_0 and on the background fermion ψ_0. (It may seem odd to think about a matrix as depending on a background Grassmann parameter like ψ_0; the meaning of this should become clearer below when we think about what we're going to do with this matrix.) This is an anti-symmetric matrix and we can choose a basis in which it takes block diagonal form

$$\Omega^i_{\ j} = \begin{pmatrix} \mathbb{W}_1 & & & \\ & \mathbb{W}_2 & & \\ & & \ddots & \\ & & & \mathbb{W}_{n/2} \end{pmatrix} \quad \text{with} \quad \mathbb{W}_a = \begin{pmatrix} 0 & \omega_a \\ -\omega_a & 0 \end{pmatrix}$$

The eigenvalues are $\pm i\omega_a$ and these depend on both x_0 and on ψ_0. We can also diagonalise the derivative term simply by working in a Fourier basis around the circle. Since we know that the end result for the Witten index must be independent of β, we'll take advantage of this and work with $\beta = 1$. We then have

$$\delta x^i(\tau) \sim e^{ik\tau} \quad \text{with} \quad k = 2\pi p \quad \text{and} \quad p \in \mathbb{Z}$$

This means that the eigenvalues of the bosonic fluctuation operator are $i(k \pm \omega)$. Restricting to any given 2×2 matrix $\mathbb{W}$, we have

$$\begin{aligned} \sqrt{\det{}'(-\partial_\tau + \mathbb{W})} &= \prod_{p \neq 0} \, (2\pi i p + i\omega)^{1/2} \, (2\pi i p - i\omega)^{1/2} \\ &= \prod_{p=1}^{\infty} (2\pi i p)^2 \left[1 + \left(\frac{i\omega}{2\pi p} \right)^2 \right] \end{aligned}$$

We've met each of these products before. The second, convergent product is given by (2.6)

$$\prod_{p=1}^{\infty} \left[1 + \left(\frac{i\omega}{2\pi i p} \right)^2 \right] = \frac{\sinh(i\omega/2)}{i\omega/2}$$

The first, divergent, product can be treated using zeta function regularisation as in (2.7) and gives

$$\prod_{p=1}^{\infty} (2\pi i p)^2 = -i$$

Putting this together, we have an expression for the partition function after integrating over the fluctuations,

$$Z = (-i)^{n/2} \prod_{a=1}^{n/2} \frac{i\omega_a/2}{\sinh(i\omega_a/2)}$$

We're now left just with the zero mode integrations: including these gives us the expression for the Witten index and hence the index of the Dirac operator

$$\mathcal{I}(\mathcal{D}) = (-i)^{n/2} \int \prod_{i=1}^{n} \frac{dx_0^i}{\sqrt{2\pi}} d\psi_0^i \prod_{a=1}^{n/2} \frac{i\omega_a/2}{\sinh(i\omega_a/2)}$$

The next step is to do the fermion zero mode integration. The idea here is that each ω_a depends quadratically on the fermion zero modes. We should expand each term in the product,

$$\frac{\omega/2}{\sinh(\omega/2)} = 1 - \frac{\omega^2}{24} + \frac{7\omega^4}{5760} + \ldots$$

Because the function is even, the expansion contains only even powers of ω and hence the fermionic variables ψ_0 come in groups of four. The fermionic integration picks out the term that saturates the Grassmann integration. We learn that the index $\mathcal{I}(\mathcal{D})$ will be non-vanishing only on manifolds whose dimension n is a multiple of four. (This also eliminates the factors of i in the expression for $\mathcal{I}(\mathcal{D})$.)

There is a more geometric way to think about this. Instead working with fermions, we turn again to forms. (The fact that our fermions don't yield forms upon quantisation is irrelevant here: this is just a formal trick.) We introduce the curvature 2-form

$$\mathcal{R}^a{}_b = R^a{}_{bcd} \hat{\theta}^c \wedge \hat{\theta}^d$$

where $\hat{\theta} = e^a{}_i dx^i$ are a basis of one-forms. (see the lectures on General Relativity for more details.) Clearly $\mathcal{R}$ has the has the same formal structure as the fermionic matrix Ω that we met before. This means that we can equally well write

$$\mathcal{I}(\mathcal{D}) = \int_M \hat{A}(M) \quad \text{with} \quad \hat{A}(M) = \frac{1}{(2\pi)^{n/2}} \sqrt{\det\left(\frac{\mathcal{R}/2}{\sinh(\mathcal{R}/2)}\right)}$$

