Clay Mathematics Monographs
Volume 1

MIRROR SYMMETRY

Kentaro Hori
Sheldon Katz
Albrecht Klemm
Rahul Pandharipande
Richard Thomas
Cumrun Vafa
Ravi Vakil
Eric Zaslow

American Mathematical Society
Clay Mathematics Institute

2000 *Mathematics Subject Classification.* Primary 14J32; Secondary 14-02, 14N10, 14N35, 32G81, 32J81, 32Q25, 81T30.

For additional information and updates on this book, visit
www.ams.org/bookpages/cmim-1

Library of Congress Cataloging-in-Publication Data
Mirror symmetry / Kentaro Hori ... [et al.].
p. cm. — (Clay mathematics monographs, ISSN 1539-6061 ; v. 1)
Includes bibliographical references and index.
ISBN 0-8218-2955-6 (alk. paper)
1. Mirror symmetry. 2. Calabi-Yau manifolds. 3. Geometry, Enumerative. I. Hori, Kentaro. II. Series.

QC174.17.S9M5617 2003
530.14'3-dc21 2003052414

Published by the American Mathematical Society, Providence, RI,
for the Clay Mathematics Institute, Cambridge, MA.
Printed in the United States of America.

∞ The paper used in this book is acid-free and falls within the guidelines established to ensure permanence and durability.
Visit the AMS home page at http://www.ams.org/
Visit the Clay Mathematics Institute home page at http://www.claymath.org/

10 9 8 7 6 5 4 3 2 1 08 07 06 05 04 03

Contents

Preface

In the spring of 2000, the Clay Mathematics Institute (CMI) organized a school on Mirror Symmetry, held at Pine Manor College, Brookline, Massachusetts. The school was intensive, running for four weeks and including about 60 graduate students, selected from nominations by their advisors, and roughly equally divided between physics and mathematics. The lecturers were chosen based on their expertise in the subject as well as their ability to communicate with students. There were usually three lectures every weekday, with weekends reserved for excursions and relaxation, as well as time to catch up with a rapidly developing curriculum. The first two weeks of the school covered preliminary physics and mathematics. The third week was devoted to the proof of mirror symmetry. The last week introduced more advanced topics.

This book is a product of that month-long school. Notes were taken for some of the lectures by Amer Iqbal, Amalavoyal Chari and Chiu-Chu Melissa Liu and put into a rough draft. Other parts were added by the lecturers themselves. Part 1 of the book is the work of Eric Zaslow (with the contribution of Ch. 7 by Sheldon Katz). Part 2 was based on the lectures of Kentaro Hori and myself (most of it is Hori's). Part 3 was based on my own lectures. Part 4 is the work of Rahul Pandharipande and Ravi Vakil, based on lectures by Rahul Pandharipande. Part 5 involves various contributions by different authors. Chs. 31, 32 and 36 were based on my lectures. Ch. 33 was written by Sheldon Katz. Ch. 34 was written by Rahul Pandharipande and Ravi Vakil, based on lectures by Rahul Pandharipande. Ch. 35 was written by Albrecht Klemm. Ch. 37 was written by Eric Zaslow. Ch. 38 is based on the lectures by Richard Thomas. Finally Ch. 39 was written by Kentaro Hori.

Given that the authors were writing in different locations, and in the interest of a more convenient mechanism of communication among various authors, CMI set up an internet-accessible system where various authors could see what each one was writing and mutually correlate their contributions. The set-up was developed by Gordon Ritter and prove to be crucial for the completion of the book. Vida Salahi was the manager of the corresponding site and set the relevant deadlines for completion and delivery. She continued to provide tremendous assistance with manuscript preparation during the months following the school.

We have also had a gratifying abundance of secretarial assistance. In particular, Dayle Maynard and John Barrett ran the daily activities of the school, registering the incoming students, producing copies of lectures for the students, taking care of financial aspects of the school, arranging excursions, etc. They were greatly assisted by Barbara Drauschke at CMI.

We are especially grateful to Arthur Greenspoon and Edwin Beschler for their expert editing of the manuscript. They read the final draft carefully and made many constructive comments and suggestions. However, the authors would be responsible for any remaining errors. We solicit help in correcting possible mistakes we have made.

Alexander Retakh did the typesetting and Arthur Greenspoon made the index for the book. Their contribution was essential to producing this volume and is greatly appreciated.

We also wish to thank Sergei Gelfand of the AMS for his editorial guidance and David Ellwood for his supervision of the editorial process through all stages of the production of this volume.

It is my pleasure to say that this book is the outcome of the CMI's generous support of all aspects of this school. I sincerely thank CMI for this contribution to science and, in particular, Arthur Jaffe for his untiring efforts in enabling this school to take place.

Cumrun Vafa
Harvard University

Introduction

Since the 1980s, there has been an extremely rich interaction between mathematics and physics. Viewed against the backdrop of relations between these two fields throughout the history of science, that may not appear to be so surprising. For example, throughout most of their history the two subjects were not clearly distinguished. For much of the 1900s, however, physics and mathematics developed to a great extent independently and, except for relatively rare and not-so-deep interconnections, the two fields went their separate ways.

With the appreciation of the importance of Yang–Mills gauge theories in describing the physics of particle interactions, and with the appreciation of its importance in the mathematics of vector bundles, renewed interaction between the two fields began to take place. For example, the importance of instantons and monopoles came to be appreciated from both the physical and mathematical points of view. With the discovery of supersymmetry and its logical completion to superstring theory, a vast arena of interaction opened up between physics and mathematics and continues today at a very deep level for both fields.

Fundamental questions in one field often turn out to be fundamental questions in the other field as well. But even today mathematicians and physicists often find it difficult to discuss their work and interact with each other. The reason for this appears to be twofold. First, the languages used in the two fields are rather different. This problem is gradually being resolved as we recognize the need to become "bilingual." The second and more serious problem is that the established scientific methods in the two fields do not converge. Whereas mathematics places emphasis on rigorous foundations and the interplay of various structures, to a physicist the relevant aspects are physical clarity and physical interconnection of ideas, even if they come at the cost of some mathematical rigor. This can lead to friction between mathematicians and physicists. While mathematicians respect physicists for their intuition, they sometimes do not fully trust how those results were obtained and so they erect their own rigorous foundations as a substitute for the physical reasoning leading to those results. At the same time, physicists, who now appreciate the importance of modern mathematics as a powerful tool for theoretical physics, feel that attempts to build on a more rigorous foundation, while noble, will distract them from their real goal of understanding nature. Thus we are at a delicate point in the history of the interaction of these two fields: While both fields desperately need each other, the relationship seems at times to be a dysfunctional codependence rather than a happy marriage!

The aim of this book is to develop an aspect of this interplay known as "mirror symmetry" from **both** physical and mathematical perspectives, in order to further interaction between the two fields. With this goal in mind, almost half of the book includes introductory mathematics and physics material, while we try to emphasize the interconnection between the two areas. Unfortunately, however, the book also reflects the present status, namely, we find two distinct approaches to understanding mirror symmetry, without a clear connection between physical and mathematical methods of proof. Even the notion of what one means by "proof" of mirror symmetry differs between the two fields.

Mirror symmetry is an example of a general phenomenon known as duality, which occurs when two seemingly different physical systems are isomorphic in a non-trivial way. The non-triviality of this isomorphism involves the fact that quantum corrections must be taken into account. Mathematically, a good analogy is the Fourier transform, where local concepts such as products are equivalent to convolution products, requiring integration over the whole space. Thus it is difficult to understand such isomorphisms in the classical context. In particular, under such an isomorphism, certain complicated quantities involving quantum corrections in one system get mapped to simple classical questions in the other. Thus, finding such dualities leads to solving complicated physical questions in terms of simple ones in the dual theory. Precisely for this reason the discovery of duality symmetries has revolutionized our understanding of quantum theories and string theory.

It is fair to say that we do not have a deep understanding of the reason for the prevalence of duality symmetries in physics. Nor do we have a proof of why a duality should exist in any given case. Most of the arguments in favor of duality symmetries involve checking consequences and seeing that they are indeed satisfied in a non-trivial way. Because there have been so many non-trivial checks, we have no doubts about their validity, but that does not mean we have a deep understanding of the inner workings of duality symmetries. The only heuristic explanation of dualities we know of is the "scarcity of rich structures," and consistent quantum theories are indeed rather rich. So different ways of coming up with similar quantum systems end up being equivalent!

There is, however, one exception to this rule, mirror symmetry; for we have a reasonably clear picture of how it works. Moreover, a mathematical framework to rigorize many of the statements arising from the physics picture has also been constructed, and the subject is in a rather mature state of development. It is our hope that by elaborating aspects of this beautiful duality to both physicists and mathematicians, we can inspire further clarifications of this duality, which may also serve as a model for a deeper understanding of other dualities and interconnections between physics and mathematics.

A History of Mirror Symmetry

The history of the development of mirror symmetry is a very complicated one. Here we give a brief account of it, without any claim to completeness. The origin of the idea can be traced back to a simple observation of [**154**], [**223**] that string theory propagation on a target space that is a circle of radius R is equivalent to string propagation on a circle of radius $1/R$ (in some natural units). This has become known as T-duality. Upon the emergence of Calabi–Yau manifolds as interesting geometries for string propagation [**41**], a more intensive study of the corresponding string theories was initiated. It was soon appreciated that $\mathcal{N}=2$ supersymmetry on the worldsheet is a key organizing principle for the study of the corresponding string theories. It was noticed by [**71**] and [**173**] that given an $\mathcal{N}=2$ worldsheet theory, it is not possible to uniquely reconstruct a corresponding Calabi–Yau manifold. Instead there was a twofold ambiguity. In other words, it was seen that there could be pairs of Calabi–Yau manifolds that lead to the same underlying worldsheet theory, and it was conjectured that perhaps this was a general feature of all Calabi–Yau manifolds. Such pairs did not even have to have the same cohomology dimensions. In fact, the Hodge numbers $h^{p,q}$ for one of them was mapped to $h^{d-p,q}$ for the mirror, where d is the complex dimension of the Calabi–Yau manifold. Moreover, it was seen that the instanton-corrected cohomology ring (i.e., quantum cohomology ring) for one is related to a classical computation for the mirror. Phenomenological evidence for this conjecture was found in [**42**], where a search through a large class of Calabi–Yau threefolds showed a high degree of symmetry for the number of Calabi–Yaus with Euler numbers that differ by sign, as is predicted by the mirror conjecture. Non-trivial examples of mirror pairs were constructed in [**123**], using the relation between Calabi–Yau manifolds and Landau–Ginzburg models [**107**], [**189**], [**124**]. It was shown in [**45**] that one could use these mirror pairs to compute the instanton corrections for one Calabi–Yau manifold in terms of the variations of Hodge structure for the mirror. The instanton corrections involve certain questions of enumerative geometry; roughly speaking, one needs to know how many holomorphic maps exist from the two-sphere to the Calabi–Yau for any fixed choice of homology class for the two-cycle image.

