12

Instantons, Supersymmetry and Morse Theory

As a final application of instanton methods, we will present an exposition of instantons in supersymmetric theories. The great simplification that occurs because of supersymmetry is the exact pairing of fermionic states with bosonic states, which makes the calculation of the fluctuation determinant very simple. The fermionic determinant exactly cancels the bosonic one.

Morse theory and the Morse inequalities concern the critical points of a function defined on a compact, Riemannian manifold, and the global topological aspects of the manifold. It was the genius of Witten [125, 123, 124] to point out that there is a deep connection between Morse theory and supersymmetric quantum mechanics defined on a manifold. This is what we hope to recount in this chapter.

We will require some familiarity with certain concepts from differential geometry which we will review here, but the reader should refer to more detailed texts [42, 60, 26] for a more complete picture.

12.1 A Little Differential Geometry

12.1.1 Riemannian Manifolds

We consider a compact, n-dimensional Riemannian manifold. A manifold is a point set with a topology (the definition of the open sets in $\mathcal{M}$) that is locally homeomorphic to $\mathbf{R}^n$. This means that each point in the manifold is contained in an open subset U_i of the manifold which can be mapped to $\mathbf{R}^n$ by a homeomorphism $\phi_i \ni \phi_i(U_i) \subseteq \mathbf{R}^n$. Homeomorphism means the mapping takes open sets in U_i to open sets in $\mathbf{R}^n$. The set of the U_is cover the manifold, *i.e.* $\cup_i U_i = \mathcal{M}$. Any such set of U_is is called an atlas and each individual U_i provides a coordinate chart. If two different coordinate charts U_i and U_j have a non-empty intersection, $U_i \cap U_j \neq \varnothing$, then the function $\phi_i \circ \phi_j^{-1}$ which maps points

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

in $\phi_j(U_i \cap U_j) \to \phi_i(U_i \cap U_j)$, *i.e.* defines a function from $\mathbf{R}^n \to \mathbf{R}^n$ must be k times differentiable. This defines a $\mathbb{C}^k$ manifold. We will always simply take $\mathbb{C}^\infty$ manifolds. The (local) coordinates of each point in a given U_i are just the coordinates of the point to which it is mapped in $\mathbf{R}^n$.

12.1.2 The Tangent Space, Cotangent Space and Tensors

The manifold has a tangent space at each point P, $T_P(\mathcal{M})$, which is defined as the space of linear mappings of real-valued functions defined on the manifold to the real numbers which satisfy the Liebniz rule, $\vec{v}(fg) = (\vec{v}f)g + f(\vec{v}g)$. The dimension of the tangent space is also n. The elements of the tangent space are called vectors. A basis of the tangent space can be trivially given in terms of a system of local coordinates. If x^i are a set of coordinates at a point P of the manifold, then any linear mapping that satisfies the Liebniz rule, on the space of functions defined on the manifold at the point P can be defined by

$$\vec{v} : f(x) \to \mathbf{R} \ni \vec{v}(f) = v^i \partial_i f(x)|_P. \tag{12.1}$$

Thus a vector is equivalent to a set of n components $\vec{v} \equiv (v^1, v^2, \cdots, v^n)$. If the components of the vector are smoothly varying functions of the coordinates $v^i(x)$, then we define a vector field. The cotangent space $T_P^*(\mathcal{M})$ at the point P is simply defined as the dual vector space of the tangent space at the point P. The dual vector space of a given vector space is simply the space of linear mappings of the vector space to the real numbers, thus $T_P^*(\mathcal{M}) : T_P(\mathcal{M}) \to \mathbf{R}$. The dimensionality of $T_P^*(\mathcal{M})$ is also n. If we have an arbitrary basis E_i of $T_p(\mathcal{M})$, then the dual basis of $T_P^*(\mathcal{M})$ is defined by the condition

$$\langle E_i, e^j \rangle = \delta_i^j. \tag{12.2}$$

We name the dual basis to the coordinate basis ∂_i using the notation dx^j so that

$$\langle \partial_i, dx^j \rangle = \delta_i^j. \tag{12.3}$$

A general dual vector or "co-vector" can be written as $\vec{u}^* = u_j dx^j$ and then for a general vector $\vec{v} = v^i \partial_i$ we have

$$\langle \vec{v}, \vec{u}^* \rangle = v^i u_j \langle \partial_i, dx^j \rangle = v^i u_j \delta_i^j = v^i u_i. \tag{12.4}$$

If we change our system of coordinates of the coordinate chart at the point p, $x^i \to x'^j$, then the coordinate basis vectors of the tangent space transform simply as $\partial_i = \frac{\partial x'^j}{\partial x^i} \partial'_j$ or equivalently $\partial'_j = \frac{\partial x^i}{\partial x'^j} \partial_i$. But then the new dual basis vectors must be given by $dx'^j = \frac{\partial x'^j}{\partial x^i} dx^i$ or equivalently $dx^i = \frac{\partial x^i}{\partial x'^j} dx'^j$ so that the inner product between ∂_i and dx^j is preserved, *i.e.*

$$\begin{aligned} \langle \partial_i, dx^j \rangle &= \left\langle \frac{\partial x'^k}{\partial x^i} \partial'_k, \frac{\partial x^j}{\partial x'^l} dx'^l \right\rangle = \frac{\partial x'^k}{\partial x^i} \frac{\partial x^j}{\partial x'^l} \langle \partial'_k, dx'^l \rangle \\ &= \frac{\partial x'^k}{\partial x^i} \frac{\partial x^j}{\partial x'^l} \delta_k^l = \frac{\partial x'^k}{\partial x^i} \frac{\partial x^j}{\partial x'^k} = \delta_i^j. \end{aligned} \tag{12.5}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

This then gives the transformation properties of the covariant and contravariant components of of vectors and co-vectors, indeed, $\vec{v} = v'^j \partial'_j = v'^j \frac{\partial x^i}{\partial x'^j} \partial_i = v^i \partial_i$ and $\vec{u}* = u'_j dx'^j = u'_j \frac{\partial x'^j}{\partial x^i} dx^i = u_i dx^i$. Hence $v^i = v'^j \frac{\partial x^i}{\partial x'^j}$ and $u_i = u'_j \frac{\partial x'^j}{\partial x^i}$ or equivalently $v'^i = v^j \frac{\partial x'^i}{\partial x^j}$ and $u'_i = u_j \frac{\partial x^j}{\partial x'^i}$. Then the inner product between arbitrary vectors and co-vectors is invariant

$$\langle \vec{v}, \vec{u}^* \rangle = v^i u_i = v'^i u'_i. \tag{12.6}$$

We note the possibly confusing nomenclature: the components of vectors are said to transform contravariantly while the components of co-vectors are said to transform covariantly.

We can also take tensor products of the tangent space k times and the cotangent space l times,

$$\underbrace{T_P \otimes \cdots \otimes T_P}_{k} \otimes \underbrace{T_P^* \otimes \cdots \otimes T_P^*}_{l} \tag{12.7}$$

to define tensors (and tensor fields)

$$t = t^{i_1 \cdots i_k}_{j_1 \cdots j_l} \partial_{i_1} \otimes \cdots \otimes \partial_{i_k} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_l}. \tag{12.8}$$

We should stress that at this point there is no relationship between the tangent spaces, the cotangent spaces and their tensor products over distinct points. The construction is independently done over each point. To use a leading terminology, there is, at the moment, no connection between tangent spaces at neighbouring points. The ensemble of the tangent spaces over all the points in the manifold defines a larger manifold called the tangent bundle, a fibre bundle over the manifold $\mathcal{M}$. The base manifold is $\mathcal{M}$ and the fibre is T_P over the point P in $\mathcal{M}$. There is also the corresponding cotangent bundle constructed with the cotangent space. The complete spaces are fibre bundles, spaces that locally permit a decomposition into a Cartesian product of a patch of the base manifold $\mathcal{M}$ cross the fibre, which would be the tangent space in the case of the tangent bundle, etc.

12.2 The de Rham Cohomology

12.2.1 The Exterior Algebra

The de Rham cohomology concerns the ensemble of the set of spaces of the completely anti-symmetric tensor products of the dual tangent space. We start with the cotangent space, T_P^*. Any two basis elements dx and dy can form an anti-symmetric two-co-tensor defined as

$$dx \wedge dy = \frac{1}{2}(dx \otimes dy - dy \otimes dx). \tag{12.9}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

The product $\wedge$ is called the Cartan wedge product or the exterior product. Then an arbitrary anti-symmetric two-co-tensor is given by

$$t = t_{ij} dx^i \wedge dx^j. \tag{12.10}$$

This construction obviously generalizes to the notion of anti-symmetric p co-tensors constructed over each point x of the manifold. The set of anti-symmetric p co-tensors forms a sub-space of the p-fold tensor product of the co-tangent space which we will call $\Lambda^p(x)$. The set of $\Lambda^p(x)$s for all the points of the manifold forms a fibre bundle over $\mathcal{M}$. The elements of $\Lambda^p(x)$ are called differential forms, or more precisely p-forms. The dimensionality of $\Lambda^p(x)$ is obviously $\binom{n}{p}$ the number of ways of choosing p basis vectors from the total set of n basis vectors. We add in $\Lambda^0(x) = \mathbf{R}$, simply the real line, and then we have $n+1$ spaces of differential forms, $\Lambda^0(x)$ to $\Lambda^n(x)$, since for $\Lambda^{n+1}(x)$ or higher, it is no longer possible to anti-symmetrize $n+1$ or more co-vectors and these spaces are just empty. The space of smooth p-forms corresponds to the choice of the anti-symmetric tensor component fields $f_{i_1 \cdots i_p}(x)$, the corresponding tensor field being $f_{i_1 \cdots i_p}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_p}$, which we write as $C^\infty(\Lambda^p)$ which is a space of dimension $\binom{n}{p}$. It is obvious that $C^\infty(\Lambda^p)$ and $C^\infty(\Lambda^{n-p})$ have the same dimensionality. The wedge product serves as a product on the full space of the direct sum of all possible anti-symmetric tensor fields

$$\Lambda^* = \Lambda^0 \oplus \Lambda^1 \oplus \cdots \oplus \Lambda^n, \tag{12.11}$$

which then defines an algebra called Cartan's exterior algebra.

12.2.2 Exterior Derivative

We can define the exterior derivative of a p-form, an operation d, which takes p forms to $p+1$ forms

$$d : C^\infty(\Lambda^p) \to C^\infty(\Lambda^{p+1}) \quad \ni$$
$$d(f_{i_1 \cdots i_p}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_p}) = \left(\frac{\partial}{\partial x^j} f_{i_1 \cdots i_p}(x) \right) dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}. \tag{12.12}$$

Note the placement of the additional dx^j by convention to the left of all the other differential forms. Obviously

$$dd\omega_p = 0 \tag{12.13}$$

for any p-form ω_p. The chain rule also simply follows, for ω_p a p-form and χ_q a q-form

$$d(\omega_p \wedge \chi_q) = (d\omega_p) \wedge \chi_q + (-1)^p \omega_p \wedge (d\chi_q). \tag{12.14}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

As $C^\infty(\Lambda^p)$ and $C^\infty(\Lambda^{n-p})$ have the same dimensionality, we can define a duality mapping between these spaces, called the Hodge $*$ duality transformation. We define

$$\begin{aligned} * &: C^\infty(\Lambda^p) \to C^\infty(\Lambda^{n-p}) \quad \ni \\ *(dx^{i_1} \wedge \cdots \wedge dx^{i_p}) &= \frac{1}{(n-p)!} \epsilon^{i_1 \cdots i_p}_{i_{p+1} \cdots i_n} dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n} \end{aligned} \tag{12.15}$$

where $\epsilon^{i_1 \cdots i_p}_{i_{p+1} \cdots i_n}$ is the completely anti-symmetric tensor in n dimensions. We have been careful about keeping indices up or down; however, it is important to point out that nothing we are doing requires the definition of a metric on the manifold. The exterior algebra and exterior differentiation does not depend on a metric. We note that

$$** \omega_p = (-1)^{p(n-p)} \omega_p. \tag{12.16}$$

12.2.3 Integration

The space $C^\infty(\Lambda^n)$ is one-dimensional, there is only one n-form, $dx^1 \wedge \cdots \wedge dx^n$, thus it is easy to see that $dx^{i_1} \wedge \cdots \wedge dx^{i_n} = \epsilon^{i_1 \cdots i_n} dx^1 \wedge \cdots \wedge dx^n$. This form can be identified with the volume form on the manifold and we can define the integration over the manifold with this volume form; one simply integrates in $\mathbf{R}^n$ in the charts of any given atlas, making sure not to double count the contributions from regions where the charts intersect. The integration is independent of the coordinate system, since the volume form transforms exactly by the Jacobian of the coordinate transformation, $dx^1 \wedge \cdots \wedge dx^n = det\left(\frac{\partial x^i}{\partial x'^j}\right) dx'^1 \wedge \cdots \wedge dx'^n$. This integration generalizes trivially to integration over sub-manifolds of $\mathcal{M}$ of a given dimensionality p of a p-form defined over the sub-manifold. With the notion of integration, we can define an inner product on the space of p-forms

$$(\omega_p, \chi_p) = \int_{\mathcal{M}} \omega_p \wedge * \chi_p. \tag{12.17}$$

In terms of the coefficients, $\omega_p = \omega_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$ and $\chi_p = \chi_{j_1 \cdots j_p} dx^{j_1} \wedge \cdots \wedge dx^{j_p}$ then

$$\begin{aligned} (\omega_p, \chi_p) &= \int_{\mathcal{M}} \omega_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge *(\chi_{j_1 \cdots j_p} dx^{j_1} \wedge \cdots \wedge dx^{j_p}) \\ &= \int_{\mathcal{M}} \omega_{i_1 \cdots i_p} \chi_{j_1 \cdots j_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge \frac{1}{(n-p)!} \epsilon^{j_1 \cdots j_p}_{j_{p+1} \cdots j_n} dx^{j_{p+1}} \wedge \cdots \wedge dx^{j_n} \\ &= \int_{\mathcal{M}} \omega_{i_1 \cdots i_p} \chi_{j_1 \cdots j_p} \frac{1}{(n-p)!} \epsilon^{j_1 \cdots j_p}_{j_{p+1} \cdots j_n} \epsilon^{i_1 \cdots i_p j_{p+1} \cdots j_n} dx^1 \wedge \cdots \wedge dx^n \\ &= p! \int_{\mathcal{M}} \omega_{i_1 \cdots i_p} \chi_{j_1 \cdots j_p} \delta^{i_1 j_1} \cdots \delta^{i_p j_p} dx^1 \wedge \cdots \wedge dx^n, \end{aligned} \tag{12.18}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

The inner product is symmetric, $(\omega_p, \chi_p) = (\chi_p, \omega_p)$.