This is the Atiyah-Singer index theorem. The expression $\hat{A}(M)$ is referred to as the *A-roof genus* or *A-hat genus* of M (or, more correctly, the tangent bundle of M). The expression should be viewed as expanding out the determinant until we find a top form that can be integrated over the manifold M. The terms that appear in this expansion are

$$\int_M \hat{A}(M) = 1 - \frac{p_1}{24} + \frac{1}{16}\left(\frac{7}{360}p_1^2 - \frac{1}{90}p_2\right) + \ldots$$

where the various terms in the expansion arise for manifolds of dimension $n = 0, 4, 8$ and so on. The p_i are known as *Pontryagin numbers* and can be expressed as integrals of the curvature 2-form over the manifold M. For a manifold of dimension $\dim(M) = 4$, we have

$$\int_M \hat{A}(M) = -\frac{p_1}{24} \quad \text{with} \quad p_1 = -\frac{1}{8\pi^2}\int_M \operatorname{tr}\mathcal{R}^2 \tag{3.34}$$

The simplest examples of 4-manifolds are the torus and sphere, which have $p_1(\mathbf{T}^4) = p_1(\mathbf{S}^4) = 0$. This tells us that the index of the Dirac operator vanishes. For the torus $\mathbf{T}^4$, with periodic boundary conditions for spinors, this is because both $\mathcal{D}$ and $\mathcal{D}^\dagger$ have zero modes. For the sphere $\mathbf{S}^4$, it is because there are no zero modes.

To find a compact manifold M with $p_1 \neq 0$, we need to turn to something more exotic. A nice example is provided by the manifold known as $M = K3$, which can be viewed as a smooth quartic surface in $\mathbf{CP}^3$. This is the only non-trivial Calabi-Yau 4-manifold and has $p_1(K3) = -48$.

The factor of $1/24$ in (3.34) is telling us something interesting. Because the left-hand side is counting something, the right-hand-side must be an integer. This suggests that the integral p_1 must be a multiple of 24. In fact, things are a little more subtle. The thing that we're counting is the number of solutions to the Dirac equation and to pose this question at all, we must be able to put Dirac spinors on the manifold M. This, it turns out, is not possible for all manifolds. Those for which is is possible are called *spin manifolds*. And for any orientable spin manifold, it turns out that p_1 is always divisible by 48.

The canonical example of a non-spin 4-manifold is complex projective space $\mathbf{CP}^2$. It's not possible to consistently patch spinor fields over this space, and so the question of counting solutions to the Dirac equation is irrelevant. It turns out that $p_1(\mathbf{CP}^2) = 3$. Moreover, one can show that for any orientable 4-manifold, p_1 is always divisible by 3.

For manifolds with $\dim M = 8$, we need the result

$$p_2 = \frac{1}{128\pi^4}\int_M \left((\operatorname{tr}\mathcal{R}^2)^2 - 2\operatorname{tr}\mathcal{R}^4\right)$$

There are generalisations to higher dimensional manifolds.

3.3.3 Adding a Gauge Field

There is an interesting generalisation of the action (3.28) that retains $N = 1$ supersymmetry. This is

$$L = \int dt\ \frac{1}{2}\, g_{ij}(x)\, \dot{x}^i \dot{x}^j + \frac{i}{2} g_{ij} \psi^i \nabla_t \psi^j + i\eta_\alpha^\dagger D_t \eta^\alpha + \frac{1}{2} (F_{ij})^\alpha{}_\beta \eta_\alpha^\dagger \eta^\beta \psi^i \psi^j \tag{3.35}$$

with $i, j = 1, \ldots, n$ as before and $\alpha, \beta = 1, \ldots, r$. Here the ψ^i are Majorana fermions, with covariant derivative given by the same expression (3.29) that we had before. Meanwhile, the η_α are complex fermions with covariant derivative

$$D_t \eta^\alpha = \frac{d\eta^\alpha}{dt} + (A_i)^\alpha_\beta \frac{dx^i}{dt} \eta^\beta$$

We should think of A_i as a $U(r)$ gauge connection over the manifold M. Just as the metric g_{ij} is something fixed, so too is this gauge field. In more mathematical language, we should think of A as a connection of a vector bundle E. In the action (3.35) F is the associated field strength

$$(F_{ij})^\alpha{}_\beta = \partial_i (A_j)^\alpha{}_\beta - \partial_j (A_i)^\alpha{}_\beta + [A_i, A_j]^\alpha{}_\beta$$

Note that the four-fermion term involves F, which is the curvature of A. This entirely analogous to the situation in $N = 2$ supersymmetric quantum mechanics where the four-fermion term involves the Riemann tensor, which is the curvature of the spin connection.