The notion of topological strings was introduced in [**262**] where it abstracted from the full worldsheet theory only the holomorphic maps to the target. It was noted in [**245**] and [**264**] that mirror symmetry descends to a statement of the equivalence of two topological theories. It is this latter statement that is often taken to be the definition of the mirror conjecture in the mathematics literature. In [**16**] and [**17**] it was suggested that one could use toric geometry to propose a large class of mirror pairs. In [**265**] linear sigma models were introduced, which gave a simple description of a string propagating on a Calabi–Yau, for which toric geometry was rather natural. In [**267**] it was shown how to define topological strings on Riemann surfaces with boundaries and what data is needed to determine the boundary condition (the choice of the boundary condition is what we now call the choice of a D-brane and was first introduced in [**67**]). In [**24**] and [**25**], it was shown how one can use mirror symmetry to count holomorphic maps from higher genus curves to Calabi–Yau threefolds. In [**164**] a conjecture was made about mirror symmetry as a statement about the equivalence of the derived category and the Fukaya category. In [**163**] it was shown how one can use localization ideas to compute the "number" of rational curves directly. It was shown in [**108, 109**] and [**180, 181, 182, 183**] how one may refine this program to find a more effective method for computation of the number of rational curves. Moreover, it was shown that this agrees with the predictions of the number of rational curves based on mirror symmetry (this is what is now understood to be the "mathematical proof of mirror symmetry"). In [**234**] it was shown, based on how mirror symmetry acts on D0-branes, that Calabi–Yau mirror pairs are geometrically related: One is the moduli of some special Lagrangian submanifold (equipped with a flat bundle) of the other. In [**246**] the implications of mirror symmetry for topological strings in the context of branes was sketched. In [**114**] the integrality property of topological string amplitudes was discovered and connected to the physical question of counting of certain solitons. In [**135**] a proof of mirror symmetry was presented based on T-duality applied to the linear sigma model. Work on mirror symmetry continues with major developments in the context of topological strings on Riemann surfaces with boundaries, which is beyond the scope of the present book.

The Organization of this Book

This book is divided into five parts. Part 1 deals with mathematical preliminaries, including, in particular, a brief introduction to differential and algebraic geometry and topology, a review of Kähler and Calabi–Yau geometry, toric geometry and some fixed point theorems. Part 2 deals with physics preliminaries, including a brief definition of what a quantum field theory is, with emphasis on dimensions 0, 1, and 2 and the introduction of supersymmetry and localization and deformation invariance arguments for such systems. In addition, Part 2 deals with defining linear and non-linear sigma models and Landau–Ginzburg theories, renormalization group flows, topological field theories, D-branes and BPS solitons. Part 3 deals with a physics proof of mirror symmetry based on T-duality of linear sigma models. Part 4 deals with a mathematics proof of the mirror symmetry statement about the quantum cohomology ring. This part includes discussions of moduli spaces of curves and moduli spaces of stable maps to target spaces, their cohomology and the use of localization arguments for computation of the quantum cohomology rings. Even though the basic methods introduced in Parts 3 and 4 to prove mirror symmetry are rather different, they share the common feature of using circle actions. In Part 3, the circle action is dualized, whereas in Part 4 the same circle actions are used to localize the cohomology computations. Part 5 deals with advanced topics. In particular, topological strings at higher genera and the notion of holomorphic anomaly are discussed, as well as how one can carry out explicit computations at higher genera. In addition, integral invariants are formulated in the context of topological strings. Applications of mirror symmetry to questions involving QFTs that are geometrically engineered, as well as black hole physics, are discussed. Also discussed is a large N conjecture relating closed and open topological string amplitudes. Aspects of D-branes and their role in a deeper understanding of mirror symmetry are discussed, including the relevant categories in the mathematical setup as well as the relevance of special Lagrangian fibrations to a geometric understanding of mirror symmetry.

Throughout the book we have tried to present exercises that are useful in gaining a better understanding of the subject material, and we strongly encourage the reader to carry them out. Whenever feasible, we have tried to connect the various topics to each other, although it is clear that more work remains to be done to develop deeper connections among the various topics discussed – whose further development is, after all, one of the goals of this book.

There are a number of textbooks that nicely complement the topics covered here. In particular, quantum field theories are presented for a mathematical audience in [**68**]. An expository book on mirror symmetry, with emphasis on the mathematical side, is [**63**].

Part 1

Mathematical Preliminaries

CHAPTER 1

Differential Geometry

In this chapter we review the basics of differential geometry: manifolds, vector bundles, differential forms and integration, and submanifolds. Our goal is a quick understanding of the tools needed to formulate quantum field theories and sigma models on curved spaces. This material will be used throughout the book and is essential to constructing actions for quantum field theory in Part 2.

1.1. Introduction

Atlases of the Earth give coordinate charts for neighborhoods homeomorphic (even diffeomorphic) to open subsets of $\mathbb{R}^2$. One then glues the maps together to get a description of the whole manifold. This is done with "transition functions" (as in "see map on page 36"). Vector bundles are constructed similarly, except that at every point lives a vector space of fixed rank, so one needs not only glue the points together, but also their associated vector spaces. The transition functions, then, have values in isomorphisms of fixed-rank vector spaces. Differentiating a vector field, then, is a chart-dependent operation. In order to compare the vector space over one point to a neighboring point (to take a derivative), one must therefore have a way of connecting nearby vector spaces. Assigning to each direction an endomorphism representing the difference (from the identity) between "neighboring" vector spaces is the notion of a connection.

The notions of lengths and relative angles of vectors are provided by a position-dependent inner product, or "metric." This allows us to compute the sizes of vector fields and create actions.

Other notions involving vector spaces in linear algebra and high school vector calculus can be adapted to curved manifolds. While we will use coordinates to describe objects of interest, meaningful quantities will be independent of our choice of description.

1.2. Manifolds

As stated above, we describe a manifold by coordinate charts. Let $\{U_\alpha\}$ be an open covering of the topological space M. We endow M with the structure of an n-dimensional manifold with the following information. Let $\varphi_\alpha : U_\alpha \hookrightarrow \mathbb{R}^n$ be a coordinate chart (one may think of coordinates $x_\alpha = (x_\alpha)^i$, $i = 1, \ldots, n$ as representing the points themselves, i.e., their pre-images under φ_α). On $U_\alpha \cap U_\beta$, we can relate coordinates (x_α) to coordinates (x_β) by $x_\alpha = \varphi_\alpha \circ \varphi_\beta^{-1}(x_\beta)$.

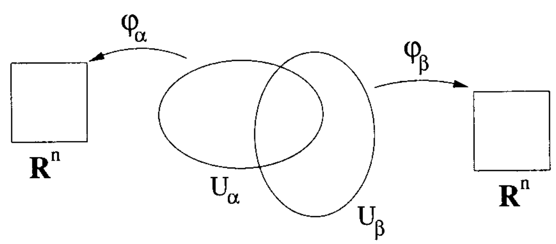


FIGURE 1. Two open sets U_α and U_β, with coordinate charts φ_α and φ_β

The map $g_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}$ is a transition function. Note that $g_{\alpha\beta} = g_{\beta\alpha}^{-1}$ and $g_{\alpha\beta} g_{\beta\gamma} g_{\gamma\alpha} = 1$. As an alternative to this structure, we could form a manifold based solely on the data of patches and transition functions satisfying the above relations. A manifold is called "differentiable" if its transition functions are differentiable, and "smooth" if the transition functions are smooth (C^∞). If $n = 2k$ and one can (and does) choose $\varphi_\alpha : U_\alpha \hookrightarrow \mathbb{C}^k$ with holomorphic transition functions, the manifold is called "complex." Note that this extra structure is restrictive. Two complex manifolds may be diffeomorphic as real manifolds (meaning there are invertible, onto, differentiable maps between the two), but there may be no complex analytic mapping between them (we then say they have different complex structures). Likewise, two homeomorphic manifolds may have different structures as differentiable manifolds. Differentiablility depends on the coordinate chart maps φ_α.

EXAMPLE 1.2.1 (S^2). *On the two-sphere we can choose coordinates (θ, ϕ), but these are "singular" at the poles (the azimuthal angle ϕ is not well defined).*

Instead we consider two patches. Let U_s be $S^2 \setminus \{n\}$ and U_n be $S^2 \setminus \{s\}$, where n and s are the north and south poles.

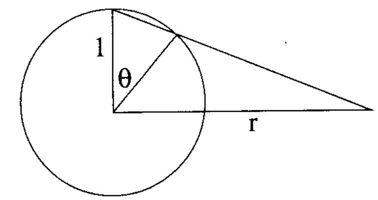


FIGURE 2. The stereographic projection of the sphere. Note that projecting from the south pole can be effected by sending $\theta \to \pi - \theta$.

Projecting as shown in the figure gives a map from U_s to $\mathbb{R}^2$. In terms of θ and ϕ (which we keep here only for convenience),

$$x = \cot(\theta/2)\cos\phi, \qquad y = \cot(\theta/2)\sin\phi.$$

We can also define a complex coordinate $z = \cot(\theta/2)e^{i\phi}$.

On U_n, we can project onto $\mathbb{R}^2$ from the bottom. In order to preserve the "handedness" of the coordinates, it is convenient to view $\mathbb{R}^2$ "from below." The maps may be easily obtained by replacing $\theta \rightsquigarrow \pi - \theta$ and $\phi \rightsquigarrow -\phi$, so that in this patch the coordinates are

$$\widetilde{x} = \tan(\theta/2)\cos\phi = \frac{x}{x^2+y^2}, \qquad \widetilde{y} = \tan(\theta/2)\sin(-\phi) = \frac{-y}{x^2+y^2}.$$

Note that $\widetilde{z} = \tan(\theta/2)e^{-i\phi} = 1/z$.