The next structure we will define is the adjoint of the exterior derivative, which we call δ. The inner product

$$\begin{aligned}(\omega_p, d\chi_{p-1}) &= \int_{\mathcal{M}} \omega_p \wedge *d\chi_{p-1} = \int_{\mathcal{M}} (d\chi_{p-1}) \wedge *\omega_p \\ &= \int_{\mathcal{M}} d(\chi_{p-1} \wedge *\omega_p) - (-1)^{p-1} \chi_{p-1} \wedge d * \omega_p \\ &= \int_{\mathcal{M}} -(-1)^{p-1} \chi_{p-1} \wedge (-1)^{(n-p+1)(n-n+p-1)} * *d * \omega_p \\ &= \int_{\mathcal{M}} ((-1)^{np+n+1} * d * \omega_p) \wedge *\chi_{p-1} \\ &\equiv (\delta\omega_p, \chi_{p-1}),\end{aligned} \tag{12.19}$$

where we have used trivial identities such as $(-1)^{2n} = 1$. Thus $\delta = (-1)^{np+n+1} * d*$, and note for n even the sign is always -1 and $\delta = - * d*$, while for n odd we get $\delta = (-1)^p * d*$. It is also easy to see $\delta\delta\omega_p = 0$.

The exterior algebra naturally gives rise to a Stokes theorem for manifolds with boundaries. If $\partial\mathcal{M}$ is the boundary of a p-dimensional manifold $\mathcal{M}$ and ω_{p-1} is an arbitrary $(p-1)$-form, then Stokes theorem states

$$\int_{\mathcal{M}} d\omega_{p-1} = \int_{\partial\mathcal{M}} \omega_{p-1}. \tag{12.20}$$

This theorem contains and generalizes all three of the usual Green, Gauss and Stokes theorems that are taught in an elementary course on vector calculus.

12.2.4 The Laplacian and the Hodge Decomposition

The Laplacian is now defined as

$$\nabla^2 = (d+\delta)^2 = d\delta + \delta d. \tag{12.21}$$

The Laplacian does not change the degree of the form. The Laplacian is a positive operator as

$$(\omega_p, \nabla^2\omega_p) = (\omega_p, d\delta\omega_p + \delta d\omega_p) = (\delta\omega_p, \delta\omega_p) + (d\omega_p, d\omega_p) \geq 0, \tag{12.22}$$

assuming there are no boundaries. Therefore, $\nabla^2\omega_p = 0$, and then ω is called a harmonic p-form, if and only if both $d\omega_p = 0$ (we say ω_p is closed) and $\delta\omega_p = 0$ (we say ω_p is co-closed).

A p-form that can be globally written as the exterior derivative of a $p-1$-form, *i.e.*

$$\omega_p = d\zeta_{p-1} \tag{12.23}$$

is called an exact p-form while if ω can be globally written as

$$\omega_p = \delta\xi_{p+1} \tag{12.24}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

then it is called a co-exact p-form. The Hodge theorem states that on a compact manifold without boundary any p-form ω_p can be uniquely decomposed as the sum of an exact form, a co-exact form and a harmonic form

$$\omega_p = d\zeta_{p-1} + \delta\xi_{p+1} + \rho_p, \tag{12.25}$$

where ρ_p is a harmonic form, meaning that $\nabla^2\rho_p = 0$.

12.2.5 Homology

The homology of a manifold is the set of equivalence classes of sub-manifolds called cycles, boundaryless collections of sub-manifolds of dimension p, which differ only by boundaries. We start with our initial n-dimensional manifold $\mathcal{M}$. Then we define a p-chain as a formal sum of p-dimensional, smooth, oriented, sub-manifolds, $\mathcal{N}_i^p$, the formal, finite sum being written as $\alpha_p = \sum_i c_i \mathcal{N}_i^p$, where c_i are real, complex or integer, or even in the group $\mathbb{Z}_2$, giving rise to the corresponding p-chain. We continue to use the symbol ∂ as the operator that corresponds to taking the boundary, $\partial\alpha_p = \sum_i c_i \partial\mathcal{N}_i^p$, which is evidently a $(p-1)$-chain. Let Z_p be the set of boundaryless p-chains, which are called p-cycles. This means $\alpha_p \in Z_p \Rightarrow \partial\alpha_p = \varnothing$. Let X_p be the set of p-chains that are boundaries, *i.e.* $\alpha_p \in X_p \Rightarrow \alpha_p = \partial\alpha_{p+1}$. Since the boundary of a boundary is always empty, $X_p \subseteq Z_p$. Then the simplicial homology of $\mathcal{M}$ is defined as the set of equivalence classes H_p

$$H_p = Z_p / X_p, \tag{12.26}$$

i.e. the set of p-cycles that only differ from each other by boundaries are considered equivalent, $\alpha_p \sim \alpha_p' \Rightarrow \alpha_p = \alpha_p' + \partial\alpha_{p+1}$. H_p is obviously a group under the formal addition. The formal sum of two p-cycles commutes with the process of making equivalence classes with respect to p-cycles which are boundaries. The integral homology groups are the most fundamental, we can get the real, complex or $\mathbb{Z}_2$ homologies from them. We will write the homology groups as $H_p(\mathcal{M}, G)$, where $G = \mathbb{C}, \mathbf{R}, \mathbb{Z}, \mathbb{Z}_2$. $H_p(\mathcal{M}, G) = \varnothing$ for $p > n$.

$H_0(\mathcal{M}, G) = G$ if $\mathcal{M}$ is connected, since 0-cycles are just collections of points, the boundary of a point is empty. We can reduce any finite collection of points with arbitrary coefficients to a 0-cycle consisting of single point, $P \in \mathcal{M}$. Any 0-cycle, $\alpha_0 = \sum_i c_i P_i$, can be reduced to single point P with a coefficient $\sum_i c_i \in G$, using

$$\alpha_0 = \sum_i c_i P_i = \sum_i (c_i P_i - c_i P + c_i P) \sim \left(\sum_i c_i\right) P \tag{12.27}$$

as every pair of points with alternating coefficient, as appears above $c_i P_i - c_i P$, is the boundary of a 1-cycle corresponding to any curve joining the two points. However, cP is not equivalent to $\tilde{c}P$ for $c \neq \tilde{c} \in G$, thus the elements of $H_0(\mathcal{M}, G)$ are in a one-to-one correspondence with G. Obviously $H_n(\mathcal{M}, G) = G$ also, since

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

there is only one sub-manifold of dimension n, $\mathcal{M}$ itself, in $\mathcal{M}$, and we see that $H_0(\mathcal{M},G) = H_n(\mathcal{M},G)$. This generalizes to what is called Poincaré duality, when $G = \mathbf{R}, \mathbb{C}$ or $\mathbb{Z}_2$ (all fields) we have $H_p(\mathcal{M},G) = H_{n-p}(\mathcal{M},G)$.

Finally, for $G = \mathbf{R}, \mathbb{C}$ or $\mathbb{Z}_2$ the homology group $H_p(\mathcal{M},G)$ is clearly a vector space over G. We define the cohomology group $H^p(\mathcal{M},G)$ simply as the dual vector space to $H_p(\mathcal{M},G)$.

12.2.6 De Rham Cohomology

We define the de Rham cohomology group with respect to differential forms for $G = \mathbf{R}, \mathbb{C}$. With the definitions Z^p as the set of closed p-forms and X^p as the set of exact p-forms, the de Rham cohomolgy group is defined as

$$H^p_{dR}(\mathcal{M},G) = Z^p/X^p, \tag{12.28}$$

i.e. the equivalence classes of closed modulo exact p-forms, $\omega_p \sim \omega'_p \Rightarrow \omega_p = \omega'_p + d\alpha_{p-1}$. For the special case of $H^0_{dR}(\mathcal{M},G)$ we define this as the space of constant functions, as their exterior derivative vanishes. A zero-form cannot be the exterior derivative of any "-1" form, as these do not exist. The spectacular conclusion of the de Rham theorem asserts that these cohomology groups are in fact identical to the simplicial cohomology groups and hence dual to the simplicial homology groups.

We define the inner product of a p-cycle $\alpha_p \in Z_p$ with a closed p-form $\omega_p \in Z^p$ through the integral

$$\pi(\alpha_p,\omega_p) = \int_{\alpha_p} \omega_p. \tag{12.29}$$

It is easy to see that this inner product only depends on the equivalence class of α_p and of ω_p. Indeed,

$$\int_{\alpha_p} (\omega_p + d\chi_{p-1}) = \int_{\alpha_p} \omega_p + \int_{\alpha_p} d\chi_{p-1} = \int_{\alpha_p} \omega_p + \int_{\partial\alpha_p} \chi_{p-1} = \int_{\alpha_p} \omega_p \tag{12.30}$$

and

$$\int_{\alpha_p+\partial\beta_{p+1}} \omega_p = \int_{\alpha_p} \omega_p + \int_{\partial\beta_{p+1}} \omega_p = \int_{\alpha_p} \omega_p + \int_{\beta_{p+1}} d\omega_p = \int_{\alpha_p} \omega_p \tag{12.31}$$

as $\partial\alpha_p = \varnothing$ and $d\omega_p = 0$. Thus π gives a mapping

$$\pi : H_p(\mathcal{M},G) \otimes H^p_{dR}(\mathcal{M},G) \to G. \tag{12.32}$$

De Rham proved the following theorems. Let $\{c_i\}, i = 1,\cdots,dim(H_p(\mathcal{M},R)$, be a set of independent p-cycles that form a basis of $H_p(\mathcal{M},R)$. Then

1. For any given set of periods $\nu_i, i = 1,\cdots,dim(H_p(\mathcal{M},R)$ there exists a closed p-form ω_p such that

$$\nu_i = \pi(c_i,\omega_p) = \int_{c_i} \omega_p. \tag{12.33}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

2. If all the periods vanish for a give p-form ω_p, then ω_p is exact, *i.e.* $\omega_p = d\chi_{p-1}$.

This means that if $\{\omega_j\}$ is a basis of p-forms of $H^p_{dR}(\mathcal{M}, \mathbf{R})$ then the period matrix $\pi_{ij} = \pi(c_i, \omega_j)$ is invertible. This is equivalent to saying that $H^p_{dR}(\mathcal{M}, \mathbf{R})$ is dual to $H_p(\mathcal{M}, \mathbf{R})$. Consequently, the de Rham cohomology and the simplicial cohomology are naturally isomorphic and can be identified.

The Hodge theorem asserts that for each de Rham cohomology class there is an essentially unique harmonic form that can be taken as the representative of the class. Indeed, we have from the Hodge decomposition

$$\omega_p = d\zeta_{p-1} + \delta\xi_{p+1} + \rho_p. \tag{12.34}$$

Then evidently, the exact form $d\zeta_{p-1}$ is irrelevant in determining the equivalence class. ω_p being closed and ρ_p being harmonic, thus $d\omega_p = d\rho_p = 0$ which implies that $d\delta\xi_{p+1} = 0$, but then $0 = (\xi_{p+1}, d\delta\xi_{p+1}) = (\delta\xi_{p+1}, \delta\xi_{p+1})$ requires $\delta\xi_{p+1} = 0$. Thus $\omega_p = d\zeta_{p-1} + \rho_p$ and the de Rham cohomology class of ω_p is determined by the unique harmonic form ρ_p in its Hodge decomposition. This fact will be very important in the supersymmetric quantum mechanics that we will analyse in the later sections.

We define the Betti numbers as the dimension of the homology groups and consequently also the cohomology groups

$$B_p = dim(H_p(\mathcal{M}, \mathbf{R})) = dim(H^p_{dR}(\mathcal{M}, \mathbf{R})) = dim(H^p(\mathcal{M}, \mathbf{R})), \tag{12.35}$$

where B_p is the pth Betti number. The alternating sum of the Betti number is the Euler characteristic

$$\chi(\mathcal{M}) = \sum_{p=0}^{n} (-1)^p B_p \tag{12.36}$$

and we will see it is a topological invariant of the manifold. Morse theory relates the critical points of functions defined on a manifold to its Betti numbers.

12.3 Supersymmetric Quantum Mechanics

After this brief, condensed exposition of manifolds, structures defined on them and of the de Rham cohomology we can now move on to show how supersymmetry and instantons can be used to prove the global topological results framed in the Morse inequalities [62].