At first glance, it is surprising that the action (3.35) admits supersymmetry. Until now, all our examples of supersymmetry involve a matching between bosonic and fermionic degrees of freedom. But here we have introduced additional fermionic degrees of freedom η^α without the corresponding bosonic partners. Nonetheless, you can check that the action is invariant under the following $N = 1$ supersymmetry,

$$\begin{aligned} \delta x^i &= \epsilon \psi^i \\ \delta \psi^i &= -\epsilon \dot{x}^i \\ \delta \eta^\alpha &= -\epsilon \psi^i (A_i)^\alpha{}_\beta \psi^\beta \end{aligned}$$

Together with the conjugate expression $\delta \eta_\beta^\dagger = \epsilon \eta_\alpha^\dagger (A_i)^\alpha{}_\beta \psi^i$. It turns out that adding extra fermions, without bosonic counterparts, is an option only for supersymmetric theories in $d = 0 + 1$ and $d = 1 + 1$ dimensions.

How should we think about the theory (3.35). Quantising the real fermions ψ^i gives us a spinor over M as before. Quantising the complex fermions η^α gives us forms, but they're not forms over the manifold M since they don't carry the $i = 1, \ldots, n$ index. Instead, they are forms over the gauge bundle E. In more mundane language, this simply means that the states transform in different representations of the $U(r)$ gauge connection. Ignoring the ψ^i fermions for now, we start with the state $|0\rangle$ such that

$$\eta^\alpha |0\rangle = 0$$

This is a singlet (i.e. neutral) under the $U(r)$ gauge bundle. Next, we have $\eta^{\dagger\,\alpha}|0\rangle$ which sit in the fundamental representation of $U(r)$, and the $\eta^{\dagger\,\alpha}\eta^{\dagger\,\beta}|0\rangle$ which sits in the anti-fundamental representation, and so on. Furthermore, there is a $U(1)$ symmetry that rotates $\eta^\alpha \to e^{i\theta}\eta^\alpha$ and this ensures that energy eigenstates have fixed number of $\eta^\dagger$ excitations and so sit in a fixed representation of $U(r)$.

The upshot is that the Hilbert space consists of a collection of spinors over M, each transforming in the p^{th} antisymmetric representation of $U(r)$. The $E = 0$ ground states obey $Q|\chi\rangle = 0$ which, translated into the geometric language, becomes the Dirac equation $\not{D}\chi = 0$ with

$$D_i = \partial_i + \frac{1}{2}(\omega_i)^{bc}S_{bc} + A_i$$

with A_i in the appropriate representation. Once again, this can be put in block off-diagonal form (3.31) and we can compute the index $\mathcal{I}(\mathcal{D})$ which, as in (3.32), coincides with the Witten index. Restricting to the fundamental representation, a similar calculation to the one above now yields the Atiyah-Singer index theorem

$$\mathcal{I}(\mathcal{D}) = \int_M \hat{A}(M) \wedge \mathrm{ch}(F) \tag{3.36}$$

where F it the *Chern character*,

$$\mathrm{ch}(F) = \mathrm{Tr}\, e^{F/2\pi} = r + c_1(F) + c_2(F) + \ldots$$

The individual Chern classes are topological invariants of the gauge field when integrated over the manifold M. The first two are

$$c_1 = \frac{1}{2\pi}\mathrm{Tr}\, F \quad \text{and} \quad c_2 = \frac{1}{8\pi^2}\mathrm{Tr}\, F \wedge F - \mathrm{Tr}\, F \wedge \mathrm{Tr}\, F$$

There are close connections here to the physics of solitons. In particular, the integral of the first Chern number counts the number of vortices in a gauge theory. The integral of the second Chern number counts the number of Yang-Mills instantons. Both of these have an interesting zero mode structure, even in flat spacetime, as follows from the Atiyah-Singer index theorem (3.36).