On $U_s \cap U_n \subset U_s$, coordinatized by $\{(x,y) \neq (0,0)\}$, we have

$$g_{ns} : (x,y) \mapsto (x/(x^2+y^2), -y/(x^2+y^2)).$$

In complex coordinates, $g_{ns} : z \mapsto 1/z$, and we see that the two-sphere can be given the structure of a one-dimensional complex manifold (Riemann surface). Note that the dimension as a complex manifold is half the dimension as a real manifold.

1.3. Vector Bundles

As mentioned in the introduction, vector bundles are constructed similarly, only now every point carries an additional structure of a vector space ("fiber") over it. Clearly, by retaining the information of the point but forgetting the information of the vector space, we get a map to the underlying manifold. In this section, we will focus on smooth vector bundles.[1]

From the description above, it is clear that the simplest vector bundle, E, will be a product space $E = M \times V$, where M is a manifold and V is an r-dimensional vector space. E is said to be a rank r vector bundle. E is equipped with the map $\pi : E \to M$, namely $\pi((m, v)) = m$. Such a vector bundle is called "trivial."

Locally, all vector bundles are trivial and look like products. So a rank r vector bundle E is a smooth manifold with a map $\pi : E \to M$ to a base manifold, M, such that every point $x \in M$ has a neighborhood $U_x \ni x$ with $\pi^{-1}(U_x) \cong U_x \times \mathbb{R}^r$.

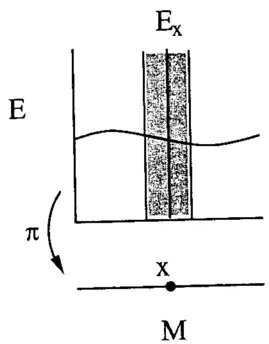


FIGURE 3. A vector bundle, E, with its map π to a base manifold, M. $E_x = \pi^{-1}(x)$ is the fiber over $x \in M$. The shaded region represents $\pi^{-1}(U_x)$, where $U_x \ni x$. The curvy line represents a section of E

From now on we assume we have a cover $\{U_\alpha\}$ of M along each chart of which E is locally trivial. The choice of isomorphism $\rho_\alpha : \pi^{-1}(U_\alpha) \xrightarrow{\sim} U_\alpha \times \mathbb{R}^r$ is analogous to a choice of coordinates, so geometric structures will undergo transformations when different "local trivializations" are chosen. Writing $\rho_\alpha = (\pi, \psi_\alpha)$, we have $\psi_\alpha : \pi^{-1}(U_\alpha) \to \mathbb{R}^r$.

By analogy with manifolds, we glue together vectors using $s_{\beta\alpha} = \psi_\beta \circ \psi_\alpha^{-1}$, so that $(m, v_\alpha) \in U_\alpha \times \mathbb{R}^r$ will be identified with $(m, \psi_\beta \circ \psi_\alpha^{-1}(v_\alpha))$. Of

[1]More generally, fiber bundles have fibers a fixed topological space, and principal bundles have Lie group-valued fiber spaces. We focus on vector bundles here.

course, we insist that $\psi_\beta \circ \psi_\alpha^{-1}$ be a linear map on the $\mathbb{R}^r$ fibers, and

$$\begin{aligned} s_{\alpha\beta} &= s_{\beta\alpha}^{-1}, \\ s_{\alpha\beta} s_{\beta\gamma} s_{\gamma\alpha} &= \mathbf{1}. \end{aligned}$$

Conversely, these data can be used to construct the vector bundle by gluing:

$$E = \coprod_\alpha U_\alpha \times \mathbb{R}^r / \sim, \qquad (x, v_\alpha) \sim \big(x, s_{\alpha\beta}(x)(v_\beta)\big).$$

A "section" v_α over U_α is a map $v_\alpha : U_\alpha \to \mathbb{R}^r_\alpha$ (think of a vector-valued function $(v_\alpha^1, v_\alpha^2, ..., v_\alpha^r)$). Thus there is a specific point on the fiber for each point on the base. Two sections v_α and v_β over U_α and U_β make up a section over $U_\alpha \cup U_\beta$ if they coincide along the intersection $U_\alpha \cap U_\beta$, i.e., $v_\beta = s_{\beta\alpha}(v_\alpha)$. A "global section" is a map $\nu : M \to E$ such that $\pi \circ \nu$ is the identity on M. One can check that this is equivalent to ν being a section on $\bigcup_\alpha U_\alpha$. We will denote the space of sections of E over $U \subset M$ by $\Gamma(U, E)$. $\Gamma(E)$ will denote global sections. Note that sections can be multiplied by functions: the value of the section over a point gets multiplied by the value of the function at that point. Put differently, $\Gamma(U, E)$ is a $C^\infty(U)$ module.

EXAMPLE 1.3.1. *A section of a trivial bundle $M \times V$ is a V-valued function, $f : M \to V$.*

A complex vector bundle is a locally trivial family of complex vector spaces, and again its rank is half its rank as a real vector bundle. Such a bundle over a complex manifold is called "holomorphic" if all the transition functions are holomorphic.

1.3.1. The Tangent Bundle. The classic vector bundle is the tangent bundle of a manifold. If the manifold is a surface embedded in $\mathbb{R}^3$ this is easy to visualize by thinking of the tangent plane at a point as its associated vector space (though intersections of different tangent planes should be disregarded). More formally, a vector field v is a differential operator on the space of functions via the directional derivative: $v(f) = D_v f$ (or $\nabla f \cdot v$ in calculus notation). In coordinates x^a, the obvious differential operators are $\frac{\partial}{\partial x^a}$, and these provide a local trivialization of the tangent bundle. Namely, in this coordinate patch, we may express any vector field (differential) as

$v = v^a \frac{\partial}{\partial x^a}$.[2] Clearly $v(fg) = gv(f) + fv(g)$. Between coordinate patches x^a and $\widetilde{x}^k(x)$ the chain rule provides transition functions s:

$$\frac{\partial}{\partial x^a} = \frac{\partial \widetilde{x}^k}{\partial x^a} \frac{\partial}{\partial \widetilde{x}^k},$$

so

$$s_a{}^k = \frac{\partial \widetilde{x}^k}{\partial x^a},$$

where we have written $s : \widetilde{\mathbb{R}}^r \to \mathbb{R}^r$ in matrix notation.[3] A global section is a global vector field. Note that every vector bundle has the zero section as a global section. The existence of non-vanishing sections is non-trivial, especially if we are working in the holomorphic category.

EXAMPLE 1.3.2. *We recall from Example 1.2.1 that the two-sphere can be considered as a one-dimensional complex manifold. Let us look for global, holomorphic vector fields. By "holomorphic" we mean a vector field* $v = v^z \frac{\partial}{\partial z}$, *with* v^z *holomorphic. It lives in the holomorphic piece of* $TM \otimes \mathbb{C} = T_{\text{hol}} \oplus T_{\text{anti-hol}}$, *where* T_{hol} *is generated by* $\frac{\partial}{\partial z}$ *and* $T_{\text{anti-hol}}$ *by* $\frac{\partial}{\partial \bar{z}}$. *Moving to the patch coordinatized by* $w = 1/z$, *we see that* $v^w = v^z \frac{\partial w}{\partial z} = -v^z/z^2$, *and since this must be non-singular at* $w = 0$ *(i.e.,* $z \to \infty$*),* v^z *must be at most quadratic in* z *(note then that* v^w *is also quadratic in* w*). Therefore there is a three-dimensional space of global, holomorphic vector fields on the complex sphere:* $v = a + bz + cz^2$, *with* a, b, c *constant complex numbers.*

EXAMPLE 1.3.3. *The total space of the Möbius bundle is* $[0,1] \times \mathbb{R}/\sim$, *where* $(0,r) \sim (1,-r)$. *It is a one-dimensional vector bundle (line bundle) over the circle* S^1. *Note that* $x \mapsto \{x, 0\}$ *is the zero section, its image isomorphic to* S^1. *This bundle has no nowhere-vanishing sections — an issue related to the non-orientability of the Möbius strip.*

EXAMPLE 1.3.4. *Consider a path* $\gamma : \mathbb{R} \to M$. *Choose* t *a coordinate on* $\mathbb{R}$. *Then the vector field* $\partial_t \equiv \frac{\partial}{\partial t}$ *trivializes the tangent bundle of* $\mathbb{R}$, *since every vector field has the form* $f\partial_t$, *where* f *is a function. Along the image* $\gamma(\mathbb{R})$ *the coordinates (locally) depend on* t, *so a function* f *along the image*

[2] Here we sum over repeated indices, a convention we use throughout this book. Note, however, that when an index is the label of a coordinate chart (such as α, β) then there is no summation.

[3] Note that choosing active or passive representation of the linear transformation s will affect the indices. We often denote vectors by their components, for example. In any case, consistency is key.

can be thought of as the function $f \circ \gamma : \mathbb{R} \to \mathbb{R}$. *In particular,* $\frac{df}{dt}$ *makes sense. Therefore we can push the vector* ∂_t *forward with* γ *to create the vector field* $\gamma_* \partial_t$, *a vector on the image defined by* $\gamma_*\partial_t(f) = \frac{df}{dt}$. *In coordinates* y^α *on* M, *the map* γ *looks like* $t \mapsto x^\alpha(t)$, *and the chain rule gives*

$$\gamma_* \partial_t = \frac{\partial x^\alpha}{\partial t} \frac{\partial}{\partial x^\alpha}.$$

The vector $\gamma_* \partial_t$ *is often written* $\dot{\gamma}$.

The example above can be generalized. Instead of a path, we can have any map $\varphi : N \to M$ of N to M. Locally, the map can be written as $y^\alpha(x^k)$, where y^α and x^k are coordinates on M and N, respectively. This allows us to define the push-forward

$$\varphi_* \frac{\partial}{\partial x^k} = \frac{\partial y^\alpha}{\partial x^k} \frac{\partial}{\partial y^\alpha}.$$

Note that, in general, one cannot pull vectors back.