12.3.1 The Supersymmetry Algebra

In any quantum theory we can separate the Hilbert space $\mathbb{H}$ into $\mathbb{H} = \mathbb{H}^+ \oplus \mathbb{H}^-$, where $\mathbb{H}^+$ and $\mathbb{H}^-$ are the subspaces of bosonic and fermionic states, respectively. A supersymmetry corresponds to a transformation generated by conserved hermitean operators $Q_i, i = 1, \cdots, N$ that maps $\mathbb{H}^+$ to $\mathbb{H}^-$ and vice versa. We also

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

define the operator $(-1)^F$, where F is the fermion number. Then $(-1)^F|\psi\rangle = |\psi\rangle$ for $|\psi\rangle \in \mathbb{H}^+$ while $(-1)^F|\psi\rangle = -|\psi\rangle$ for $|\psi\rangle \in \mathbb{H}^-$. The supersymmetry generators must anti-commute with $(-1)^F$, $\{Q_i, (-1)^F\} = 0$, which means that they are fermionic operators. On the other hand, they commute with the Hamiltonian $\mathcal{H}$, $[Q_i, \mathcal{H}] = 0$, which means that they are conserved. Finally, to define a supersymmetric theory we also impose $Q_i^2 = \mathcal{H}$ for any i and $\{Q_i.Q_j\} = 0$ for $i \neq j$, together giving

$$\{Q_i.Q_j\} = 2\delta_{ij}\mathcal{H}. \tag{12.37}$$

This definition of supersymmetry does not allow for Lorentz-invariant theories. This is because Lorentz transformations combine the Hamiltonian to the momentum generators. In $1+1$ dimensions we have only one momentum generator, P. The simplest algebra preserving Lorentz symmetry requires two supersymmetry operators, Q_1 and Q_2 and the algebra

$$Q_1^2 = \mathcal{H} + P, \quad Q_2^2 = \mathcal{H} - P, \quad \{Q_1, Q_2\} = 0. \tag{12.38}$$

This is compatible with the idea that $(\mathcal{H}, P)$ transform as a vector and (Q_1, Q_2) transform as a spinor. There is just one generator of Lorentz transformation M, taken hermitean, and

$$[M, \mathcal{H}] = iP, \quad [M, P] = i\mathcal{H}, \quad [M, Q_1] = i\frac{1}{2}Q_1, \quad [M, Q_2] = -i\frac{1}{2}Q_2. \tag{12.39}$$

Then, for example, $[M, \mathcal{H} + P] = i(\mathcal{H} + P)$, which is compatible with

$$[M, Q_1^2] = [M, Q_1]Q_1 + Q_1[M, Q_1] = i\frac{1}{2}Q_1Q_1 + Q_1 i\frac{1}{2}Q_1 = iQ_1^2 = i(\mathcal{H} + P). \tag{12.40}$$

From Equation (12.38) we easily find

$$\mathcal{H} = \frac{1}{2}(Q_1^2 + Q_2^2) \tag{12.41}$$

and therefore the Hamiltonian is positive semi-definite. Also, $[\mathcal{H}, (-1)^F] = [P, (-1)^F] = 0$ as they are quadratic in the supercharges, hence the Hamiltonian and the momentum generator are bosonic operators.

If there exists a single state $|0\rangle$ in the Hilbert space that is annihilated by the supercharges

$$Q_i|0\rangle = 0 \quad i = 1, 2, \tag{12.42}$$

then the supersymmetry is unbroken. Such a state obviously has zero energy and, since the Hamiltonian is positive semi-definite, $|0\rangle$ is the vaccum state. If there are many solutions to Equation (12.42), then the supersymmetry is also unbroken, but presumably the Hilbert space separates into superselection sectors of states constructed over each vacuum. If there are no states that satisfy Equation (12.42), then the supersymmetry is spontaneously broken. It is generally quite difficult to directly prove the existence of solutions to

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

Equation (12.42), or the lack thereof. However, the following indirect method sheds light on the question: one computes the index of one of the supersymmetry generators.

We are looking for states that are annihilated by both supersymmetry generators, $Q_i|0\rangle = 0$. Then with the algebra (12.38) it is easy to see that $P|0\rangle = 0$, thus we can restrict to the subspace $\mathbb{H}_{P=0}$, which is all states annihilated by P. This subspace also splits into a bosonic and a fermionic subspace, $\mathbb{H}_{P=0} = \mathbb{H}^+_{P=0} \oplus \mathbb{H}^-_{P=0}$. Within this subspace, $Q_1^2 = Q_2^2 = \mathcal{H}$, restricted to $\mathbb{H}_{P=0}$, we can look for states that are annihilated by one of the supercharges, call it $\tilde{Q}$, where $\tilde{Q}$ could be Q_1 or Q_2 or a linear combination of the two. $\tilde{Q}$ necessarily can only take a state in $\mathbb{H}^+_{P=0} \to \mathbb{H}^-_{P=0}$ and a state in $\mathbb{H}^-_{P=0} \to \mathbb{H}^+_{P=0}$. $\tilde{Q}$ has no other action. This fact then allows for the decomposition $\tilde{Q} = Q_+ + Q_-$, where Q_+ acts only on and maps $\mathbb{H}^+_{P=0} \to \mathbb{H}^-_{P=0}$ while Q_- acts only on and maps $\mathbb{H}^-_{P=0} \to \mathbb{H}^+_{P=0}$. Q_- is the adjoint of Q_+. The index of $\tilde{Q}$ restricted to $\mathbb{H}_{P=0}$ is then defined as

$$index(\tilde{Q}) = dim(Ker(Q_+)) - dim(Ker(Q_-)), \tag{12.43}$$

where $Ker(Q_\pm)$ is the subspace of $\mathbb{H}^\pm_{P=0}$ that is annihilated by $Q_\pm$. If the index is non-zero then we know for sure that there are states that are annihilated by $\tilde{Q}$ and hence supersymmetry is unbroken. The $index(\tilde{Q})$ can be written as

$$index(\tilde{Q}) = Tr(-1)^F = n_B(E=0) - n_F(E=0) \tag{12.44}$$

as the bosonic zero modes in $\mathbb{H}^+_{P=0}$ count as $+1$ for each mode and the fermionic zero modes in $\mathbb{H}^-_{P=0}$ count as -1 for each mode. The non-zero energy modes are necessarily paired because of the supersymmetry, and hence cancel pairwise in their contribution to the trace. The index being non-zero requires necessarily that there exists at least one zero energy state and hence we can conclude that in this case the supersymmetry is unbroken. In the sequel we will drop the subscript $P = 0$ and take as given that we are working in the subspace with $P = 0$.

12.3.2 Supersymmetric Cohomology

The Hamiltonian is given by

$$\mathcal{H} = QQ^\dagger + Q^\dagger Q \tag{12.45}$$

with the superalgebra

$$Q^2 = Q^{\dagger 2} = 0, \tag{12.46}$$

with consequently

$$[\mathcal{H}, Q] = [\mathcal{H}, Q^\dagger] = 0. \tag{12.47}$$

There also exists the operator $(-1)^F$ and usually the fermion number operator F, which both commute with the Hamiltonian. The states in the Hilbert spaces are graded by the eigenvalue of $(-1)^F$. Bosonic states, a subspace denoted by

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

$\mathbb{H}^+$, take eigenvalue $+1$ while fermionic states, a subspace denoted by $\mathbb{H}^-$, take the eigenvalue -1. The fermion number operator is integer-valued, with bosonic states having an even number of fermions and fermionic states having an odd number. The Hamiltonian maps bosonic states to bosonic states and fermionic states to fermionic states, while the supercharges switch the two, mapping bosonic states to fermionic states and fermionic states to bosonic states.

$$\begin{aligned} \mathcal{H} &: \mathbb{H}^+ \to \mathbb{H}^+, \quad \mathbb{H}^- \to \mathbb{H}^- \\ Q, Q^\dagger &: \mathbb{H}^+ \to \mathbb{H}^-, \quad \mathbb{H}^- \to \mathbb{H}^+. \end{aligned} \tag{12.48}$$

If we write the energy levels in an ordered list $E_0 < E_1 < \cdots$ then the Hamiltonian preserves the energy-level subspace and the Hilbert space can be decomposed in terms of subspaces $\mathbb{H}_m$ of fixed energy levels

$$\mathbb{H} = \bigoplus_m \mathbb{H}_m \tag{12.49}$$

with the action of the Hamiltonian, the supercharges and $(-1)^F$ satisfying

$$\mathcal{H}|_{\mathbb{H}_m} = E_m, \quad Q, Q^\dagger, (-1)^F : \mathbb{H}_m \to \mathbb{H}_m. \tag{12.50}$$

The energy-level subspace further decomposes into bosonic and fermionic subspaces $\mathbb{H}_m = \mathbb{H}_m^+ \oplus \mathbb{H}_m^-$ and while the Hamiltonian preserves the bosonic and fermionic subspaces (they are indeed eigensubspaces of the Hamiltonian) the supercharges exchange the two

$$Q, Q^\dagger : \mathbb{H}_m^+ \to \mathbb{H}_m^-, \quad \mathbb{H}_m^- \to \mathbb{H}_m^+. \tag{12.51}$$

The action of the operator Q twice, vanishes, $Q^2 = 0$. Thus we have the exact sequence:

$$\mathbb{H}^- \xrightarrow{Q} \mathbb{H}^+ \xrightarrow{Q} \mathbb{H}^- \xrightarrow{Q} \mathbb{H}^+ \tag{12.52}$$

An exact sequence means that the image of a given map in the sequence is the kernel of the subsequent map. This is called a $\mathbb{Z}_2$-graded complex of vector spaces as the fermionic and bosonic Hilbert spaces are graded with the $\mathbb{Z}_2$ charge with respect to the operator $(-1)^F$. This gives rise to the cohomology groups:

$$\begin{aligned} H^+(Q) &= Kernel\left\{Q : \mathbb{H}^+ \to \mathbb{H}^-\right\} / Image\left\{Q : \mathbb{H}^- \to \mathbb{H}^+\right\} \\ H^-(Q) &= Kernel\left\{Q : \mathbb{H}^- \to \mathbb{H}^+\right\} / Image\left\{Q : \mathbb{H}^+ \to \mathbb{H}^-\right\}. \end{aligned} \tag{12.53}$$

We can further refine this complex by noting that at energy level $E_m \neq 0$, the action of Q does not take you out of the energy sector, since Q commutes with the Hamiltonian. If a vector $|E_m\rangle$ is Q closed, $Q|E_m\rangle = 0$, *i.e.* in the kernel of Q, then it is necessarily exact, *i.e.* in the image of the previous map, since

$$|E_m\rangle = \mathcal{H}|E_m\rangle/E_m = \left(QQ^\dagger + Q^\dagger Q\right)|E_m\rangle/E_m = Q\left(Q^\dagger|E_m\rangle/E_m\right). \tag{12.54}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

Hence all states that are closed are also exact for all the non-zero energy levels, and thus cohomology groups are just determined by the states in the zero energy sector. For a state of zero energy $|E_0\rangle$ we have

$$0 = \langle E_0|\mathcal{H}|E_0\rangle = \langle E_0|\left(QQ^\dagger + Q^\dagger Q\right)|E_0\rangle = \left|Q^\dagger|E_0\rangle\right|^2 + |Q|E_0\rangle|^2, \tag{12.55}$$

which is only possible if both $Q|E_0\rangle = 0$ and $Q^\dagger|E_0\rangle = 0$. Thus the zero energy states are annihilated by Q and hence closed. But none of them are exact, $|E_0\rangle \neq Q|\alpha\rangle$, since, if this were true, $Q^\dagger|E_0\rangle = 0 = Q^\dagger Q|\alpha\rangle$, which implies $\langle\alpha|Q^\dagger Q|\alpha\rangle = |Q|\alpha\rangle|^2 = 0$, which is only possible if $Q|\alpha\rangle = 0$. Thus the cohomology groups can be identified with the set of zero energy states:

$$\begin{aligned} H^+(Q) &= H_0^+ \\ H^-(Q) &= H_0^- \end{aligned} \tag{12.56}$$

where $H_0^\pm$ are the states of zero energy.

Q takes states of p fermions to states of $p+1$ fermions. It is reasonable to assign vanishing fermion number to states without fermions, and the action of Q an even number of times always gives back a bosonic subspace, while an odd number of times give us a fermionic subspace, hence with the notation that $\mathbb{H}^p$ is the subspace of states of p fermions, we have:

$$\begin{aligned} \mathbb{H}^+ &= \oplus_{p\ \text{even}}\mathbb{H}^p \\ \mathbb{H}^- &= \oplus_{p\ \text{odd}}\mathbb{H}^p. \end{aligned} \tag{12.57}$$

Then the $\mathbb{Z}_2$-graded exact sequence in Equation (12.52) becomes a $\mathbb{Z}$-graded exact sequence

$$\cdots \xrightarrow{Q} \mathbb{H}^{p-1} \xrightarrow{Q} \mathbb{H}^p \xrightarrow{Q} \mathbb{H}^{p+1} \xrightarrow{Q} \cdots \tag{12.58}$$

and we can define the cohomology group at each p:

$$H^p(Q) = Kernel\left\{Q : \mathbb{H}^p \to \mathbb{H}^{p+1}\right\} / Image\left\{Q : \mathbb{H}^{p-1} \to \mathbb{H}^p\right\}. \tag{12.59}$$

The Witten index then becomes the "Euler" characteristic of the complex

$$Tr\,(-1)^F = \sum_{p=0,1,\cdots} (-1)^p dim\,(H^p(Q)). \tag{12.60}$$

12.3.3 1-d Supersymmetric Quantum Mechanics

Consider the action

$$S = \int dt L(t) = \int dt \left(\frac{1}{2}\dot{x}^2 - \frac{1}{2}\left(h'(x)\right)^2 + \frac{i}{2}\left(\psi^\dagger\dot{\psi} - \dot{\psi}^\dagger\psi\right) - h''(x)\psi^\dagger\psi\right). \tag{12.61}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

The variables ψ and $\psi^\dagger$ are anti-commuting variables which will eventually be realized by the exterior derivative or some deformation of the exterior derivative. For the moment we just impose

$$\left\{\psi, \psi^\dagger\right\} = 0. \tag{12.62}$$

The supersymmetric transformation is

$$\begin{aligned} \delta x &= \epsilon\psi^\dagger - \epsilon^\dagger\psi \\ \delta\psi &= \epsilon\left(i\dot{x} + h'(x)\right) \\ \delta\psi^\dagger &= \epsilon^\dagger\left(-i\dot{x} + h'(x)\right), \end{aligned} \tag{12.63}$$

where $\epsilon = \epsilon_1 + i\epsilon_2$ is a complex fermionic parameter. It is reasonably easy to see that the action is invariant under the supersymmetry transformation. The conserved supercharges can be obtained by Noether's theorem

$$\begin{aligned} Q &= \psi^\dagger\left(i\dot{x} + h'(x)\right) \\ Q^\dagger &= \psi\left(-i\dot{x} + h'(x)\right). \end{aligned} \tag{12.64}$$