3.4 What Comes Next?

We have barely scratched the surface of supersymmetric theories and their connection to mathematics. In this final section we give a brief sketch of the next steps. Roughly speaking, there are two directions that are particularly rich: increase the number of supersymmetries, and increase the spacetime dimension of the quantum theory.

First the supersymmetries. The stories we told above revolved around $N = 2$ supersymmetry (for Morse theory) and $N = 1$ supersymmetry (for the index theorem). We briefly met theories with $N = 4$ supersymmetry in Section 1.4.2 where we saw that they naturally come with complex fields and a holomorphic superpotential. This suggests that a sigma model with $N = 4$ supersymmetry should have a target space M that is, in some sense, a complex manifold.

This is indeed what happens. Sigma models with $N = 4$ supersymmetry have target spaces that are *Kähler*. These target spaces necessarily have even dimension, with coordinates that can be paired together into complex numbers consistently over the entire manifold. This structure is best seen through the introduction of superfields, where the kinetic terms in the Lagrangian are written directly in terms of a function known as the *Kähler potential* $K(\phi, \bar{\phi})$ that is related to the metric by

$$g_{i\bar{j}} = 2\,\frac{\partial^2 K}{\partial \phi^i \bar{\phi}^{\bar{j}}} \tag{3.37}$$

Superfields and the Kähler potential are both described in the accompanying lectures on Supersymmetric Field Theory.

More supersymmetry brings yet more structure to the geometry, at least to a point. There are theories with $N = 8$ supersymmetry whose target spaces are *hyperKähler*. Such manifolds have a dimension that is a multiple of four, with three inequivalent ways of pairing coordinates together into complex numbers. It's a little like having a quaternionic structure on the manifold (although beware that there is a different kind of object in mathematics known as a "quaternionic manifold"). However, the interesting things don't keep happening forever and by the time we get to $N = 16$ supersymmetry the restriction becomes too strong and the target space is obliged to be flat and largely boring.

The full riches of supersymmetry really come when we consider theories in higher dimensions, meaning quantum field theories rather than quantum mechanics. While there are interesting stories for field theories in any spacetime dimension $d \leq 6$ (and, if you include gravitational theories, for $d \leq 11$) there is, as I now explain, a reason why QFTs in $d = 1 + 1$ dimensions are special.

Consider a sigma model in d spacetime dimensions. We introduce coordinates x^μ, with $\mu = 0, 1, \ldots, d-1$ for spacetime. The action is again based on some target space M. This means that the fields $\phi^i(x)$, with $i = 1, \ldots, n$ should be thought of as coordinates on M. The sigma model takes the form

$$S = \int d^dx \; g_{ij}(\phi)\partial_\mu\phi^i\partial^\mu\phi^j + \text{fermions} \tag{3.38}$$

Note that there are two conceptually different spaces in this action. The spacetime of the quantum field theory has dimension d, while the target space M has dimension $\dim M = n$.

With no potential $V(\phi)$ to preference one point on the manifold M from another, this theory has many classical ground states: each point on M or, equivalently, each constant value of ϕ^i should be viewed as a different classical ground state of the system.

But what happens in the quantum theory? We've already seen that for $d = 1$, which is just quantum mechanics, the ground state wavefunction spreads over M. This is important as it means that the ground state knows something about the entire manifold M and may therefore encode some information about its topology. We've seen examples of this throughout these lectures.

What happens in higher dimensions with $d > 1$? It turns out that $d = 2$ is just like quantum mechanics: the ground state wavefunction spreads over the whole manifold M. Meanwhile, at least in this respect, quantum theories in dimensions $d \geq 3$ behave like the classical theory: each point on M defines a different ground state.

I won't prove this statement in these lectures. It sometimes goes by the name of the (Coleman)-Mermin-Wagner theorem and is closely related to the concept of a "lower critical dimension" in Statistical Field Theory. At heart, it boils down to a property of the Poisson equation $\nabla^2\phi = \delta(x)$ in d Euclidean dimensions. At long distances, the solution grows in $d = 1$ and $d = 2$. (It is $\phi \sim x$ in $d = 1$ and $\phi \sim \log x$ in $d = 2$.) Conversely, the solution decays at long distance as $\phi \sim 1/x^{d-2}$ in $d \geq 3$. Physically, this translates into the fact that wavefunctions spread over M for sigma models with $d = 1$ and $d = 2$, while the ground state remains localised at a point in M for $d \geq 3$.