1.3.2. The Cotangent Bundle. Every bundle E has a dual E^* whose vector space fibers are the dual vector spaces to the fibers of E, so if $E_x = \pi^{-1}(x)$, then $E_x^* = \mathrm{Hom}(E_x, \mathbb{R}) =$ *linear maps from* E *to* $\mathbb{R}$ is a vector space of the same dimension.

Dual to the tangent bundle TM is the "cotangent bundle" T^*M, and it, too, has a natural trivialization in a coordinate patch. One defines the basis dx^a to be dual to the basis $\frac{\partial}{\partial x^b}$, so that the natural pairing is

$$dx^a\left(\frac{\partial}{\partial x^b}\right) = \langle dx^a, \frac{\partial}{\partial x^b} \rangle = \delta^a{}_b.$$

Here we have sloppily, though conventionally, used the same symbol $\langle\, ,\, \rangle$ for the natural pairing as for the inner product. An arbitary cotangent vector (also called a "one-form") θ can be written in this basis as $\theta = \theta_a dx^a$.

Now the transition functions for the tangent bundle determine those of the cotangent bundle, both a consequence of the chain rule. If in a new coordinate basis we rewrite $\frac{\partial}{\partial x^a}$ as $\frac{\partial \widetilde{x}^k}{\partial x^a} \frac{\partial}{\partial \widetilde{x}^k} = s_a{}^k \frac{\partial}{\partial \widetilde{x}^k}$ in the relation $\langle dx^b, \frac{\partial}{\partial x^a} \rangle = \delta^b{}_a$, and rewrite $dx^b = \Lambda^b{}_l d\widetilde{x}^l$, then using $\langle d\widetilde{x}^l, \frac{\partial}{\partial \widetilde{x}^k} \rangle = \delta^l{}_k$, we must have $\Lambda^b{}_l s^k_a \delta^l{}_k = \delta^b{}_a$. From this we see $\Lambda = (s^T)^{-1}$, which is of course how the elements of the dual space should transform.

Note that we could have used an arbitrary position-dependent set of basis vectors to trivialize the tangent and cotangent bundles, but the coordinate vectors are particularly natural.

Tangents push forward, and cotangents pull back. So if $\varphi : M \to N$, dy^k is a local basis, and $\theta = \theta_k dy^k$ is a cotangent section of T^*N, we define the pull-back $\varphi^*\theta$ to be a cotangent section of T^*M. We define a covector by its action on a vector, v, so define $\langle \varphi^*\theta, v\rangle = \langle \theta, \varphi_* v\rangle$. Let us set $v = \frac{\partial}{\partial x^a}$ so $\langle \varphi^*\theta, v\rangle$ equals the component $(\varphi^*\theta)_a$. Now the push-forward equation gives

$$(\varphi^*\theta)_a = \frac{\partial y^k}{\partial x^a}\theta_k.$$

1.3.3. More Bundles. In the last section, we used dual vector spaces to construct a new bundle, and its transition functions followed naturally from the original ones through linear algebra. Similarly, we get a whole host of bundles using duals and tensor products. For example, starting with E we can form the vector bundle $E \oplus E$, whose fiber at x is $E_x \oplus E_x$. If $s_{\alpha\beta}$ is the transition function for E, then $s_{\alpha\beta} \oplus s_{\alpha\beta}$ is the transition function for $E \oplus E$.

Given two vector bundles E and F over M, we can define $E \oplus F$, $\mathrm{Hom}(E, F)$, E^*, $E \otimes F$, etc.[4] Note that $E \otimes E$ decomposes as $(E \otimes_s E) \oplus (E \otimes_a E)$, where s and a indicate symmetric and anti-symmetric combinations. Recall that if V is a vector space, then $\Lambda^2 V$ or $V \wedge V$ or $V \otimes_a V$ is formed by the quotient $V \otimes V/I$ where I is the subspace generated by $v_i \otimes v_j + v_j \otimes v_i$. The equivalence class $[v_i \otimes v_j]$ is usually written $v_i \wedge v_j$, and equals $-v_j \wedge v_i$, as can easily be checked. Thus we write $E \otimes_a E$ as $\Lambda^2 E$. $\Lambda^p E$ can be defined similarly. If E and F are two bundles over M, then a map $f : E \to F$ is a bundle map if it is a map of the total spaces of the bundles, linear on the fibers, and commutes with projections. In such a case we can define the bundle $\mathrm{Ker}(f) \subset E$ and $\mathrm{Coker}(f) = F/\mathrm{Im}(f)$ whose fibers have the natural linear algebra interpretation.

The bundles $\Lambda^p T^*M$ are particularly important and can be thought of as totally anti-symmetric p-multi-linear maps (p-tensors) on tangent vectors. If $\dim M = n$, $\Lambda^p T^*M$ is a rank $\binom{n}{p}$ bundle of anti-symmetric p-tensors, or "p-forms." The sections of $\Lambda^p T^*M$ are often written as $\Omega^p(M)$. Note that $\Lambda^0 V = \mathbb{R}$ for any vector space V, so $\Omega^0(M)$ are sections of the trivial line bundle, i.e., functions.

[4] To form the transition functions for $\mathrm{Hom}(E, F)$, simply use the relation for finite-dimensional vector spaces $\mathrm{Hom}(A, B) = A^* \otimes B$.

EXAMPLE 1.3.5. *From any function f we can form the one-form differential, $df = \frac{\partial f}{\partial x^a}dx^a$, which one checks is independent of coordinates. More invariantly, the value of df on a vector $v = v^a\partial_a$ is $\langle df, v\rangle = v^a\partial_a f = D_v f$, the directional derivative. So the directional derivative provides a map*

$$d : \Omega^0 \to \Omega^1,$$

where we have suppressed the M. It appears that sections of the bundles can be related in a natural way. We will return to this idea later in this chapter.

A "metric" (more in the next section) is a position-dependent inner product on tangent vectors. That is, it is a symmetric, bilinear map from pairs of vector fields to functions. From the discussion before the example, we learn that g is a global section of $T^*M \otimes_s T^*M$. Therefore it makes sense to express g in a coordinate patch as $g = g_{ab}dx^a \otimes dx^b$; so g_{ab} is symmetric under $a \leftrightarrow b$.

A "principal bundle" is entirely analogous to a vector bundle, where instead of "vector space" we have "Lie group," and transition functions are now translations in the group. Given a representation of a group, we can glue together locally trivial pieces of a vector bundle via the representation of the transition functions and create the "vector bundle associated to the representation." This is important in gauge theories. However, since particles are associated to vector bundles defined by representations as just discussed, we will focus on vector bundles exclusively.

Another important way to construct bundles is via "pull-back." If $f : M \to N$ is a map of manifolds and E is a vector bundle over N, then the pull-back bundle f^*E is defined by saying that the fiber at $p \in M$ is equal to the fiber of E at $f(p)$, that is, $f^*E|_p = E|_{f(p)}$. In terms of transition functions, the (s_E) pull back to transition functions $(s_{f^*E}) = (s_E \circ f)$. As a trivial example, a vector space V can be considered to be a vector bundle over a point $*$. Any manifold induces the map $f : M \to *$, and the pull-back is trivial: $f^*V = M \times V$. If E is a bundle on N and $f : * \to N$, then $f^*E = E|_{f(*)}$. If $f : M \hookrightarrow N$ is a submanifold, then f^*E is the restriction of E to M.

1.4. Metrics, Connections, Curvature

The three subjects of this section are the main constructions in differential geometry.

1.4.1. Metrics on Manifolds. On a vector space, an inner product tells about the sizes of vectors and the angles between them. On a manifold, the tangent vector spaces (fibers of the tangent bundle) can vary (think of the different tangent planes on a sphere in three-space) — hence so does the inner product. Such an inner product is known as a "metric," g, and provides the notion of measurement inherent in the word *geometry.* So if v and w are two vectors at x, then $g(v, w)$ is a real number. If $v(x)$ and $w(x)$ are vector fields, then g and $\langle v, w\rangle$ are x-dependent. Of course, we require g to be bilinear in the fibers and symmetric, so

$$g(v,w) = g(w,v), \qquad g(\lambda v, w) = \lambda g(v,w) = g(v, \lambda w).$$

In a coordinate patch x^a, we can write $v = v^a \frac{\partial}{\partial x^a}$, so

$$g(v,w) = g(v^a \frac{\partial}{\partial x^a}, w^b \frac{\partial}{\partial x^b}) = v^a w^b g(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}).$$

We define

$$g_{ab} \equiv g(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}),$$

and we see that, in a patch, g is defined by the matrix component functions $g_{ab}(x)$, and $\langle v, w\rangle = v^a w^b g_{ab}$. A manifold with a positive-definite metric ($g_{ab}(x)$ is a positive-definite matrix for all x) is called a "Riemannian manifold."

EXAMPLE 1.4.1. *What is the round metric on a sphere of radius r in terms of the coordinates (θ, ϕ)? Since $\frac{\partial}{\partial\theta}$ represents a vector in the latitudinal θ direction, some trigonometry shows that this is perpendicular to the longitudinal ϕ direction and should be assigned a length-squared equal to r^2. Analogously, the length-squared of ϕ is $r^2 \sin^2\theta$. So*

$$g_{\theta\theta} = r^2, \quad g_{\theta\phi} = g_{\phi\theta} = 0, \quad g_{\phi\phi} = r^2 \sin^2\theta.$$

The independence of ϕ is an indication of the azimuthal symmetry of the round metric; indeed, $\phi \to \phi + const$ is an "isometry." Note that other rotational isometries are not manifest in these coordinates.

EXERCISE 1.4.1. *Using the chain rule (equivalently, transition functions) to rewrite $\frac{\partial}{\partial\theta}$ and $\frac{\partial}{\partial\phi}$ in terms of the real and imaginary parts x and y of the complex coordinate $z = \cot(\theta/2)e^{i\phi}$, show that the metric takes the form $(4/(|z|^2+1)^2)\,[dx\otimes dx + dy \otimes dy]$. We can write this metric as the symmetric part of $(4/(|z|^2+1)^2) dz \otimes d\bar{z}$, where $dz = dx + idy$, etc. We will have more to say about the anti-symmetric part in future chapters.*

A metric is an inner product on the tangent bundle. If $v^a\partial_a$ and $w^b\partial_b$ are two vectors v and w, their inner product is $g(v,w) = g(v^a\partial_a, w^b\partial_b) = v^a w^b g_{ab}$. It is convenient to define $w_a \equiv w^b g_{ab}$, namely we "lower the index by contracting with the metric." Then the cotangent vector or one-form, $w_a dx^a$, has a natural pairing with v equal to the inner product of v and w. In short, the metric provides an isomorphism between the tangent and the cotangent bundles, exactly as an inner product defines an isomorphism between V and V^*.