Quantizing the system corresponds to imposing the canonical commutation and anti-commutation relations

$$\begin{aligned} [x, p] &= i \\ \left\{\psi, \psi^\dagger\right\} &= 1 \end{aligned} \tag{12.65}$$

as the canonically conjugate momenta are $p = \partial L/\partial\dot{x}$ and $\pi_\psi = \partial L/\partial\dot{\psi} = i\psi^\dagger$, with $\{\psi, \pi_\psi\} = i$. (The convention taken with Grassmann derivatives is action from the left, $\partial\psi_1\psi_2/\partial\psi_1 = -\psi_2$, in the final analysis, it is just the algebra of the operators that counts.) The Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2}p^2 + \frac{1}{2}\left(h'(x)\right)^2 + \frac{1}{2}h''(x)\left(\psi^\dagger\psi - \psi\psi^\dagger\right). \tag{12.66}$$

The fermion number operator is $F = \psi^\dagger\psi$ and satisfies the commutation relations

$$[F, \psi] = -\psi, \quad [F, \psi^\dagger] = \psi^\dagger. \tag{12.67}$$

As $\{\psi, \psi\} = 0 = \left\{\psi^\dagger, \psi^\dagger\right\}$ the fermionic fields satisfy the algebra of fermionic annihilation and creation operators and, if there exists the state $|0\rangle$ that is annihilated by ψ, which we assume $\psi|0\rangle = 0$, then the state $\psi^\dagger|0\rangle$ is the only other independent state in the theory. Evidently $\psi\psi^\dagger|0\rangle = |0\rangle$ and $\psi^\dagger\psi^\dagger|0\rangle = 0$. Thus we can write the fermionic operators as

$$\psi = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \psi^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{12.68}$$

The full Hilbert space of the theory will be the Hilbert space of the bosonic variable x, which is the space of complex-valued square-integrable functions of

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

the variable x denoted by $L^2(\mathbf{R},\mathbb{C})$, multiplying (tensored with the states $|0\rangle$ and the state $\psi^\dagger|0\rangle$),

$$\mathbb{H} = L^2(\mathbf{R},\mathbb{C})|0\rangle \oplus L^2(\mathbf{R},\mathbb{C})\psi^\dagger|0\rangle \tag{12.69}$$

the first component is identified with the bosonic subspace and the second with the fermionic subspace. The supercharges remain form-invariant from their classical expressions

$$\begin{aligned} Q &= \psi^\dagger\,(ip + h'(x)) \\ Q^\dagger &= \psi\,(-ip + h'(x))\,, \end{aligned} \tag{12.70}$$

and commute with the Hamiltonian. We can compute, with a little straightforward algebra, that indeed

$$\left\{Q, Q^\dagger\right\} = 2\mathcal{H}, \tag{12.71}$$

hence the supersymmetry algebra is satisfied.

The supersymmetric ground states are determined by the two conditions:

$$\begin{aligned} Q|E_0\rangle &= \begin{pmatrix} 0 & 0 \\ d/dx + h'(x) & 0 \end{pmatrix} |E_0\rangle = 0 \\ Q^\dagger|E_0\rangle &= \begin{pmatrix} 0 & -d/dx + h'(x) \\ 0 & 0 \end{pmatrix} |E_0\rangle = 0. \end{aligned} \tag{12.72}$$

Expanding $|E_0\rangle = \xi_1(x)|0\rangle + \xi_2(x)\psi^\dagger|0\rangle$ gives

$$\begin{aligned} \left(\frac{d}{dx} + h'(x)\right)\xi_1(x) &= 0 \\ \left(-\frac{d}{dx} + h'(x)\right)\xi_2(x) &= 0, \end{aligned} \tag{12.73}$$

which are trivially solved as

$$\begin{aligned} \xi_1(x) &= c_1 e^{-h(x)} \\ \xi_2(x) &= c_2 e^{h(x)}. \end{aligned} \tag{12.74}$$

Obviously these solutions cannot both be square-integrable and the square-integrability depends on the behaviour of $h(x)$ as $x \to \pm\infty$. The four cases are $\lim_{x\to\pm\infty} h(x) = \pm\infty$, $\lim_{x\to\pm\infty} h(x) = \mp\infty$, $\lim_{x\to\pm\infty} h(x) = +\infty$ and $\lim_{x\to\pm\infty} h(x) = -\infty$, the first case being equivalent to the second. The first two cases yield no square-integrable solution and hence there are no supersymmetric ground states and $Tr(-1)^F = 0$. The latter two yield a solution with either $c_2 = 0$ or $c_1 = 0$; in each case there is one supersymmetric ground state, bosonic if $c_2 = 0$ yielding $Tr(-1)^F = 1$ and fermionic if $c_1 = 0$ yielding $Tr(-1)^F = -1$. Thus we know exactly that in the first two cases there are no supersymmetric ground states, while in the latter two cases there is exactly one, which is bosonic if the potential rises to $+\infty$ as $x \to \pm\infty$ and fermionic if the potential falls to $-\infty$ as $x \to \pm\infty$. We underline that these are exact results.

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

12.3.3.1 Supersymmetric harmonic oscillator The example of a harmonic oscillator is particularly simple; here we take $h(x) = \frac{\omega}{2}x^2$. Then the potential in our Hamiltonian is $\frac{1}{2}(h'(x))^2 = \frac{\omega^2}{2}x^2$ while the coefficient of the fermionic term is $h''(x) = \omega$. Thus the Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2}p^2 + \frac{\omega^2}{2}x^2 + \frac{1}{2}\omega\left(\psi^\dagger\psi - \psi\psi^\dagger\right). \tag{12.75}$$

The harmonic oscillator has spectrum

$$\mathcal{E}_n = \left(n + \frac{1}{2}\right)\omega \quad n = 0, 1, 2, \cdots \tag{12.76}$$

for eigenstate $\phi_n(x)$, which are known to be Hermite polynomials multiplied by a Gaussian. The fermionic part yields the matrix

$$\frac{\omega}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{12.77}$$

which commutes with the harmonic oscillator and has the spectrum $\tilde{\mathcal{E}} = \left(-\frac{\omega}{2}, \frac{\omega}{2}\right)$. Thus the spectrum of the Hamiltonian is for $\omega > 0$,

$$E_n = \begin{cases} n\omega & \text{for } \phi_n(x)|0\rangle \\ (n+1)\omega & \text{for } \phi_n(x)\psi^\dagger|0\rangle \end{cases} \quad n = 0, 1, 2, \cdots \tag{12.78}$$

and for $\omega < 0$

$$E_n = \begin{cases} (n+1)|\omega| & \text{for } \phi_n(x)|0\rangle \\ n|\omega| & \text{for } \phi_n(x)\psi^\dagger|0\rangle \end{cases} \quad n = 0, 1, 2, \cdots. \tag{12.79}$$

We notice that for positive ω we have a bosonic zero mode but for negative ω the supersymmetric zero mode is fermionic.

12.3.4 A Useful Deformation

We will next consider a deformation where the supersymmetric harmonic oscillator corresponds to the lowest-level approximation. Consider the theory with $h(x)$ replaced with $th(x)$, where t is just a parameter (in no sense the time).

$$h(x) \to th(x). \tag{12.80}$$

Then the Hamiltonian becomes

$$\mathcal{H}_t = \frac{1}{2}p^2 + \frac{t^2}{2}\left(h'(x)\right)^2 + \frac{t}{2}h''(x)\left(\psi^\dagger\psi - \psi\psi^\dagger\right) \tag{12.81}$$

and we are interested in what happens as $t \to \infty$. In this limit, the potential $\frac{t^2}{2}\left(h'(x)\right)^2$ becomes very large for most values of x, and the wave function is pushed into regions where the potential is small. The potential is small only at

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

the places x_i where $h'(x_i) = 0$, *i.e.* critical points of the function $h(x)$. Around critical points, the potential can be approximated in the lowest approximation as a quadratic polynomial $\sim (x - x_i)^2$, which brings us back to the harmonic oscillator that we have just analysed. The frequency of the harmonic oscillator becomes $t\omega$, where $\omega = h''(x_i)$ at the critical point, and then the energy levels are linear in t. The fermionic term now has coefficient $th''(x_i) = t\omega$, and thus also gives a linear contribution in t to the energy, which exactly cancels the oscillator ground-state zero-point energy for the bosonic case if $\omega > 0$ and for the fermionic case if $\omega < 0$, just as we have seen explicitly above for the exact harmonic oscillator.

Thus we are left with exactly one energy level at each critical point whose energy does not scale linearly with t. The energy of the state is zero in the approximation that we have employed. It may well be exactly zero, but this is not yet determined. However, we do know that without any approximations there is only one or no exact supersymmetric ground state in the theory, depending on the asymptotic behaviour of $h(x)$. Thus all or all but one of the zero-energy levels that we have found approximately must in fact have non-zero energy. What will be clear is that the exact energy levels of the corresponding exact eigenstates, which are concentrated about the critical points of $h(x)$ (as we have found to be approximately the case), will not scale linearly with t. To first order in the approximation, they are zero-energy modes. Perturbatively, they will actually remain zero-energy modes to all orders. Their energy can only become non-zero through non-perturbative corrections. These non-perturbative corrections are just instanton corrections, corresponding to tunnelling transitions between the perturbative zero-energy modes.

Expanding the function $h(x)$ about a critical point x_i where $h'(x_i) = 0$, and assuming $h''(x_i) \neq 0$, which simply means that the critical points are non-degenerate, we have

$$h(x) = h(x_i) + \frac{1}{2}h''(x_i)(x - x_i)^2 + \frac{1}{6}h'''(x_i)(x - x_i)^3 + \cdots \tag{12.82}$$

and evidently

$$h'(x) = h''(x_i)(x - x_i) + \frac{1}{2}h'''(x_i)(x - x_i)^2 + \cdots . \tag{12.83}$$

Scaling $x - x_i \to \tilde{x} - \tilde{x}_i = (x - x_i)/\sqrt{t}$ and correspondingly $p \to \tilde{p} = \sqrt{t}p$ gives

$$h(x) = h(x_i) + \frac{1}{2t}h''(x_i)(\tilde{x} - \tilde{x}_i)^2 + \frac{1}{6t^{3/2}}h'''(x_i)(\tilde{x} - \tilde{x}_i)^3 + o\left(\frac{1}{t^2}\right) \tag{12.84}$$

and for the Hamiltonian

$$\begin{aligned}\mathcal{H}_t = t\left(\frac{1}{2}\tilde{p}^2 + \frac{1}{2}\left(h''(x_i)\right)^2(\tilde{x} - \tilde{x}_i)^2 + \frac{1}{2}h''(x_i)\left(\psi^\dagger\psi - \psi\psi^\dagger\right)\right) \\ + o(\sqrt{t}) + o(1) + o\left(\frac{1}{\sqrt{t}}\right) + \cdots .\end{aligned} \tag{12.85}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

Thus we can imagine computing perturbatively in $1/\sqrt{t}$ where the leading term is given by

$$\begin{aligned}\mathcal{H}_{local} &= t\left(\frac{1}{2}\tilde{p}^2 + \frac{1}{2}\left(h''(x_i)\right)^2(\tilde{x}-\tilde{x}_i)^2 + \frac{1}{2}h''(x_i)\left(\psi^\dagger\psi - \psi\psi^\dagger\right)\right) \\ &= \frac{1}{2}p^2 + t^2\frac{1}{2}\left(h'(x_i)\right)^2 + t\frac{1}{2}h''(x_i)\left(\psi^\dagger\psi - \psi\psi^\dagger\right) \end{aligned} \tag{12.86}$$

where in the last equality we have put back $\tilde{x} \to x = \sqrt{t}\tilde{x}$ after shifting so that the critical point occurs at $x = 0$. Obviously this is the supersymmetric harmonic oscillator that we have just treated and completely understand. There will be, in this approximation, one bosonic supersymmetric ground state of zero energy, as in Equation (12.78), for each critical point with $h''(x_i) > 0$ and one fermionic supersymmetric ground state of zero energy, as in Equation (12.79), for each critical point with $h''(x_i) < 0$. The eigenstate, say if bosonic, will be of the form (unnormalized)

$$|E_0\rangle \approx e^{-\frac{t}{2}h''(x_i)(x-x_i)^2}|0\rangle, \tag{12.87}$$

which is the first approximation to the exact zero-energy state (unnormalized) which in this case is

$$|E_0\rangle = e^{-th(x)}|0\rangle \tag{12.88}$$

but with $h(x)$ expanded about x_i with the constant value of $h(x_i)$ absorbed into the normalization. Evidently, if we compute the perturbative corrections to the energy state in Equation (12.87), using the perturbatively (in $1/\sqrt{t}$) expanded Hamiltonian (12.85), we will simply rebuild the exact zero-energy state given in Equation (12.88) from a Taylor expansion of $h(x)$. However, at each stage of the perturbative calculation the wave function will be concentrated around $x = x_i$, a Gaussian multiplied by polynomial corrections corresponding to the higher levels of the harmonic oscillator. The energy admits an expansion in even powers of $1/\sqrt{t}$ since the contribution from odd powers vanishes due to parity. However, the energy must actually remain zero at all stages of the perturbation, since we know that the energy of the exact wave function is exactly zero. Perturbative contributions at a higher order cannot correct a non-zero contribution to the energy at a lower order, hence the correction to the energy must be absent at each order. We can do this calculation around each critical point and, hence, perturbatively we will construct as many zero-energy modes as there are critical points.