This means that if we want to find some interesting physics that captures topological properties of M, we should first look at $d = 1$ and $d = 2$. We've now spent over 100 pages studying supersymmetric quantum mechanics. So the next step is to look at supersymmetric sigma models of the form (3.38) in $d = 1 + 1$.

There is a rather special feature of quantum field theories in $d \geq 2$ dimensions that distinguishes them from quantum mechanics in $d = 1$: renormalisation. This means that the coupling constants that characterise a theory are not, in fact, constant. Instead they change with scale. Sigma models in $d = 1 + 1$ dimensions are no exception. Here, the couplings of the theory are encoded in the metric $g_{ij}(\phi)$. Under renormalisation, this metric changes and depends on the scale μ at which you look. The manner in which the metric changes is governed by the beautifully geometric beta function equation, known as *Ricci flow*

$$\mu \frac{\partial g_{ij}}{\partial \mu} = R_{ij} \tag{3.39}$$

This equation also plays an important role in String Theory, where it is ultimately responsible for the emergence of the Einstein equations of general relativity.

Taking into account the renormalisation (3.39), there are three types of behaviour that can occur depending on the type of target space M. Those spaces with positive Ricci curvature, $R > 0$, will shrink under RG flow.. In this case, the theory becomes increasingly strongly coupled in the infra-red and the impact on the physics is rather dramatic with the seemingly massless scalars ϕ^i developing a quantum-generated mass. Examples include $M = \mathbf{S}^n$ and $M = \mathbf{CP}^n$. You can read more about this in the lectures on Statistical Field Theory and the lectures on Gauge Theory.

Target spaces M with negative Ricci curvature, $R < 0$, will typically expand under RG flow. In this case they become more and more weakly coupled as the flow to they infra-red. Hyperbolic spaces provide a simple example.

The sweet spot are those target spaces M for which the metric is Ricci flat, with $R_{ij} = 0$. In fact, rather wonderfully, you don't have to start in the UV with a manifold M with a Ricci flat metric. If the manifold admits a Ricci flat metric, then the quantum theory will typically find it through the RG flow. The long-wavelength physics of such a theory is then governed by an interacting conformal field theory. Or, if we're dealing with a supersymmetric sigma model, an interacting supersymmetric conformal field theory or SCFT for short.

At this point, there is again an lovely intersection with results from mathematics. If we have $N = 4$ supersymmetry so that the target space is Kähler, then there is famous class of compact manifolds M that admit a Ricci flat metric known as *Calabi-Yau manifolds*. (For what it's worth, they are defined by having vanishing first Chern class.) Our discussion above means that for each Calabi-Yau manifold M there is an associated SCFT.

That, it turns out, is interesting. While the quantum mechanical sigma models described earlier in these lecture notes capture well-known stories of geometry, the $d = 1 + 1$ sigma models give a new lens through which to look at the manifolds M. And this lens gives us new information about the manifolds that goes beyond what mathematicians originally knew (although they very quickly caught up!) It is appropriate to refer to this view of the target space M, as seen by a SCFT, as "quantum geometry".

There are many novelties that come from associating a SCFT to a Calabi-Yau manifold M. But one stands out. There is not a unique association between manifolds M and SCFTs. Instead, pairs of manifolds M and N turn out to give rise to the *same* SCFT. These two manifolds M and N are topologically distinct and, naively, one wouldn't have thought that they have anything to do with each other. But, perhaps surprisingly, "quantum geometry" turns out to be more myopic than classical geometry and the SCFT approach cannot distinguish objects which appear obviously different to a classical geometer. At first glance, this myopia might appear to be a weakness, but closer examination shows that it is very much a strength. The myopia only arises because there are deep connections between the two manifolds M and N, with the geometric information of one encoded in a hidden and subtle form in the other. In technical language, the complex structure of one manifold is mapped to the symplectic structure of the other. This mapping is known as *mirror symmetry*. It is, sadly, a topic for another course.