The inner product of any vector space can be extended to arbitrary tensor products, wedge (or anti-symmetric) products, and dual spaces. The metric on the dual space is the inverse metric (this then respects the inner products between two vectors and their corresponding one-forms). If θ and ψ are two one-forms, their point-wise inner product is $g(\theta_a dx^a, \psi_b dx^b) = \theta_a\psi_b g^{ab}$, where we have paired the inverse matrix to g_{ab} with g^{ab} (i.e., $g^{ac}g_{cb} = \delta^a{}_b$). Note $\theta_a = \langle\theta, \partial_a\rangle$.

On arbitrary tensor products of vectors or forms of the same degree, we obtain the inner product by using the metric to raise indices, then contracting.

1.4.2. Metrics and Connections on Bundles. The notion of a metric makes sense for any vector bundle. Thus, given two sections r, s of E, we can ask for the inner product $h(r,s)$ as a function on the base. In a local trivialization, one specifies a "frame" of basis vectors e_a, $a = 1, ..., \mathrm{rank}(E)$. In terms of this basis, the metric is given by components

$$h_{ab}(x) = h(\mathrm{e}_a, \mathrm{e}_b).$$

Now let us try to differentiate vectors. Taking a hands-on approach, it is tempting to try to define the derivative of a vector v at a point x as a limit of

$$\text{“}\frac{v(x+\epsilon) - v(x)}{\epsilon}\text{”}.$$

However, this expression makes no sense! First of all, $+\epsilon$ makes no sense on a manifold. Instead, we shall have to specify a vector direction along which to compare nearby values of the vector. Let us choose to look in the ith direction, and denote the point whose ith coordinate has been advanced by ϵ as $x + \epsilon\partial_i$. Secondly, subtraction of vectors living in different spaces makes no sense either. We will need a way to relate or *connect* the vector space at

$x + \epsilon\partial_i$ to that at x. That is, we need an i-dependent automorphism. Since ϵ is small, we require our automorphism to be close to the identity (in any frame chosen to describe the vector spaces), so we write it as $1 + \epsilon A_i$, and it will be invertible for A_i an arbitrary endomorphism. Note the i-dependence. Differentiation, then, requires a direction-dependent endomorphism of tangent vectors — i.e., an endomorphism-valued one-form. Such a form is called a "connection."

Now let us try to differentiate in the ith direction. We want to write

$$D_i v = \frac{(1 + \epsilon A_i)(v(x + \epsilon\partial_i)) - v(x)}{\epsilon}.$$

Let us write v as $v^a\partial_a$ and expand (to linear order) the components v^a of the shifted argument by Taylor expansion. We get $v^a(x + \epsilon\partial_i) = v^a(x) + \epsilon\partial_i v^a(x)$. Thus, keeping the ath component of the vector and writing the endomorphism A_i as a matrix,

$$(D_i v)^a = \partial_i v^a + (A_i)^a{}_b v^b.$$

Recapping, given a direction, D maps vectors to vectors: $v \mapsto D_i v$. More generally, the vector w sends $v \mapsto D_w v = w^i D_i v = \langle Dv, w\rangle$. In the last expression, we have defined the vector-valued one-form $Dv = (D_i v)dx^i$. Now we can write the shorthand formula $Dv = (d + A)v$, or $D = d + A$.

The same procedure holds *mutatis mutandis* for arbitrary vector bundles (nothing special about tangent vectors). Given an $\mathrm{End}(E)$-valued one-form A (a "connection") and a direction, we compare values of a section s at nearby points and find the derivative. Then $D = d + A$, $Ds = (D_i s)dx^i$, and in a frame e_a, $D_i s = \left[\partial_i s^a + (A_i)^a{}_b s^b\right] \mathrm{e}_a = \langle Ds, \partial_i\rangle$. Thus, Ds is a one-form with values in E, or $D : \Gamma(E) \to \Omega^1 \otimes \Gamma(E)$. Note that, by our definition, if f is a function and s a section, $D(fs) = (df) \otimes s + f \cdot Ds$. A connection can also be defined as any map of sections $\Gamma(E) \to \Omega^1 \otimes \Gamma(E)$ obeying this Leibnitz rule.

A vector field/section s is called "covariantly constant" if $Ds = 0$, meaning that its values in nearby fibers are considered the same under the automorphisms defined by A.

If $D_{\dot\gamma} s = 0$ for all tangent vectors $\dot\gamma$ along a path γ, then s is said to be "parallel translated" along γ. Since parallel translation is an ordinary differential equation, all vectors can be parallel translated along smooth paths.

The curvature measures the non-commutativity of parallel translation along different paths, as we shall see.

1.4.3. The Levi–Civita Connection. The tangent bundle TM of a Riemannian manifold M has a natural connection denoted $\nabla : \Gamma(TM) \to \Omega^1 \otimes \Gamma(TM)$, which we will define shortly. This connection can be extended to the cotangent bundle or arbitrary tensor bundles. Given a metric, we define the connection ∇ with the properties that it is "torsion-free," i.e., $\nabla_X Y - \nabla_Y X = [X, Y]$ for all vector fields X and Y and further $\nabla g = 0$, where g is the metric considered as a section of $\mathrm{Sym}^2 T^*M$. To find this connection, let us work in local coordinates x^i with $\partial_i \equiv \frac{\partial}{\partial x^i}$, $i = 1, \ldots, n$, as a basis for tangent vectors. We write $\nabla_i X$ for $\langle dx^i, \nabla X\rangle$, the ith covector component of ∇X. Define Γ by $\nabla_{\partial_i}\partial_j = \Gamma^k{}_{ij}\partial_k$. Then the torsion-free condition says $\Gamma^k{}_{ij} = \Gamma^k{}_{ji}$. Let us denote $\langle X, Y\rangle = g(X, Y)$.

EXERCISE 1.4.2. *Start with* $\partial_i g_{jk} = \partial_i\langle\partial_j, \partial_k\rangle = \langle\nabla_{\partial_i}\partial_j, \partial_k\rangle + \langle\partial_j, \nabla_{\partial_i}\partial_k\rangle = \langle\Gamma^m{}_{ij}\partial_m, \partial_k\rangle + \langle\partial_j, \Gamma^m{}_{ik}\partial_m\rangle = \Gamma^m_{ij} g_{mk} + \Gamma^m{}_{ik} g_{jm}$. *Now add the equation with* $i \leftrightarrow j$ *and subtract the equation with* $i \leftrightarrow k$. *Using the torsion-free condition* $\Gamma^k{}_{ij} = \Gamma^k{}_{ji}$, *show that*

$$\Gamma^i{}_{jk} = \frac{1}{2} g^{im}(\partial_j g_{mk} + \partial_k g_{jm} - \partial_m g_{jk}).$$

Using the result from this exercise, we define $\nabla_j X = (\partial_j X^i)\frac{\partial}{\partial x^i} + \Gamma^i{}_{jk} X^k \frac{\partial}{\partial x^i}$.

EXERCISE 1.4.3. *For practice in pulling back metrics and using the Levi–Civita connection, it is instructive to derive the geodesic equations. Consider a curve* $\gamma : \mathbb{R} \to M$, *where* M *is a Riemannian manifold with its Levi–Civita connection,* ∇. γ *is called a geodesic if* $\nabla_{\dot\gamma}\dot\gamma = 0$. *This provides a notion of straightness. Prove that this equation, with components* $\gamma^i(t)$, *yields the "geodesic equation"*

$$\frac{d^2\gamma^i}{dt^2} + \Gamma^i{}_{jk}\frac{d\gamma^j}{dt}\frac{d\gamma^k}{dt} = 0.$$

In a flat metric with $\Gamma \equiv 0$, *we recover the usual notion of straight lines.*

EXERCISE 1.4.4. *Consider the metric* $g = \frac{1}{y^2}(dx \otimes dx + dy \otimes dy)$ *on the upper half-plane* $y > 0$. *Prove that all geodesics lie on circles centered on the* x*-axis (or are vertical lines).* **Hints:** *First show that only* $\Gamma^x{}_{xy} = \Gamma^y{}_{yy} = -\Gamma^y{}_{xx} = -1/y$ *are nonzero. Now write down the geodesic equations explicitly and re-express all* t*-derivatives of the path* $x(t), y(t)$ *in terms of*

$y' = dy/dx$ and y''. Show that the geodesic equation implies $y'' = -[(y')^2 + 1]/y$, which is solved by curves along $(x-a)^2 + y^2 = R^2$.

This metric has constant scalar curvature. By excising circular geodesics and a few identifying points, one can construct constant scalar curvature metrics on regions in the shape of "pairs of pants." Sewing these "pants" together along like seams leads to the constant curvature metrics on Riemann surfaces. It is not too hard to see that there are $6g-6$ real parameters to choose how to do the sewing for a Riemann surface of genus $g \geq 2$. These parameters describe the moduli space of Riemann surfaces, as we will see in future chapters. More generally, any Riemann surface can be obtained as the quotient of the upper half-plane by a discrete group of isometries.

One can also map this metric onto the unit disc by choosing coordinates $z = -i\frac{w-i}{w+i}$, where $w = x+iy$. Then $g = (4/(|z|^2-1)^2)dz \otimes_s d\bar{z}$.

1.4.4. Curvature. Of course, there is a lot to say about curvature. When the curvature is nonzero, lines are no longer "straight," triangles no longer have angles summing to π, etc. We won't have time to explore all the different meanings of curvature: for example, in general relativity, curvature manifests itself as "tidal" forces between freely falling massless particles. All of these deviations from "flatness" are a consequence of the fact that on a curved space, if you parallel translate a vector around a loop, it comes back shifted. For example, on a sphere, try always pointing south while walking along a path which goes from the north pole straight down to the equator, then a quarter way around the equator, then straight back up to the north pole (see Fig. 4). Your arm will come back rotated by $\pi/2$.