Since we know that in fact there is at most only one exact zero-energy mode, all but one combination of these perturbatively found zero modes must be non-perturbatively corrected to finite energy. The Witten index will be given as

$$Tr(-1)^F = \sum_{i=1}^{N} \text{sign}(h''(x_i)) \tag{12.89}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

as each bosonic zero mode for $h''(x_i) > 0$ contributes $+1$ and each fermionic zero mode for $h''(x_i) < 0$ contributes -1. Evidently, this sum must equal ± 1 or 0, as we have found, dependent on the asymptotic behaviour of $h(x)$. This makes perfect sense as the number of concave and convex critical points can only change in equal numbers if we deform $h(x)$ locally, as long as the asymptotic behaviour of $h(x)$ is kept invariant. As we have said, these perturbatively found zero modes must not be exact zero modes, thus they must lift away from zero energy due to non-perturbative corrections. But then supersymmetry imposes that for each bosonic mode lifting away from zero energy there must be a corresponding fermionic one that is exactly degenerate in (non-zero) energy. Thus the non-perturbative corrections must simultaneously lift the bosonic and fermionic perturbatively found zero modes away from zero energy in pairs.

The generalization of this theory to n dimensions and on a Riemannian manifold will bring us to Morse theory in the next section.

12.4 Morse Theory

There is a connection between the Betti numbers and critical points of real-valued functions defined on a manifold [94, 10, 89, 125, 62, 19, 64]. We do not consider arbitrary real-valued functions, but an essentially generic class of functions that are called Morse functions. Morse functions, for which we will use the notation $h(x)$, are defined to be those real-valued functions that have a finite number of non-degenerate, isolated critical points. A critical point is where the first derivative of the function vanishes, which evidently is independent of the system of coordinates that are used. Thus the critical points of a Morse function occur at a finite number of discrete points, P_a, and the condition that they be non-degenerate means that the determinant of the matrix of second derivatives in any system of local coordinates containing P_a, the so-called Hessian matrix, has a non-zero determinant. This means the eigenvalues of the Hessian are non-zero. We can diagonalize the Hessian, a real symmetric matrix, by an orthogonal transformation of the coordinates, and shift the coordinates so that the critical point occurs at the origin of the coordinates. We can also rescale the resulting coordinates so that the positive eigenvalues are $+1$ and the negative eigenvalues are -1. Then around a critical point P_a with p negative directions, there exists a coordinate system in which a Morse function appears as

$$h(x) = c_a - \sum_{i=1}^{p} x_i^2 + \sum_{i=p+1}^{n} x_i^2, \tag{12.90}$$

where c_a is the value of the Morse function at the critical point. This rather reasonable fact corresponds to the Morse Lemma. It is clear that with an infinitesimal deformation of the Morse function, all values of the function c_a at the critical points can be taken to be distinct. Furthermore, we can assume

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

that the values of the Morse function at the critical points are labelled in a monotone, ascending order, $c_l < c_{l+1}$, $l = 0, \cdots, N$. p is called the Morse index of a critical point, and the number of critical points with Morse index p is called the M_p.

A surprising fact corresponds to the understanding that the manifold can be reconstructed out of any Morse function that is defined on it. One considers the inverse map defining the submanifold (not including its boundary)

$$\mathcal{M}_c = \{x \in \mathcal{M} \ni h(x) < c\}. \tag{12.91}$$

Clearly for $c < c_0$, where c_0 is the global minimum of the Morse function, $\mathcal{M}_{c<c_0} = \varnothing$. The global minimum must exist as the manifold is assumed to be compact. As we increase c, when we pass c_0, but stay below the next critical point where the value of the Morse function is c_1, the manifold $\mathcal{M}_{c_1>c>c_0}$ is topologically always the same and what is called a 0-cell. The nomenclature, 0-cell, corresponds to the fact that the critical point which is the global minimum has 0 negative directions. A 0-cell is in fact topologically an n-dimensional ball, without its boundary. It is evident that the topology of $\mathcal{M}_{c>c_0}$ does not change as we increase c, until we come to the value of c_1, the next critical point of the Morse function. At the critical point c_1, there are p negative directions and $n-p$ positive directions. The manifold must attach a p-cell to the 0-cell that rises from the global minimum and the topology of the manifold must change as c passes from below c_1 to above c_1 by the attachment of a p-cell. A p-cell corresponds to a topological manifold that has p negative directions and $n-p$ positive directions, such a manifold is sometimes called a p-handle.

This construction will continue at each critical point of the Morse function. The topology of the set $\mathcal{M}_c$ will be invariant for $c_l < c < c_{l+1}$, the topology change occurring exactly and only at the critical points of the Morse function with values c_l. At each critical point of the Morse index p we will have to attach a p-cell. Finally, for $c > c_N$, where c_N is the global maximum of the Morse function,

$$\mathcal{M}_{c>c_N} = \mathcal{M} \tag{12.92}$$

and at this point we will have reconstructed the entire manifold. As we approach the final critical point, we must attach an n-cell, as the global maximum has n negative directions. An n-cell is also, topologically, an n-dimensional ball, as was the 0-cell at the global minimum, except that it now has n negative directions. Nothing precludes the attachment of n-cells, 0-cells or in general any number of p-cells at intermediate critical points; if there are local critical points with p negative directions, that is what is required. Indeed, in principle, for a critical point of the Morse index p, we must add a p-cell. The detailed description of this attachment of p-cells, or p-handles as they are sometimes called, is rather straightforward and unremarkable. We will not describe it in any more detail. The reader can consult the literature cited above for the full details.

Obviously, the reconstruction of the manifold based on a given Morse function must obey some constraints imposed on it due to the actual global topology of the manifold. The actual global topology of the manifold cannot arbitrarily change by its reconstruction based on a given Morse function. The actual topology of the manifold specifically constrains how many p-cells exist in the manifold. Hence the reconstruction based on a Morse function must be in some sense redundant. This gives the first hint that the number of critical points with Morse index p must be restricted by the global topology of the manifold.

The crudest example of such a restriction is, for example, the condition that there must exist only one global maximum and one global minimum for the Morse function. The topology of the manifold, that it is compact, imposes this condition. As any Morse function on the manifold can be interpreted as a height function, with a corresponding topology preserving deformation of the manifold, we can easily see that it is possible to eliminate pairwise, for example, a local maximum and a local minimum by simply deforming the Morse function or equivalently the manifold. Indeed, we will be able to show that the number of critical points, M_p, of the Morse index p is bounded below by exactly the topological properties of the manifold expressed in the Betti number B_p,

$$M_p \geq B_p. \tag{12.93}$$

These correspond to the weak Morse inequalities. There are also strong Morse inequalities, which we will introduce when appropriate in the sequel.

12.4.1 Supersymmetry and the Exterior Algebra

The realization of supersymmetry that we will use corresponds to the following identification in the exterior algebra of a Riemannian manifold, $\mathcal{M}$, of dimension, n, where we will further assume that it is equipped with a smooth metric g_{ij}. Let $Q = d$, $Q^\dagger = \delta$, and

$$Q_1 = d + \delta \quad Q_2 = i(d - \delta), \quad \mathcal{H} = d\delta + \delta d. \tag{12.94}$$

Then

$$\mathcal{H} = Q_1^2 = Q_2^2 \quad \text{and} \quad \{Q_1, Q_2\} = 0, \tag{12.95}$$

i.e. the supersymmetry algebra is satisfied. p-forms are bosonic or fermionic depending on whether p is even (bosonic) or odd (fermionic). The Q_i map bosonic states to fermionic states.

What are the supersymmetric ground states for this quantum-mechanical theory? Evidently they are the zero modes of the Laplacian, those p-forms that are annihilated by the Laplacian, $\mathcal{H} = d\delta + \delta d = \nabla^2$. But these are just the harmonic forms. The harmonic forms satisfy exactly

$$\nabla^2 \rho_p = 0. \tag{12.96}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

The number of harmonic p-forms is exactly the dimension of the pth homology group, $dim H_p(\mathcal{M}, \mathbf{R})$. Hence the number of supersymmetric ground states, $dim H^p(Q)$, of a supersymmetric quantum mechanics defined on a Riemannian manifold, $\mathcal{M}$, is exactly equal to the Betti numbers, B_p, of the manifold. Interestingly, supersymmetry has some relation to the global topology of the manifold as defined by the Betti numbers.

The Witten index is obviously a topological invariant, the number of supersymmetric ground states can only change by pairs of bosonic–fermionic states lifting away from zero energy or coming down to zero energy. Therefore, we see that the Euler characteristic

$$\chi(\mathcal{M}) = \sum_{p=0}^{n} (-1)^p B_p = \sum_{p=0}^{n} (-1)^p dim H^p(Q) = Tr\,(-1)^F \tag{12.97}$$

is in fact a topological invariant of the manifold.

12.4.2 The Witten Deformation

We deform the exterior algebra with an additional real parameter, t, and an arbitrary smooth real-valued function, $h(x)$, defined on $\mathcal{M}$, which will be the appropriate Morse function, and then we let

$$d_t = e^{-ht} d e^{ht} \quad \delta_t = e^{ht} \delta e^{-ht}. \tag{12.98}$$

These operators continue to satisfy $d_t^2 = 0 = \delta_t^2$, and so we define

$$Q_{1t} = d_t + \delta_t, \quad Q_{2t} = i(d_t - \delta_t), \quad \mathcal{H}_t = d_t \delta_t + \delta_t d_t \tag{12.99}$$

and the supersymmetry algebra is satisfied for each t

$$Q_{1t}^2 = Q_{2t}^2 = \mathcal{H}_t, \quad \{Q_{1t}, Q_{2,t}\} = 0. \tag{12.100}$$

Then with $Q_t = (Q_{1t} - iQ_{2t})/2$, deformed supercharges are given by

$$Q_t = d + t dh \wedge \quad Q_t^\dagger = \delta + t(dh\wedge)^\dagger. \tag{12.101}$$

As before, the exact supersymmetric ground states are those that are exactly annihilated by Q_t and by $Q_t^\dagger$. These would be the analogue of the harmonic forms. In the local coordinate system these are easily determined; for example, for states annihilated by Q_t we need to find p-forms that satisfy

$$Q_t \omega_p = (d + t dh)\omega_p = 0. \tag{12.102}$$

Writing $\omega_p = \omega_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$ we get

$$(\partial_i \omega_{i_1 \cdots i_p} + t \partial_i h \omega_{i_1 \cdots i_p}) dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} = 0. \tag{12.103}$$

This has an evident solution

$$\omega_{i_1 \cdots i_p} = e^{-th} c_{i_1 \cdots i_p}, \tag{12.104}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

where $c_{i_1\cdots i_p}$ is constant, and similarly for $Q_t^\dagger$. However, this does not mean that we have actually found a harmonic form, the coordinate system is in principle only a patch on the manifold. To find the set of harmonic forms is, in general, a complicated exercise. The set of exact supersymmetric ground states does exist and their numbers are given by the corresponding Betti numbers.

We define the Betti numbers, B_p, analogous to the definition of the de Rham cohomology, as the number of linearly independent p-forms that satisfy $d_t\omega_p = 0$, *i.e.* closed with respect to d_t, but which cannot be written as the exterior derivative of a $p-1$-form, *i.e.* $\omega_p \neq d_t\chi_{p-1}$, *i.e.* that are not exact with respect to d_t. The point is that this definition of the Betti numbers is actually independent of the parameter t, the Betti numbers so defined must be equal to their usual values at $t=0$. d_t differs from d by conjugation with an invertible operator e^{ht}, thus the mapping $\omega_p \to e^{-ht}\omega_p$ is an invertible mapping of closed but not exact p-forms in the sense of d, mapped to closed but not exact p-forms in the sense of d_t. The dimensions of these spaces are independent of t.

At each point, P, of the manifold, $\mathcal{M}$, choose a basis, $\{a_k\}$, of the tangent space, T_P. We will also consider the dual basis $\{a^{*k}\}$ of the cotangent space T_P^*. The tangent space basis vectors and the dual space basis vectors can be thought of as operators on the exterior algebra, acting through what is called interior product for the $\{a_k\}$ and through the usual exterior product for the $\{a^{*k}\}$. Thus explicitly we have

$$\begin{aligned} a^{*i} &= dx^i \wedge \\ a_i &= \iota_{\partial/\partial_i}, \end{aligned} \tag{12.105}$$

where the interior product ι_V is defined as

$$\iota_V(\omega_p) = \chi_{p-1} \ni \chi_{p-1}(V_1,\cdots,V_{p-1}) = \omega_p(V,V_1,\cdots,V_{p-1}). \tag{12.106}$$

This is just a fancy way of saying that we contract the vector index on the first index of the differential form. Thus for the present case $V = \delta^i_k a_i = a_k$ and $\omega_p = \omega_{i_1\cdots i_p} a^{*i_1} \wedge \cdots \wedge a^{*i_p}$ then $\iota_{a_k}(\omega_p) = \omega_{k,i_2\cdots i_p} a^{*i_2} \wedge \cdots \wedge a^{*i_p}$. Even more explicitly

$$a_k(a^{*i_1} \wedge \cdots \wedge a^{*i_p}) = \sum_{l=1}^{p} (-1)^{l-1} \delta^{i_l}_k a^{*i_1} \wedge \cdots \wedge a^{*i_{l-1}} \wedge a^{*i_{l+1}} \wedge \cdots a^{*i_p}. \tag{12.107}$$

The operators $\{a^{*k}\}$ are dual to the $\{a_k\}$, and their action on the exterior algebra corresponds simply to exterior multiplication. Explicitly, the action on a given p-form is simply given by $a^{*k}(\omega_p) = a^{*k} \wedge \omega_p$. These operators play the role of fermion creation and annihilation operators. The supercharges can be written in this notation as

$$Q_t = d + t\partial_i h a^{*i} \wedge \quad \text{and} \quad Q_t^\dagger = \delta + t g^{ij} \partial_i h a_j. \tag{12.108}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

The function $h(x)$ can be differentiated in the coordinate system, then one can calculate in a straightforward, but somewhat tedious, manner,

$$\mathcal{H}_t = d\delta + \delta d + t^2 g^{ij}\partial_i h \partial_j h + t g^{jk} D_i D_j h \left[a^{*i}, a_k\right], \tag{12.109}$$

where g^{ij} is the assumed Riemannian metric on the manifold $\mathcal{M}$ and D_i is the covariant derivative with respect to the Levi–Civita connection associated to the metric, explicitly, $D_i D_j h = D_i \partial_j h = \partial_i \partial_j h - \Gamma^l_{ij}\partial_l h$ with $\Gamma^i_{jk} = \frac{1}{2} g^{il}\left(\partial_i g_{lj} + \partial_j g_{li} - \partial_l g_{ij}\right)$. For large t, the potential $t^2 g^{ij}\partial_i h \partial_j h$ dominates, and the wave function concentrates about the minima (critical points) of this potential. Corrections can be computed as an expansion in powers of $1/\sqrt{t}$, exactly as in the one-dimensional case.

12.4.3 The Weak Morse Inequalities

$h(x)$ will be called the Morse function, and we will assume it is non-degenerate, meaning that it only has isolated critical points at coordinates x^a, at which $\partial_j h(x^a) = 0$. Therefore, at each critical point the matrix of second derivatives, $D_i D_j h$, must be non-singular, *i.e.* it does not have any vanishing eigenvalues. We define M_p to be the number of critical points with p negative eigenvalues. The first Morse inequality states that $M_p \geq B_p$, which we will be able to prove with our supersymmetric quantum mechanical model.

Let $\lambda_p^{(n)}(t)$ be the nth smallest eigenvalue of $\mathcal{H}_t$ acting on p-forms. We will see that

$$\lambda_p^{(n)}(t) = t A_p^{(n)} + o(1) + o(1/t), \tag{12.110}$$

which admits an expansion in powers of $1/t$ due to parity. The Betti number, B_p, is equal to the number of exactly zero eigenvalues. For large t, the number of the eigenvalues that vanish can be no larger than the number of $A_p^{(n)}$ that vanish, simply because a vanishing eigenvalue requires $A_p^{(n)} = 0$. We will show that the number of $A_p^{(n)}$ that vanish is equal to the number of critical points of the Morse function with p negative eigenvalues, which means that $M_p \geq B_p$.

At each critical point we can use Gaussian normal coordinates x^i, coordinates in which the metric is simply δ_{ij} and shift the origin so that they are chosen to vanish at the position of the critical point. The Morse function can be expanded in a Taylor series; in general, this gives

$$h(x) = h(0) + \frac{1}{2}\sum_{i,j=0}^{n}\left(\frac{\partial^2}{\partial_i \partial_j} h(0)\right) x^i x^j + \cdots. \tag{12.111}$$

A further orthogonal rotation of the coordinates keeps the metric δ_{ij}; however, the real symmetric matrix of second partial derivatives can be diagonalized, with eigenvalues $\lambda_i = \partial_i \partial_i h(0)$ in the new coordinates. Then we get

$$h(x) = h(0) + \frac{1}{2}\sum_i \lambda_i (x^i)^2 + \cdots \tag{12.112}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

and

$$\partial_i h(x) = \lambda_i x^i + \cdots . \tag{12.113}$$

The Hamiltonian then also admits a local expansion about each critical point, using the general expression Equation (12.109) and noting that the metric is δ_{ij}, the Levi–Civita connection vanishes so that covariant derivatives are ordinary derivatives and $\partial_i h(x^i) = \lambda_i x^i + \cdots$ using Equation (12.112)

$$\mathcal{H}_t = \sum_{i=1}^{n} \left(-\frac{\partial^2}{\partial x^i \partial x^i} + t^2 \lambda_i^2 (x^i)^2 + t\lambda_i [a^{*i}, a_i] \right) + \cdots . \tag{12.114}$$

The explicitly written term, although an approximation to the full Hamiltonian, is sufficient to compute the $A_p^{(n)}$. To compute the expansion of the eigenvalues in powers of $1/t$ requires calculating the higher-order terms in the Hamiltonian and continuing the perturbative expansion.

As the operators a_i and a^{*i} are also simply linear operators on the exterior algebra by exterior or interior multiplication, they commute with the simple harmonic oscillator part and hence the local Hamiltonian in lowest approximation can be written as two commuting terms

$$\mathcal{H}_{local} = \sum_{i=1}^{n} H_i + t\lambda_i K_i \tag{12.115}$$

with

$$H_i = -\frac{\partial^2}{\partial x^i \partial x^i} + t^2 \lambda_i^2 (x^i)^2 \tag{12.116}$$

while

$$K_i = [a^{*i}, a_i]. \tag{12.117}$$

H_i is the Hamiltonian of the simple harmonic oscillator, with the well-known spectrum $E_i(N_i) = t|\lambda_i|(1+2N_i)$, where $N_i = 0, 1, 2, \cdots$, taking into account that the λ_i are not necessarily positive. The corresponding eigenfunctions are Hermite polynomials multiplied by Gaussians centred at the origin, and hence rapidly fall off for $|\lambda_i x^i| \gg 1/\sqrt{t}$.

The eigenvalues of K_i are simply ± 1. The action of K_i on a p-form ω is $K_i\omega = [a^{*i}, a_i]\omega = a^{*i}a_i\omega - a_i a^{*i}\omega = 2a^{*i}a_i\omega - \{a_i, a^{*i}\}\omega = (2a^{*i}a_i - 1)\omega$. The first operator is simply the fermionic Hamiltonian for one degree of freedom for each i, which has eigenspectrum 0 or 2, acting on the fermionic vacuum state or the one fermion state, which yields the eigenspectrum ± 1 for K_i. Another way to see this is to realize that the action of K_i on a p-form $\omega = \omega_{i_1 \cdots i_p} a^{*i_1} \wedge \cdots \wedge a^{*i_p}$ obviously gives back ω if $i \in (i_1 \cdots i_p)$ but gives back $-\omega$ if $i \notin (i_1 \cdots i_p)$.

As K_i and H_i commute, the eigenvalues simply add; thus, the spectrum of $\mathcal{H}_{local}$ is

$$E_t(N_i, n_i) = t \sum_{i=1}^{n} (|\lambda_i|(1+2N_i) + n_i\lambda_i), \quad N_i = 0, 1, 2, \cdots, n_i = \pm 1. \tag{12.118}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

If we restrict the action of $\mathcal{H}_{local}$ to p-forms, then the sum over K_i in the Hamiltonian contains p terms for which the eigenvalue of K_i is $+1$; thus, the number of n_i that equal $+1$ must be equal to p. The remaining K_i will have eigenvalue -1, thus the number of these will be $n-p$, where n is the dimension of $\mathcal{M}$.

The only way it is possible for the energy $E_t(N_i, n_i)$ to vanish is if all $N_i = 0$, $n_i = 1$ for each negative λ_i and $n_i = -1$ for each positive λ_i. We can solve this constraint if we choose the p-form to consist of the p-fold exterior product of coordinate differentials of exactly those coordinate directions which correspond to the negative eigenvalues. Thus the energy eigenvalue is (allowing for a minor relabelling of the independent directions in the manifold)

$$E_t(N_i, n_i) = t\sum_{i=1}^{p}(|\lambda_i|(1) + \lambda_i) + t\sum_{i=p+1}^{n}(|\lambda_i|(1) - \lambda_i) = 0 \tag{12.119}$$

as $n_i = +1$ for the first p directions with negative eigenvalues and $n_i = -1$ for the $n-p$ remaining directions for which the eigenvalues are positive.

Thus for a critical point with Morse index equal to p, *i.e.* with p negative directions at the critical point of the Morse function h, it is possible to satisfy these conditions. We choose a p-form with a coefficient function given by the ground state of the harmonic oscillator (which puts all the $N_i = 0$), and which consists of exactly those coordinate differentials which correspond to the p negative directions, λ_i, which gives the desired $n_i = +1$. Thus at a critical point of Morse index p, we can construct exactly one eigenfunction which could have a zero eigenvalue. These are zero-energy eigenfunctions of the approximate Hamiltonian given in Equation (12.115). We could, in principle, compute the corrections that are brought to these approximate zero-energy levels, but we can be assured that they will remain low-lying levels even as $t \to \infty$, the key point being that $A_p^{(n)}$ vanishes for all of these levels. The dimension of the subspace spanned by these levels is M_p, the number of critical points with Morse index p.

For an actual vanishing eigenvalue of the full Hamiltonian (12.109), all higher perturbative and non-perturbative corrections must also vanish. This will happen for the exact supersymmetric ground states. The number of exact supersymmetric ground states is given by the Betti number, B_p, which is equal to the number of p-forms with zero eigenvalues of the Laplacian, $d\delta + \delta d$, or the deformed Laplacian, $d_t\delta_t + \delta_t d_t$. For each actual zero eigenvalue, we know that the $A_p^{(n)}$ must also vanish, as the computation of $A_p^{(n)}$ is the first step of computing the exact eigenvalue in perturbation. We have determined that the number of approximate states corresponding to $A_p^{(n)} = 0$ is M_p, the number of critical points of Morse index p. Hence the number of actual zero eigenvalue states must be less than or equal to M_p. Thus we obtain the result

$$M_p \geq B_p. \tag{12.120}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

These are called the weak Morse inequalities.

12.4.4 Polynomial Morse Inequalities

We actually wish to prove something stronger, that the Morse numbers always dominate the Betti numbers as encapsulated in the polynomial Morse inequality which states that there exists a set of non-negative integers Q_p such that

$$\sum_{p=0}^{n} M_p t^p - \sum_{p=0}^{n} B_p t^p = (1+t)\sum_{p=0}^{n-1} Q_p t^p. \tag{12.121}$$

This is an inequality in the sense that $Q_p \geq 0$. As the weak Morse inequalities give us that $M_p \geq B_p$, it is clear that the coefficient of t^p on the left-hand side is necessarily positive semi-definite. The right-hand side has the coefficient $Q_p + Q_{p-1}$ (with $Q_n = Q_{-1} = 0$) for t^p, which then must be positive semi-definite.

The polynomial Morse inequality is equivalent to the following two assertions, called the strong Morse inequalities (as originally proven by Morse):

$$\sum_{p=0}^{m}(-1)^{p+m} M_p \geq \sum_{p=0}^{m}(-1)^{p+m} B_p \quad \text{for} \quad m = 0, 1, \cdots, n \tag{12.122}$$

$$\sum_{p=0}^{n}(-1)^{p} M_p = \sum_{p=0}^{n}(-1)^{p} B_p. \tag{12.123}$$

We can prove the equivalence as follows. If we take the second equality, Equation (12.123), we have

$$\left(\sum_{p=0}^{n} M_p t^p - \sum_{p=0}^{n} B_p t^p\right)\Bigg|_{t=-1} = 0, \tag{12.124}$$

i.e. $t = -1$ is a root of the polynomial $\sum_{p=0}^{n} M_p t^p - \sum_{p=0}^{n} B_p t^p$ and hence it is divisible by $1+t$. Thus we have immediately and trivially

$$\sum_{p=0}^{n} M_p t^p - \sum_{p=0}^{n} B_p t^p = (1+t)\sum_{p=0}^{n-1} Q_p t^p. \tag{12.125}$$

Since the coefficients are integers on the left-hand side, the Q_n must also be integers. It remains to show that $Q_n \geq 0$. To see this we analyse the identity power by power in t. We start with the t^0 term. This term gives

$$M_0 - B_0 = Q_0, \tag{12.126}$$

which the first inequality, Equation (12.122), for $m = 0$ requires $Q_0 \geq 0$. Next, for the t term we have

$$M_1 - B_1 = Q_1 + Q_0 \tag{12.127}$$

or replacing for Q_0 from above

$$M_1 - M_0 - (B_1 - B_0) = Q_1, \tag{12.128}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

which the first inequality, Equation (12.122), for $m = 1$ then requires $Q_1 \geq 0$. Doing one more step, before concluding the general relation, we have for the coefficient of t^2

$$M_2 - B_2 = Q_2 + Q_1 \tag{12.129}$$

replacing for Q_1 from above

$$M_2 - M_1 + M_0 - (B_2 - B_1 + B_0) = Q_2. \tag{12.130}$$

Again from the first inequality, Equation (12.122), for $m = 2$ then requires $Q_2 \geq 0$. We see then that in general

$$\sum_{p=0}^{m} (-1)^{p+m} M_p - \sum_{p=0}^{m} (-1)^{p+m} B_p = Q_m \quad \text{for} \quad m = 0, 1, \cdots, n-1 \tag{12.131}$$

and hence we can conclude that $Q_m \geq 0$ for all $m = 0, 1, 2, \cdots n-1$.

To prove the converse, the polynomial Morse inequality, Equation (12.121), by comparing powers of t, as we have just seen, implies

$$\sum_{p=0}^{n} (-1)^{p+m} M_p - \sum_{p=0}^{n} (-1)^{p+m} B_p = Q_m \quad \text{for} \quad m = 0, 1, \cdots, n-1 \tag{12.132}$$

but now we assume that the $Q_m \geq 0$. Hence we recover the first inequalities in Equation (12.122) trivially. To recover the second equality, Equation (12.123), we simply put $t = -1$ in Equation (12.121).