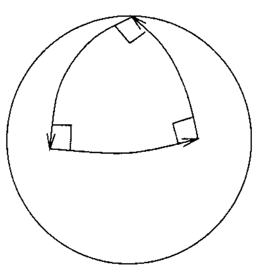


FIGURE 4. Holonomy is encountered upon parallel transport of a vector around a closed loop.

This process can be measured infinitesimally by associating an infinitesimal rotation (i.e., an endomorphism of the tangent space) to an infinitesimal loop (i.e., one defined from two vectors by a parallelogram). The curvature tensor is then an endomorphism-valued two-form,[5] which gives the infinitesimal rotation associated to any pair of directions. By the same reasoning, using parallel translation we can define a curvature associated to any vector bundle equipped with a connection.

Note that $\nabla_{\partial_i}\partial_k$ represents the infinitesimal difference between the vector field ∂_k and its parallel translate in the ∂_i direction. Therefore,[6] $(\nabla_{\partial_j}\nabla_{\partial_i} - \nabla_{\partial_i}\nabla_{\partial_j})\partial_k$ represents the difference in the closed loop formed by travel around a small i-j parallelogram. Generalizing from ∂_i and ∂_j to arbitrary vectors, we define

$$R(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]},$$

which maps $\Gamma(TM) \to \Gamma(TM)$. (If X and Y are coordinate vectors, then $[X,Y] = 0$ and the last term can be ignored. Here the first commutator really means the difference $\nabla_X\nabla_Y - \nabla_Y\nabla_X$.) This definition makes sense for any vector bundle with connection, if we replace the Levi–Civita connection ∇ by the connection $D = d + A$. Note that we input two vectors into the curvature and get an infinitesimal rotation out. Further, it is clearly anti-symmetric with respect to the input vectors. Thus, the curvature is an endomorphism-valued two-form.

EXERCISE 1.4.5. *Given the above definition, compute the Riemann tensor ${R_{ijk}}^l$ defined by $[\nabla_i, \nabla_j]\partial_k = {R_{ijk}}^l\partial_l$. Note that the tangent space at the identity to the space of rotations is the space of anti-symmetric matrices, and infer from the norm-preserving property of the Levi–Civita connection that the Riemann tensor obeys the anti-symmetry $R_{ijkl} = -R_{ijlk}$ (we had to use the metric to identify a matrix with a bilinear form, or "lower indices").*

EXERCISE 1.4.6. *For the tangent bundle, use the definition of curvature to derive ${(R_{ij})^k}_l$ in terms of the ${\Gamma^k}_{ij}$'s.*

We can use a shorthand to write $R = D^2$, where $D = d + A$. Then $R = dA + A \wedge A$. Here one must use the wedge product in conjunction with

[5] A two-form returns a number (or in this case endomorphism) given any pair of vectors.

[6] This is hardly a derivation; we are merely trying to capture the gist of curvature in giving its definition.

the commutator of endomorphisms. To make sense of this formula, it may be best to work out the previous exercise.

EXERCISE 1.4.7. *On a sphere, we can write the Riemannian curvature as* $R = \begin{pmatrix} 0 & \mathcal{R} \\ -\mathcal{R} & 0 \end{pmatrix} d\theta \wedge d\phi$. *Note that an infinitesimal* $SO(2)$ *matrix is an anti-symmetric matrix, as indicated (*$SO(2)$ *is a consequence of the norm-preserving or* metric *condition of the Levi-Civita connection). There is one independent component,* $\mathcal{R}$, *the scalar curvature. Show, using any choice of metric (e.g., the round metric), that* $\int_{S^2}(\mathcal{R}/2\pi)d\theta d\phi = 2$. *This is called the Euler characteristic and is our first taste of differential topology.*

1.5. Differential Forms

In this section, we look at some constructions using differential forms, the principal one being integration. In the previous exercise, we were asked to perform an integration over several coordinates. Of course, we know how to integrate with arbitrary coordinates, after taking Jacobians into consideration. This can be cumbersome. The language of differential forms makes it automatic.

1.5.1. Integration. Consider $\int f(x,y)dxdy$ on the plane. In polar coordinates, we would write the integrand as $f(r,\theta)rdrd\theta$, where r is the Jacobian

$$J = \det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r.$$

Note, though, that as a differential form, the two-form

$$dx \wedge dy = \left(\frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta\right) \wedge \left(\frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta\right) = r dr \wedge d\theta.$$

Therefore, differential forms actually *encode* Jacobians as transformation rules for changing coordinates (or patches). Here lies their beauty. If we write $\theta = f(x,y)dx \wedge dy = \theta_{ab}dx^a \wedge dx^b$ (take $x^1 = x, x^2 = y$), so that $\theta_{xy} = -\theta_{yx} = f/2$, then $\theta_{r\theta} = \theta_{ab}\frac{\partial x^a}{\partial r}\frac{\partial x^b}{\partial \theta} = (f/2)\,\epsilon_{ab}\frac{\partial x^a}{\partial r}\frac{\partial x^b}{\partial \theta} = (rf/2)$,[7] and we see that the Jacobian emerges from the anti-symmetry property of differential forms. More generally, if θ is an n-form on an n-manifold, then $\theta = f\,dx^1 \wedge \cdots \wedge dx^n$ in local coordinates, and in a new coordinate system $\widetilde{x}$,

[7] Here $\epsilon_{12} = -\epsilon_{21} = 1$, all others vanishing. In general ϵ is the totally anti-symmetric tensor in n indices, so $\epsilon_{1234\ldots n} = 1$, $\epsilon_{2134\ldots n} = -1$, etc.

$\theta = f\det\left(\frac{\partial x}{\partial \widetilde{x}}\right) d\widetilde{x}^1 \wedge \cdots \wedge d\widetilde{x}^n$, and the Jacobian is automatic. Now if instead θ is an m-form on an n-manifold (so $m < n$), then we can integrate θ over an m-dimensional submanifold C, since the restriction of θ to C makes sense. (Technically, on C we have the pull-back $i^*\theta$ of θ under the inclusion map $i : C \hookrightarrow M$, but this notation is often omitted.) In short, *we can integrate n-forms over n-folds.* No reference to coordinates is necessary.

An important form is the *volume form* associated to any metric. Note, as above, that the top form $dx^1 \ldots dx^n$ on an n-manifold is expressed as $|\det(\frac{\partial x}{\partial y})|dy^1 \ldots dy^n$ in a new coordinate system. Noting that the metric g_y in the y coordinates obeys $\sqrt{\det(g_y)} = \sqrt{\det(g_x)}|\det\left(\frac{\partial x}{\partial y}\right)|$, we see that the expression

$$dV = \sqrt{\det(g)}dx^1 \ldots dx^n \tag{1.1}$$

has the same appearance in any coordinate frame, up to a sign which is determined if we have an orientation. This is the volume form. It is natural, too, in that $\det(g)$ is the inner product (as a 1×1 matrix) inherited from g on totally anti-symmetric n-tensors. The norm, then, is given by the square root.

The volume form allows us to compute a global inner product on vector fields, forms, etc., defined over the whole manifold. We define $(\theta, \psi) = \int_M \langle\theta,\psi\rangle dV$, for any two forms θ and ψ of the same degree. Note that $(\psi,\psi) \geq 0$, with equality if and only if $\psi \equiv 0$.

EXERCISE 1.5.1. *Show that the area of a sphere of radius* R *is* $4\pi R^2$. *Use several sets of coordinates.*

1.5.2. The de Rham Complex. The main tool of differential topology is the de Rham complex. This is an elegant generalization to arbitrary manifolds of the three-dimensional notions of divergence, gradient, curl, and the identities $\mathbf{curl\,grad} = \mathbf{0}$, $\mathbf{div\,curl} = 0$.

We define the exterior derivative d to generalize the total differential of a function (df) to arbitrary forms. Define

$$df = \frac{\partial f}{\partial x^a}dx^a,$$

where again dx^a are cotangent vectors (note $dx^a = d(x^a)$, when x^a is considered as a coordinate function, so there is no abuse of notation). Note that $d(fg) = (df)g + f(dg)$ by the product rule, so d is a derivation. We extend

d to arbitrary forms by defining

$$d(\theta_1 \wedge \theta_2) = d\theta_1 \wedge \theta_2 + (-1)^{|\theta_1|}\theta_1 \wedge d\theta_2,$$

where the forms θ_1 and θ_2 are taken to be of homogeneous degree and $|\theta_1|$ represents the degree of the form. Arbitrary forms are sums of homogeneous forms, and d is taken to be linear. These rules uniquely specify d. For example, if $\theta = \theta_a dx^a$ is a one-form, then

$$d\theta = d(\theta_a) \wedge dx^a - \theta_a d(dx^2) = \frac{\partial \theta_a}{\partial x^b} dx^b \wedge dx^a + 0 \equiv \theta_{ba} dx^b \wedge dx^a,$$

where we define $\theta_{ba} = \frac{1}{2}(\frac{\partial}{\partial x^b}\theta_a - \frac{\partial}{\partial x^a}\theta_b)$ (the equality holds due to anti-symmetry). In general, if $\theta = \theta_{a_1 \dots a_p} dx^{a_1} \wedge \cdots \wedge dx^{a_p}$ is a p-form, then $d\theta = \sum_k (\partial_k \theta_{a_1 \dots a_p}) dx^k \wedge dx^{a_1} \wedge \cdots \wedge dx^{a_p}$. Most importantly, one checks that $d^2 = 0$

EXERCISE 1.5.2. *Prove that commutativity of partial derivatives is essential.*

Let $\Omega^p(M)$ be the space of p-forms on an n-dimensional manifold M. Then $d : \Omega^p \to \Omega^{p+1}$ and $d^2 = 0$. We can then form the complex

$$0 \longrightarrow \Omega^0 \longrightarrow \Omega^1 \longrightarrow \Omega^2 \longrightarrow \dots \longrightarrow \Omega^n \longrightarrow 0,$$

with d providing the maps. The complex terminates because there are no anti-symmetric $(n+1)$-tensors on an n-manifold. The composition of successive maps is zero, so we see that $\operatorname{Im} d \subset \operatorname{Ker} d$ at any given stage. Forms in $\operatorname{Ker} d$ are called "closed"; forms in $\operatorname{Im} d$ are called "exact." The de Rham cohomology is defined as closed modulo exact forms:

$$H^p(M) \equiv \{\operatorname{Ker} d\}/\{\operatorname{Im} d\}|_{\Omega^p}.$$

EXAMPLE 1.5.1. *Consider the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$. $H^0(T^2) = \mathbb{R}$, since closed zero-forms are constant functions, and there are as many of them as there are connected components of the manifold. The one-form dx is well defined and closed, but is* not *the derivative of a function, since x is not single-valued on the torus (e.g., x and $x+1$ represent the same point). One can show that any other closed one-forms are either exact or differ from $adx + bdy$ (a, b constants) by an exact form, so $H^1(T^2) = \mathbb{R}^2$. Likewise, $dx \wedge dy$ generates H^2. There are other representatives of $H^1(T^2)$. For example, consider $\delta(x)dx$, where $\delta(x)$ is a delta function. This is not exact, since $\int_{x=-1/2}^{x=1/2} df = 0$ for any function on the torus, but $\int_{x=-1/2}^{x=1/2} \delta(x)dx = 1$, just as $\int_{x=-1/2}^{x=1/2} dx = 1$. Note that $\delta(x)dx$ has the property that it is only supported along the* circle *$\{x = 0\}$, likewise, for $\delta(y)dy$. Also, $\delta(x)dx \wedge \delta(y)dy$ is only supported at a* point, *the intersection of the two circles. The relation between wedging de Rham cohomology classes and intersecting homology cycles will be explored in further chapters.*

The "Betti number" $b_k(M)$ is defined to be the dimension of $H^k(M)$.

1.5.3. The Hodge Star. The "Hodge star" operator $*$ encodes the inner product as a differential form. For any p-form ψ, define $*$ by the formula

$$\langle \theta, \psi \rangle dV = \theta \wedge *\psi,$$

where dV is as in Eq. (1.1), for any θ of the same degree. Clearly, $* : \Omega^p(M) \to \Omega^{n-p}(M)$. Defining this operation in terms of indices can be rather ugly. If $\theta_1, \dots, \theta_n$ is an *orthonormal* basis of one-forms, then $*\theta_1 = \theta_2 \wedge \cdots \wedge \theta_n$, etc., and $*\theta_I = \theta_{I^c}$, where I is some subset of $\{1, \dots, n\}$ and I^c is its (signed) complement. Then $*dx^I = \sqrt{|\det(g)|}dx^{I^c}$ (sometimes we will simply write g for $\det(g)$). Clearly $** = \pm 1$, and counting minus signs gives $** = (-1)^{p(n-p)}$. Note that since $*$ is invertible, it identifies Ω^p with Ω^{n-p}.

EXERCISE 1.5.3. *Rewrite the operators of divergence, curl, and gradient in terms of the exterior derivative, d. You will need to use the Euclidean metric on $\mathbb{R}^3$ to identify vectors and one-forms, and $*$ to identify two-forms with one-forms (e.g., $*(dy \wedge dz) = dx$ and the like) and three-forms with functions. Rewrite the relations* **curl grad = 0** *etc., in terms of $d^2 = 0$. This exercise is essential. Note that not every vector field on a region U whose curl is zero comes from a function. The extent to which such vectors exist is measured by $H^1(U)$.*

Using this exercise, we can understand the fundamental theorem of calculus, Stokes's theorem, and the divergence theorem as the single statement

$$\int_C d\theta = \int_{\partial C} \theta,$$

where ∂C represents the boundary of C (we have neglected some issues of orientation).

We can use the Hodge star operator and the global inner product to define the adjoint to the exterior derivative. Define the adjoint $d^\dagger$ of d by

$(d\theta, \psi) = (\theta, d^\dagger \psi)$. We state without proof that for p-forms on an n-manifold,

$$d^\dagger = (-1)^{np+n+1} * d * .$$

Note that $d^\dagger : \Omega^p \to \Omega^{p-1}$. It is clear that $d^2 = 0$ implies $(d^\dagger)^2 = 0$.

The Laplacian is defined as $\triangle = dd^\dagger + d^\dagger d$. Note that since the adjoint operator (equivalently, $*$) depends on the inner product, *the Laplacian depends on the metric.* We now show that the kernel of $\triangle$ is constituted by precisely those forms that are closed (annihilated by d) and co-closed (annihilated by $d^\dagger$). For if $\triangle\phi = 0$, then $(\phi, (dd^\dagger + d^\dagger d)\phi) = 0$, and by the definition of *adjoint* this equals $(d\phi, d\phi) + (d^\dagger\phi, d^\dagger\phi)$, which is zero if and only if $d\phi = 0$ and $d^\dagger\phi = 0$. Let $\mathcal{H}^p(M)$ denote the vector space of harmonic p-forms.

EXAMPLE 1.5.2. *On the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, with the metric defined from Euclidean $\mathbb{R}^2$, $\mathcal{H}^1$ is two-dimensional and generated by dx and dy. Note that here there is no choice of representatives (up to a choice of basis for $\mathbb{R}^2$). It is no coincidence that $\mathcal{H}^2 \cong H^2$, as we see below.*

Hodge decomposition is the theorem that every form θ (on a compact manifold with positive-definite metric) has a *unique* decomposition as

$$\theta = h + d\alpha + d^\dagger\beta,$$

where h is harmonic. Uniqueness follows by showing that zero (the difference of two decompositions) is uniquely written as the zero composition — namely, $0 = h + d\alpha + d^\dagger\beta$ implies $d^\dagger\beta = 0$, etc. This is clear, since d in this equation gives $0 = dd^\dagger\beta$, which after taking the inner product with β says $d^\dagger\beta = 0$ (use the adjoint). Existence of the decomposition is related to the fact that $\triangle$ is invertible on the orthogonal complement of its kernel.

Since $dd^\dagger\beta \neq 0$ for $d^\dagger\beta \neq 0$, the kernel of d comprises all forms that look like $h + d\alpha$. Further, all forms $d\alpha$ are precisely the image of d. We therefore conclude that kernel mod image can be identified with harmonic forms:

$$\mathcal{H}^p(M) = H^p(M, \mathbb{R})$$

(equality as vector spaces). This identification, of course, depends on the metric. Note that harmonic forms, unlike cohomology classes, do not form a ring, since the wedge product of two harmonic forms is not a harmonic form (though it lies in a cohomology class with a unique harmonic representative).

EXAMPLE 1.5.3. *By the wave equation, a vibrating drum has frequencies corresponding to eigenvalues of the Laplacian. The set of eigenvalues of the Laplacian is a measure of the geometry of the space. However, the set of zero modes is a topological quantity.*

CHAPTER 2

Algebraic Geometry

In this chapter outline the very basic constructions of algebraic geometry: projective spaces and various toric generalizations, the hyperplane line bundle and its kin, sheaves and Čech cohomology, and divisors. The treatment is driven by examples.

The language of algebraic geometry pervades the mathematical proof of mirror symmetry given in Part 4. Toric geometry is also crucial to the physics proof in Part 3. Sec. 2.2.2 on toric geometry is only a prelude to the extensive treatment in Ch. 7.

2.1. Introduction

In this chapter we will introduce the basic tools of algebraic geometry. Many of the spaces (manifolds or topological spaces) we encounter are defined by equations. For example, the spaces $x^2 + y^2 - R^2 = 0$ for different values of R are all circles if $R > 0$ but degenerate to a point at $R = 0$. Algebraic geometry studies the properties of the space based on the equations that define it.

2.2. Projective Spaces

Complex projective space $\mathbb{P}^n$ is the space of complex lines through the origin of $\mathbb{C}^{n+1}$. Every nonzero point in $\mathbb{C}^{n+1}$ determines a line, while all nonzero multiples represent the same line. Thus $\mathbb{P}^n$ is defined by

$$\mathbb{P}^n \equiv \left(\mathbb{C}^{n+1} \setminus \{0\}\right)/\mathbb{C}^*.$$

The group $\mathbb{C}^*$ acts to create the equivalence

$$[X_0, X_1, \ldots, X_n] \sim [\lambda X_0, \lambda X_1, \ldots, \lambda X_n],$$

where $\lambda \in \mathbb{C}^*$. The coordinates $X_0, \ldots, X_n$ are called "homogeneous coordinates" and are redundant (by one) for a description of projective space. In a patch U_i where $X_i \neq 0$, we can define the coordinates $z_k = X_k/X_i$ (for $k \neq i$). These coordinates are not affected by the rescaling.

EXAMPLE 2.2.1 ($\mathbb{P}^1$). *Consider $\mathbb{P}^1$, on which $[X_0, X_1] \sim [\lambda X_0, \lambda X_1]$. We can describe the whole space with two patches (for ease of notation, we use no indices in this example): $U = \{X_0 \neq 0\}$ with coordinate $u = X_1/X_0$ (well defined) and $V = \{X_1 \neq 0\}$ with coordinate $v = X_0/X_1$. On $U \cap V$, $v = 1/u$. Note that $X_0 = 0$ is well defined, as the scaling does not affect the solutions (the solution set of any homogeneous equation is well defined). From this we see that $\mathbb{P}^1$ is the same as S^2 as a complex manifold.*

Note that a linear action on $X_0, \ldots, X_n$ induces a holomorphic automorphism of $\mathbb{P}^n$, where an overall scaling acts trivially. It turns out that $PGL(n+1, \mathbb{C})$ is precisely the group of holomorphic automorphisms of $\mathbb{P}^n$.

Any homogeneous polynomial f in $n+1$ variables defines a subspace (subvariety) of $\mathbb{P}^n$ via the equation $f(X) = 0$, which respects the scaling relation. The equation would make no sense if f were not homogeneous.

EXAMPLE 2.2.2. *Consider a degree 3 polynomial in $\mathbb{P}^2$, $f = a_1X^3 + a_2Y^3 + a_3Z^3 + a_4XYZ + a_5X^2Y + \cdots + a_{10}YZ^2$. There are ten parameters, eight of which can be removed by a homogeneous, linear change of variables (a motion induced by $PGL(3, \mathbb{C})$), and one of which corresponds to an overall scaling. In all, there is one complex parameter that cannot be removed, and this determines the complex structure of the curve defined by f. In fact, it is an elliptic curve (Riemann surface of genus 1), and the value of its complex structure parameter $j(\tau)$ is an algebraic function of the one independent combination of the a_i.*

Using the same reasoning (not always valid, but okay here), a degree 5 ("quintic") polynomial in $\mathbb{P}$ would describe a manifold with $\binom{5+5-1}{5-1} - (25-1) - 1 = 101$ parameters describing its complex structure. (Here we have used the fact that the number of independent degree d homogeneous polynomials in n variables is $\binom{d+n-1}{n-1}$.)