The second equality, Equation (12.123), is related to the Euler characteristic of the manifold. This is defined as the alternating sum of the Betti numbers

$$\chi(\mathcal{M}) = \sum_{p=0}^{n} (-1)^p B_p = \sum_{p=0}^{n} (-1)^p M_p. \tag{12.133}$$

From the weak Morse inequalities, we know that $M_p \geq B_p$. Thus the number of critical points of Morse index, p, could be greater than the Betti number, B_p, but then there must be exactly the same surplus of critical points with opposite value of $(-1)^p$, *i.e.* each additional critical point of Morse index, p, must pair with another critical point of Morse index of opposite parity. As p determines if the state is fermionic or bosonic, we identify these pairs of critical points with the approximate, supersymmetric, bosonic and fermionic zero energy pairs of states associated with each critical point, but those which must actually lift away from exact zero energy when non-perturbative corrections are taken into account, as the actual number of supersymmetric zero energy states is strictly given by the Betti numbers.

It is straightforward [64] to prove the strong Morse inequalities using the simple ideas of supersymmetry and what we have understood about the spectrum. We know that the eigenvalues and corresponding eigenstates of $\mathcal{H}_t$ separate into two subsets as $t \to \infty$; those whose energies diverge linearly as t gets large and

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

a finite number whose energies do not. There are M_p states for each p whose energies do not diverge with t. These further split into two subsets, the first B_p states whose energies are exactly zero and the remaining $M_p - B_p$ states whose energies are of $o(1)$. We will call these latter $M_p - B_p$ states the low-lying states. But now we recall that, since the low-lying states have non-zero energy, supersymmetry requires that they come in bosonic–fermionic pairs. The fermionic states correspond to odd p and the bosonic states correspond to even p, hence we must have

$$\sum_{p \text{ odd}} (M_p - B_p) = \sum_{p \text{ even}} (M_p - B_p). \tag{12.134}$$

This immediately implies the second of the strong Morse inequalities, Equation (12.123)

$$\sum_{p=0}^{n} (-1)^p M_p = \sum_{p=0}^{n} (-1)^p B_p. \tag{12.135}$$

To obtain the first strong Morse inequality we consider the mapping that Q_{1t} induces on the fermionic and bosonic subspaces of low-lying levels. As $Q_{1t}^2 = \mathcal{H}_t$ and evidently Q_{1t} commutes with the Hamiltonian, it must preserve the eigensubspaces of $\mathcal{H}_t$. Q_{1t}, being a fermionic operator, maps the eigensubspace of p-forms to the eigensubspace of $p+1$-forms and $p-1$ forms. Let Λ_t^p denote the subspace of low-lying eigenstates of p-forms, clearly of dimension $M_p - B_p$. For any state $|\psi\rangle$ in this subspace $Q_{1t}|\psi\rangle \neq 0$ and $\mathcal{H}_t Q_{1t}|\psi\rangle = Q_{1t}^3|\psi\rangle = Q_{1t}\mathcal{H}_t|\psi\rangle = EQ_{1t}|\psi\rangle$, where $\mathcal{H}_t|\psi\rangle = E|\psi\rangle$. If Q_{1t} maps two distinct states to the same state, then it must annihilate their difference, which is not possible as this does not preserve the eigenspace. Thus the mapping $Q_{1t} : \Lambda_t^p \to \Lambda_t^{p-1} \oplus \Lambda_t^{p+1}$ must be one-to-one, into (injective). Hence we can conclude

$$\begin{aligned} Q_{1t} &: \bigoplus_{p \text{ odd } p=1}^{2j-1} \Lambda_t^p \to \bigoplus_{p \text{ even } p=0}^{2j} \Lambda_t^p \\ Q_{1t} &: \bigoplus_{p \text{ even } p=0}^{2j} \Lambda_t^p \to \bigoplus_{p \text{ odd } p=1}^{2j+1} \Lambda_t^p \end{aligned} \tag{12.136}$$

for each $j \ni 0 \leq 2j < n$ and $0 \leq 2j+1 < n$. But since the mappings are injective, the dimension of the domain must be less than or equal to the dimension of the image. This yields:

$$\begin{aligned} (M_1 - B_1) + \cdots + (M_{2j-1} - B_{2j-1}) &\leq (M_0 - B_0) + \cdots + (M_{2j} - B_{2j}) \\ (M_0 - B_0) + \cdots + (M_{2j} - B_{2j}) &\leq (M_1 - B_1) + \cdots + (M_{2j+1} - B_{2j+1}). \end{aligned} \tag{12.137}$$

These inequalities are identical to the first strong Morse inequality, Equation (12.122), if we bring the Ms and Bs to opposite sides.

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

12.4.5 Witten's Coboundary Operator

The polynomial Morse inequality is equivalent to the understanding that the critical points of a Morse function form a model for the cohomology of the manifold $\mathcal{M}$. We define X_p to be a vector space of dimension M_p for each $p \in 0, 1, 2, \cdots, n$. X_p can be thought of as a vector space spanned by the critical points of Morse index p. The polynomial Morse inequality, Equation (12.121), means that there exists a coboundary operator $\delta_W : X_p \to X_{p+1}$ (we add a subscript W to honour Witten), where $\delta_W^2 = 0$ and the corresponding Betti numbers, the dimension of the cohomology groups associated to δ_W, are identical to the Betti numbers of the manifold $\mathcal{M}$. The homotopy classes in this cohomology are elements of X_p, which are closed under the action of δ_W, but differ only by elements which are obtained by the action of δ_W on some element of X_{p-1}, the analogue of the standard notion of closed modulo exact forms, or cycles, etc. The explicit expression for δ_W is not given in the original work of Morse or others; however, Witten found an appropriate expression for it.

Witten proposed the following construction. First, consider possible zero modes of the Laplacian. The number of independent such p-forms gives the Betti numbers, B_p. We have an upper bound on the Betti numbers, $M_p \geq B_p$ in the Morse inequalities. However, although perturbation theory might suggest a given mode is a zero mode, tunnelling effects can lift the degeneracy. Exact instanton effects can give energies of the order of $\sim e^{-tS}$ where S is the action of the instanton, which for large t is smaller than any perturbative correction. Thus Witten was led to consider instanton configurations that tunnel from one zero mode to another. In fact, tunnelling from putative zero modes which are p-forms to putative zero modes which are $p+1$-forms are exactly the instanton modes that are required. However, as we have seen, the $p+1$-form chosen at a given critical point of Morse index, $p+1$, requires a choice of the exterior product of all the coordinate differentials that correspond to the $p+1$ negative directions. The orientation or order of the differentials remains arbitrary. Thus a tunnelling transition from a state at a critical point of Morse index, p, to a state at a critical point of Morse index, $p+1$, must also fix a sign. We determine the sign with the following construction.

Consider instanton paths, Γ, that pass from a critical point, B, of Morse index, $p+1$, to a critical point, A, of Morse index, p. The instanton path has initial tangent vector v within V_B, the $p+1$-dimensional vector space of negative directions at B. Let $|b\rangle$, a $p+1$-form, be the state of zero energy at the critical point B. Then $|b\rangle$ chooses an orientation of V_B, and we can choose an orientation of the p-dimensional subspace, $\tilde{V}_B$, corresponding to the orthogonal complement of v within $|b\rangle$, as we can generate a p-form from $|b\rangle$ by contracting it (by interior multiplication) with v. Then the instanton path from B to A gives a mapping of $\tilde{V}_B$ to V_A, the p-dimensional vector space of negative directions at A. This

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

mapping induces an orientation of V_A. However, the state $|a\rangle$ corresponding to the perturbative zero mode at A already gave an orientation of V_A. We define

$$n_\Gamma = \begin{cases} +1 & \text{if the induced orientation agrees with that fixed by } |a\rangle \\ -1 & \text{if the induced orientation disagrees with that fixed by } |a\rangle \end{cases} \tag{12.138}$$

and

$$n(a,b) = \sum_\Gamma n_\Gamma, \tag{12.139}$$

where the sum runs over all instantons paths (paths of steepest descent) from B to A. Then we can define the coboundary operator, for any basis element $|a\rangle$ of X_p at A

$$\delta_W |a\rangle = \sum_b n(a,b)|b\rangle, \tag{12.140}$$

where the sum runs over all basis elements of X_{p+1} (in other words, this is a set of perturbative zero modes that are $p+1$-forms that are concentrated at the critical points of Morse index $p+1$ of the Morse function). The effect of the instantons is to non-perturbatively correct the energy of some of the perturbative zero modes, their energy behaves as $\sim e^{-tS}$, for large t. Thus all states in X_p are not annihilated by the Laplacian $\delta_W \delta_W^* + \delta_W^* \delta_W$.

Denoting Y_p as the number of actual zero eigenvalues of $\delta_W \delta_W^* + \delta_W^* \delta_W$ acting on X_p, then Y_p also give upper bounds on the Betti numbers, and the strong Morse inequality, Equation (12.121), remains valid with M_p replaced with Y_p. Witten conjectures that, in fact, $Y_p = B_p$.

12.4.6 Supersymmetric Sigma Model

To demonstrate that δ_W as defined in Equation (12.140) provides the appropriate coboundary operator, Witten considered the Lagrangian version of the supersymmetric quantum-mechanical model that we have been considering, that for which the supercharge is given explicitly by d_t. Canonical quantization of the model defined by the action, in Minkowski time

$$\begin{aligned} \int d\tau \mathcal{L} = \frac{1}{2} \int d\tau \Bigg(g_{ij} \left(\frac{dx^i}{d\tau} \frac{dx^j}{d\tau} + \bar{\psi}^i i \frac{D\psi^j}{D\tau} \right) + \frac{1}{4} R_{ijkl} \bar{\psi}^i \psi^k \bar{\psi}^j \psi^l \\ - t^2 g^{ij} \frac{dh}{dx^i} \frac{dh}{dx^j} - t \frac{D^2 h}{Dx^i Dx^j} \bar{\psi}^i \psi^j \Bigg) \end{aligned} \tag{12.141}$$

where a sum over all repeated indices is understood, gives the required algebraic symmetries and explicitly the supercharge. This is the Lagrangian of the 1+1-dimensional supersymmetric sigma model restricted to 0+1 dimensions. Here the ψ and $\bar{\psi}$ (the complex conjugate field to ψ) are anti-commuting fermionic fields, x^i are local coordinates, g_{ij} is the metric tensor and R_{ijkl} is the corresponding Riemann curvature tensor on $\mathcal{M}$, and D/Dx^i is the covariant derivative with

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

the Levi–Civita connection of the metric while $D/D\tau$ is the covariant derivative along the direction tangent to the time trajectory. Specifically, acting on the fermions we have

$$\frac{D}{D\tau}\psi^i = \partial_\tau \psi^i + \Gamma^i_{jk}\partial_\tau x^j \psi^k. \tag{12.142}$$

Under the supersymmetry transformations

$$\begin{aligned}
\delta x^i &= \epsilon\bar{\psi}^i - \bar{\epsilon}\psi^i \\
\delta\psi^i &= \epsilon\left(i\dot{x}^i - \Gamma^i_{jk}\bar{\psi}^j\psi^k + tg^{ij}\partial_j h\right) \\
\delta\bar{\psi}^i &= \bar{\epsilon}\left(-i\dot{x}^i - \Gamma^i_{jk}\bar{\psi}^j\psi^k + tg^{ij}\partial_j h\right)
\end{aligned} \tag{12.143}$$

for infinitesimal anti-commuting parameters ϵ and $\bar{\epsilon}$, the action is invariant, $\delta\int d\tau\mathcal{L} = 0$. The corresponding supercharges are as required

$$\begin{aligned}
Q_t &= \bar{\psi}^i(ig_{ij}\dot{x}^j + t\partial_i h) \\
\bar{Q}_t &= \psi^i(-ig_{ij}\dot{x}^j + t\partial_i h).
\end{aligned} \tag{12.144}$$

There is also a symmetry-conserving fermion number, $\psi^i \to e^{-i\theta}\psi^i$, $\bar{\psi}^i \to e^{i\theta}\bar{\psi}^i$, which gives the conserved charge, the fermion number

$$F = g_{ij}\bar{\psi}^i\psi^j. \tag{12.145}$$

In quantizing the system we will first consider the system at $t = 0$ (all the supersymmetry and other symmetries are equally valid at $t = 0$). We impose the canonical commutation and anti-commutation relations

$$\begin{aligned}
\left[x^i, p_j\right] &= i\delta^i_j \\
\left\{\psi^i, \bar{\psi}^j\right\} &= g^{ij},
\end{aligned} \tag{12.146}$$

then the conserved supercharges are simply $Q = i\bar{\psi}^i p_i$ and $\bar{Q} = -i\psi^i p_i$. The supercharges have the opposite fermion number

$$[F, Q] = Q,\ \left[F, \bar{Q}\right] = -\bar{Q}. \tag{12.147}$$

We impose that the Hamiltonian is given by the supersymmetry algebra

$$\left\{Q, \bar{Q}\right\} = 2\mathcal{H}_0 \tag{12.148}$$

and consequently the fermion number is conserved, $[F, \mathcal{H}_0] = 0$. The natural realization of this algebra is, as we have been using, provided by the exterior algebra of differential forms, $\Lambda^*(\mathcal{M}) \otimes \mathbb{C}$ equipped with its hermitian inner product from Equation (12.17)

$$(\omega, \chi) = \int_{\mathcal{M}} \bar{\omega} \wedge *\chi \tag{12.149}$$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

for two p-forms, ω and χ. Then the observables in this realization of the algebra on this Hilbert space, when acting explicitly on a p-form ω, are:

$$\begin{aligned} x^i &: x^i\omega \\ p_i &: -i\partial_i\omega \\ \bar{\psi} &: dx^i \wedge \omega \\ \psi^i &: g^{ij}\iota_{\partial/\partial x^j}\omega \end{aligned} \tag{12.150}$$

(ι_V is the interior multiplication defined in Equation (12.106)). Then with the state $|0\rangle$ denoting the form annihilated by all of the ψ^i we have the schema:

$$\begin{aligned} |0\rangle &= 1 \\ \bar{\psi}^i|0\rangle &= dx^i \\ \bar{\psi}^i\bar{\psi}^j|0\rangle &= dx^i \wedge dx^j \\ &\cdots \\ \bar{\psi}^1\cdots\bar{\psi}^n|0\rangle &= dx^1 \wedge \cdots \wedge dx^n. \end{aligned} \tag{12.151}$$

The fermion number of a state that is a p-form is simply equal to p, thus the Hilbert space separates into bosonic and fermionic subspaces depending on whether p is even or odd, respectively. Thus the canonically quantized system reproduced with complete fidelity the supersymmetric system of the exterior algebra that we studied in subsection (12.4.1).

Recall then that the supersymmetric states are just the zero-energy states, those annihilated by the Laplacian, the so-called harmonic forms. We underline that the set of harmonic forms of the manifold characterize the de Rham cohomology of the manifold. Equally well, the space of supersymmetric ground states characterize the cohomology of the Q-operator. As there is the conserved fermion number which satisfies $[F,Q] = Q$, the Q-cohomology is graded by the fermion number and equal to the degree p of the form. As Q is identified with the exterior derivative d, the graded Q-cohomology and the de Rham cohomology must be equal

$$H^p(Q) = H^p_{dR}(\mathcal{M}). \tag{12.152}$$

The Witten index, $(-1)^F$, can be evaluated and we find

$$Tr\left((-1)^F\right) = \sum_{p=0}^{n}(-1)^p dim(H^p(Q) = \sum_{p=0}^{n}(-1)^p dim(H^p_{dR}(\mathcal{M}) = \chi(\mathcal{M}), \tag{12.153}$$

where $\chi(\mathcal{M})$ is the Euler characteristic of the manifold. The Witten index only receives contributions from the supersymmetric ground states; as we have seen, the non-zero energy modes are all paired in fermionic–bosonic pairs and their contributions cancel. Thus the calculation of the topological invariant, the Euler characteristic, can be done by studying the zero-energy modes of this

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

supersymmetric quantum mechanical system. Witten's magical trick was to add an external field to this system, which causes a separation of the zero- and low-energy modes from the finite-energy modes, and in the limiting case makes the calculation of the zero mode sector very simple.

Now adding in the deformation by th, the supercharges are then given by

$$Q_t = \bar{\psi}^i(ig_{ij}\dot{x}^j + t\partial_i h) = dx^i \wedge (\frac{\partial}{\partial x^i} + t\partial_i h) = d + tdh\wedge = e^{-th}de^{th} = d_t \quad (12.154)$$

and

$$Q_t^\dagger = \delta + t(dh\wedge)^* = e^{th}\delta e^{-th} = \delta_t, \quad (12.155)$$

where $*$ denotes the adjoint. The Hamiltonian then is as before

$$\mathcal{H} = \frac{1}{2}\left\{Q_t, \bar{Q}_t\right\} = \frac{1}{2}(d_t d_t^* + d_t^* d_t), \quad (12.156)$$

chosen to satisfy the supersymmetry algebra. The supersymmetric ground states again define the Q_t-cohomology. However, since the th deformation is obtained by a similarity transformation

$$Q_t = e^{-th}Qe^{th} \quad (12.157)$$

the cohomology is isomorphic to the undeformed case. As the cohomology of the undeformed Q is isomorphic to the de Rham cohomology, we can compute the de Rham cohomology with the deformed operator Q_t.

The perturbative approximation to the Hamiltonian around a critical point is given by

$$\mathcal{H}_t = \sum_{i=1}^{n}\left(-\frac{\partial^2}{\partial x^i \partial x^i} + t^2\lambda_i^2(x^i)^2 + t\lambda_i[a^{*i}, a_i]\right) \quad (12.158)$$

with exact, zero-energy ground-state wave functions, which we will label $|\phi\omega\rangle$, corresponding to the harmonic oscillator ground state, ϕ, multiplied by an appropriate p-form, ω, where p is the Morse index of the critical point, as discussed previously. Indeed, perturbative corrections to the energy of these wave functions must vanish to all orders: the energy remains exactly zero to all orders in perturbation theory. One can find the modification of the wave function, order by order, so that its energy remains zero in each order in perturbation theory. This is because the corrections are calculated in terms of local data at the critical point. From local data it is not possible to know which critical points are actually necessary because of the global topology of the manifold and which critical points are removable by deformations. States that have zero energy to lowest order have zero energy to all orders. The same reasoning applies to the calculation of tunnelling in the double-well potential. In perturbation theory we can never get a non-zero tunnelling amplitude, these amplitudes are non-perturbative in the coupling and are not seen at any order in perturbation theory.

However, the wave functions, $|\phi_i\omega_i\rangle$, are not necessarily exact ground states, where we have added a label i to denote different critical points. The perturbative

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

zero-energy states are not necessarily exact ground states. Hence the number of exact, supersymmetric actual ground states are clearly less than or equal to the number of critical points. An exact supersymmetric ground state is annihilated by the supercharge. For the case of the perturbative ground states, although we will find, if calculated perturbatively,

$$\mathcal{H}_t|\phi_i\omega_i\rangle = (Q_tQ_t^\dagger + Q_t^\dagger Q_t)|\phi_i\omega_i\rangle = 0, \tag{12.159}$$

which also requires

$$Q_t|\phi_i\omega_i\rangle = 0 \tag{12.160}$$

to all orders in perturbation theory, we can in fact have non-perturbative corrections

$$Q_t|\phi_i\omega_i\rangle = \sum_{j=1}^{N} |\phi_j\omega_j\rangle\langle\phi_j\omega_j|Q_t|\phi_i\omega_i\rangle + \cdots, \tag{12.161}$$

where the $+\cdots$ corresponds to amplitudes to non-zero-energy states (which are suppressed by large energy denominators as $t \to \infty$). The explicit mixing that can be important is between the perturbative zero-energy states. Thus we want to compute

$$\langle\phi_j\omega_j|Q_t|\phi_i\omega_i\rangle = \int_{\mathcal{M}} \phi_j\omega_j \wedge *(d + tdh\wedge)\phi_j\omega_i. \tag{12.162}$$

But such an amplitude is exactly what we are looking for with the coboundary operator δ_W between zero modes localized at different critical points. If ω_j is a q-form and ω_i is a p-form, this matrix element can only be non-zero if $q = p+1$, *i.e.* transitions between perturbative zero-energy modes correspond to critical points of Morse indices that differ by one negative direction. This can also be seen from fermion number conservation, the action of Q_t on the state $|\phi_i\omega_i\rangle$ changes its fermion number by one unit. It also should not be surprising that the eventual δ_W that we will be able to define will satisfy $\delta_W^2 = 0$, since it is obtained from the action of $Q_t = d_t$. Clearly $Q_t^2 = 0$, hence we can expect $\delta_W^2 = 0$.

We will return below to the notation of subsection 12.4.5 with $|a\rangle$ for $|\phi_i\omega_i\rangle$ and $\langle b|$ for $\langle\phi_j\omega_j|$ and the understanding that if $|a\rangle$ corresponds to a p-form then $\langle b|$ corresponds to a $p+1$-form. It is also clear that the action of Q_t on the low-lying states annihilates any exact, supersymmetric ground state that is a p-form as these are harmonic with respect to Q_t. Thus only the $M_p - B_p$ low-lying but not exact supersymmetric ground states will be mixed with low-lying $p+1$-forms. But additionally, none of these states can be the exact supersymmetric ground states that are $p+1$-forms, since the inner product

$$\langle b|Q_t|a\rangle = (Q_t|b\rangle)^\dagger\,|a\rangle = 0 \tag{12.163}$$

if $|b\rangle$ corresponds to an exact supersymmetric ground state, as these are also harmonic with respect to Q_t. Thus the action of Q_t on the set of states $|a\rangle$ only mixes the $M_p - B_p$ not exact ground states but low-lying states with the $M_{p+1} -$

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

B_{p+1} corresponding low-lying states $|b\rangle$. This is as it should be; since mixing causes the energies to go up, this cannot happen to any exact supersymmetric ground state.

12.4.7 The Instanton Calculation

We will use the path integral to compute this amplitude, since we know that it is exactly through the path integral that we can uncover tunnelling amplitudes through the path integral, and from the amplitude we will extract the coboundary operator, δ_W. The bosonic sector of the model is governed by the Lagrangian, in Euclidean time

$$\int d\tau \mathcal{L}_b = \frac{1}{2}\int d\tau \left(g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} + t^2 g^{ij}\frac{dh}{dx^i}\frac{dh}{dx^j}\right). \tag{12.164}$$

We can show that the stationary points of the corresponding action are the paths of steepest descent using a Bogomolny-type identity [17]. Indeed,

$$\int d\tau \mathcal{L}_b = \frac{1}{2}\int d\tau g_{ij}\left(\frac{dx^i}{d\tau} \pm t g^{ik}\frac{dh}{dx^k}\right)\left(\frac{dx^j}{d\tau} \pm t g^{jl}\frac{dh}{dx^l}\right) \mp t\int d\tau \frac{dh}{d\tau}. \tag{12.165}$$

The first integral is positive semi-definite, while the second integral is equal to $t\Delta h$. Therefore, if $t\Delta h \geq 0$, we choose the plus (lower) sign, while if $t\Delta h \leq 0$, we choose the minus (upper) sign. Then the second term is always positive, and thus

$$\int d\tau \mathcal{L}_b \geq t|\Delta h| \tag{12.166}$$

with equality for (assuming Δh is positive)

$$\frac{dx^i}{d\tau} - t g^{ij}\frac{dh}{dx^j} = 0. \tag{12.167}$$

This is exactly the equation of steepest descent, physically stating that the tangent vector to the curve is parallel to the gradient, up to reparametrization. Also it should be noted that this equation is not the same as the usual instanton equation which we have seen can be interpreted as ordinary, conservative, Newtonian cinematic motion of a particle in the reversed potential. Such a motion would never follow a path of steepest descent and stop at a lower value of the potential. Here the equation of steepest descent is first order in the "time" coordinate, and thus allows such motion. The solution to Equation (12.167) obviously exists, which then implies $S_E = \int d\tau \mathcal{L}_b = t|\Delta h|$. Then for the operator d_t whose matrix elements we want to compute, they will then be proportional to $e^{-t|\Delta h|}$. If we want to compute matrix elements of the Hamiltonian $dd^* + d^*d$ then we get two factors of S_E and hence the amplitude is proportional to $e^{-2t|\Delta h|}$.

The next step in the calculation is to compute the determinant of the fluctuations in Gaussian approximation about the instanton configuration. It is

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

in this stage that the calculation dramatically simplifies due to supersymmetry. The non-zero eigenvalues are all paired in bosonic and fermionic multiplets. The fermionic determinant is exactly cancelled by the bosonic square root of the determinant. The bosonic zero mode corresponding to Euclidean time translation invariance, which would normally give rise to a diverging factor of β, is also exactly cancelled by a corresponding fermionic zero mode which would normally give rise to a vanishing determinant. These zero modes can be explicitly obtained first for the bosonic case in the usual way, the bosonic zero mode corresponds to the Euclidean time derivative of the instanton. Then the fermionic zero mode is obtained by a supersymmetry transformation of the bosonic zero mode. To show the cancellation of the contribution of the zero modes requires some care, we refer the reader to the detailed calculation in [62]. Finally the amplitude is given by the factor

$$\langle b|d_t|a\rangle = e^{-t|\Delta h|}. \tag{12.168}$$

However, we still have not determined the sign of the amplitude; the functional integral always gives rise to an ambiguous sign due to the fermions. To determine the sign, we go back to the calculation of the amplitude in the usual WKB method of Schrödinger quantum mechanics. Here we know that the states $|a\rangle$ at the critical point A and $|b\rangle$ at the critical point B rapidly die off, away from their respective critical points. Any overlap is greatest along the paths that connect the two critical points that are the semi-classical solutions to the equations of motion, the paths that keep the Euclidean action stationary. These paths are the instantons, the paths of steepest descent or ascent between the critical points. Thus the behaviour of the states along the paths of steepest descent are enough to determine the sign of the matrix element $\langle b|d_t|a\rangle$. The quantum mechanical problem becomes effectively one-dimensional along the path of steepest descent, and we find that the state $|a\rangle$ drops off as e^{-th} along the instanton that ascends from A to B. It must ascend, as $|a\rangle$ was a p-form, hence A was a critical point of p negative directions while $|b\rangle$ was a $p+1$-form, hence B was a critical point of $p+1$ negative directions. If we descend from A we can only reach other critical points with fewer negative directions, we can never reach B.

To determine the sign, we start at $|b\rangle$ at B and the orientation of the space of negative directions at $|b\rangle$, which we called V_B. Calculating $|b\rangle$ along the path of steepest descent in the WKB approximation we find the wave function of the state $|b\rangle$ at A, but it is still a $p+1$-form. However, with the limiting direction of the tangent vector as we arrive at A, we can induce an orientation of V_A, the space of negative directions at A. Then the sign of the matrix element $\langle b|d_t|a\rangle$ is $+1$ if this induced orientation of V_A matches that furnished by the state $|a\rangle$, otherwise it is -1. This is exactly the construction of the sign $n(a,b)$ that was described in section 12.4.5, but appended with the explicit transport afforded by the WKB calculation of the wave function along the path of steepest descent.

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press

Hence the Witten coboundary operator is given by

$$\delta_W|a\rangle = \sum_b e^{-t(h(B)-h(A))} n(a,b)|b\rangle. \tag{12.169}$$

Since the path descends from B to A, the exponent has the right sign. This factor can be removed by rescaling the wave functions by

$$|a\rangle \to e^{th(A)}|a\rangle, \tag{12.170}$$

which corresponds to undoing the conjugation by e^{th} which transformed d to d_t. Hence the Witten coboundary operator is given by

$$\delta_W|a\rangle = \sum_b n(a,b)|b\rangle \tag{12.171}$$

and the notion that the set of critical points of a Morse function form a model of the cohomology of the manifold $\mathcal{M}$ is verified.

https://doi.org/10.1017/9781009291248.013 Published online by Cambridge University Press