At this point, we should note that algebraic geometry can be defined over arbitrary fields, and that the "algebraic" part of the story should be taken seriously. We will mainly be interested in algebraic varieties as (possibly singular) manifolds, so for our purposes "variety" can mean a manifold or a manifold with singularities. We mainly employ the tools of algebraic geometry to simplify calculations that would be well posed in a more general setting. In a sense, algebraic geometry is simpler than differential geometry since all quantities are algebraic, therefore holomorphic, or at worst meromorphic.

Note:

- $\mathbb{P}^n$ is a quotient space, or space of $\mathbb{C}^*$ orbits.
- We remove 0 so that $\mathbb{C}^*$ acts without fixed points.
- Open sets are complements of solutions to algebraic equations (in the above, $X_0 = 0$ and $X_1 = 0$). This is the Zariski topology.
- $(\mathbb{C}^*)^n$ acts on $\mathbb{P}^n$ via the action inherited from $\mathbb{C}^{n+1}$ (in fact, all of $PGL(n+1)$ acts), with fixed points $p_i = [0, \ldots, 1, \ldots, 0]$, $i = 0, 1, \ldots, n$.
- The quotiented scaling action is encoded in the way the coordinates scale (all equally for $\mathbb{P}^n$), so this is combinatorial data.

2.2.1. Weighted Projective Spaces. Weighted projective spaces are defined via different torus actions. Consider the $\mathbb{C}^*$ action on $\mathbb{C}^4$ defined by $\lambda : (X_1, X_2, X_3, X_4) \mapsto (\lambda^{w_1}X_1, \lambda^{w_2}X_2, \lambda^{w_3}X_3, \lambda^{w_4}X_4)$ (different combinatorial data). We define

$$\mathbb{P}^3_{(w_1,w_2,w_3,w_4)} = \left(\mathbb{C}^4 \setminus \{0\}\right)/\mathbb{C}^*.$$

Suppose $w_1 \neq 1$. Then choose $\lambda \neq 1$ such that $\lambda^{w_1} = 1$. Note that $(X_1, 0, 0, 0) = (\lambda^{w_1}X_1, 0, 0, 0)$, so we see the $\mathbb{C}^*$ action is not free (there are fixed points), and we have a $\mathbb{Z}/w_1\mathbb{Z}$ quotient singularity in the weighted projective space at the point $[1, 0, 0, 0]$.[1] Since this singularity appears in codimension 3, a subvariety of codimension 1 will generically not intersect it — so it may not cause any problems. However, suppose $(w_2, w_3) \neq 1$, so that $k|w_2$ and $k|w_3$, with $k > 1$. Then choose $\lambda \neq 1$ such that $\lambda^k = 1$. Note $(0, X_2, X_3, 0) = (0, \lambda^{w_2}X_2, \lambda^{w_3}X_3, 0)$, and we have a $\mathbb{Z}/k\mathbb{Z}$ quotient singularity along a locus of points of codimension 2. We can no longer expect a hypersurface to avoid these singularities. (We will see in later chapters that there are ways to "smooth" singularities.)

[1] A quotient singularity means that the tangent space is no longer Euclidean space, but rather the quotient of Euclidean space by a finite group. For example, $\mathbb{C}^2/\mathbb{Z}_2$, where $\mathbb{Z}_2 \equiv \mathbb{Z}/2\mathbb{Z}$ acts by $(-1, -1)$, is singular at the origin. One can construct a model for this space using the invariant polynomials $a = X_1^2, b = X_2^2$, and $c = X_1X_2$, which obey $p = ab - c^2 = 0$, a quadratic polynomial in $\mathbb{C}^3$. The singularity at the origin appears as a point where both $p = 0$ and $dp = 0$ have solutions. Singularities are discussed at greater length in Sec. 7.5.

EXAMPLE 2.2.3. *We denote by* $M = \mathbb{P}^4_{1,1,1,1,2}[6]$ *the hypersurface defined by a quasi-homogeneous polynomial of degree six in* $\mathbb{P}^4_{1,1,1,1,2}$. *For example,* M *may be the zero-locus of* $f = X_1^6 + X_2^6 + X_3^6 + X_4^6 + X_5^3$, *a Fermat-type polynomial. The singular point in* $\mathbb{P}^4_{1,1,1,1,2}$ *is* $[0,0,0,0,1]$, *since* $(0,0,0,0,1)$ *is fixed under* $\mathbb{Z}/2\mathbb{Z} \subset \mathbb{C}^*$. *However, since* $f(0,0,0,0,1) \neq 0$, *the singularity does not intersect the hypersurface, and the hypersurface is smooth (one must check that* $f = 0$ *and* $df = 0$ *has no common solution in* $\mathbb{P}^4_{1,1,1,1,2}$, *and this is immediate, as the origin is excluded).*

More generally, we can construct $\mathbb{P}^{n-1}_{\vec{w}}$, and we can expect hypersurfaces in this space to be smooth if $(w_i, w_j) = 1$ for all $i \neq j$. Again, this space have a $(\mathbb{C}^*)^{n-1}$ action depending on the vector $\vec{w}$.

2.2.2. Toric Varieties. Toric varieties are defined similarly and are even more general. We start with $\mathbb{C}^N$ and an action by an algebraic torus $(\mathbb{C}^*)^m$, $m < N$. We identify and then subtract a subset U that is fixed by a continuous subgroup of $(\mathbb{C}^*)^m$, then safely quotient by this action (up to finite quotient singularities) to form

$$\mathbb{P} = (\mathbb{C}^N \setminus U)/(\mathbb{C}^*)^m.$$

The resulting space $\mathbb{P}$ is called a toric variety, as it still has an algebraic torus action by the group $(\mathbb{C}^*)^{N-m}$ descending from the natural $(\mathbb{C}^*)^N$ action on $\mathbb{C}^N$.

EXAMPLE 2.2.4. *Here we give four examples of toric varieties, along with the diagrams (fans) that encode their combinatorial data (see Fig. 1). However, we will not give a general account of going from the diagram to the construction of the variety. The reader can find a much more thorough treatment in Ch. 7.*

A) The three vectors v_i *in the toric fan (A) are not linearly independent. They satisfy the relation* $1 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3 = 0$. *The coefficients* $(1,1,1)$ *in this relation encode the scaling action under* $\lambda \in \mathbb{C}^* : z_i \mapsto \lambda^1 z_i$. *Note that we have introduced a coordinate for each vector. Note that the triple of vectors* v_1, v_2, v_3 *are not all contained in a single cone, though any two of them are (there are three cones in the picture, the white areas). This encodes the data of the set* $U = \{z_1 = z_2 = z_3 = 0\}$. *When we take* $\mathbb{C}^3 \setminus U$, *the scaling action has no fixed points, and we can safely quotient by* $\mathbb{C}^*$. *The resulting smooth variety is, of course,* $\mathbb{P}^2$.

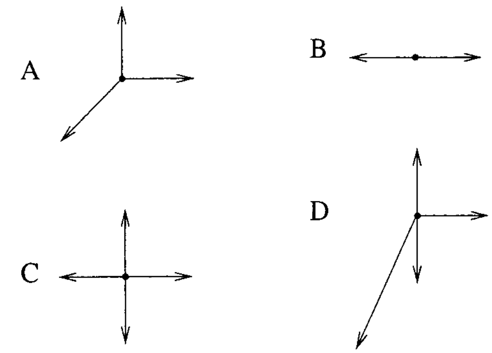


FIGURE 1. Four toric fans. A) The fan describing $\mathbb{P}^2$, consisting of three cones between three vectors: $(1,0), (0,1), (-1,-1)$. B) $\mathbb{P}^1$, described by two one-dimensional cones (vectors): 1 and -1. C) $\mathbb{P}^1 \times \mathbb{P}^1$. D) The Hirzebruch surface $F_n = \mathbb{P}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(n))$; the southwest vector is $(-1,-n)$.

This procedure is quite general, though the specifics will depend on the diagram.

B) In this one-dimensional diagram, there are two vectors that obey the relation $v_1 + v_2 = 0$. *Each vector generates a one-dimensional cone (ray). The* $\mathbb{C}^*$ *action is thus encoded by the weights* $(1,1)$: *namely,* $z_i \mapsto \lambda^1 z_i$. *As* v_1 *and* v_2 *are not contained in a common ray, we excise* $U = \{z_1 = z_2 = 0\}$. *The resulting space is* $\mathbb{P}^1$.

C.) Set $v_1 = (1,0)$, $v_2 = (-1,0)$, $v_3 = (0,1)$, *and* $v_4 = (0,-1)$. *Here there are two relations:* $v_1 + v_2 = 0$ *and* $v_3 + v_4 = 0$. *There are therefore two* $\mathbb{C}^*$ *actions encoded by the vectors* $(1,1,0,0)$ *and* $(0,0,1,1)$. *Namely,* $(\lambda_1, \lambda_2) \in (\mathbb{C}^*)^2$ *maps* $(z_1, z_2, z_3, z_4) \mapsto (\lambda_1^1 z_1, \lambda_1^1 z_2, \lambda_2^1 z_3, \lambda_2^1 z_4)$. *The set* U *is the union of two sets:* $U = \{z_1 = z_2 = 0\} \cup \{z_3 = z_4 = 0\}$. *Then* $(\mathbb{C}^4 \setminus U)/(\mathbb{C}^*)^2 = \mathbb{P}^1 \times \mathbb{P}^1$.

D.) The southwest vector here is $v_2 = (-1,-n)$, *all others the same as in (C), which is the special case* $n = 0$. *The construction of the toric variety proceeds much as in (C), except the first relation is now* $1 \cdot v_1 + 1 \cdot v_2 + n \cdot v_3 = 0$, *so the first* $\mathbb{C}^*$ *acts by* $(1,1,n,0)$. *The toric space is called the* n*th Hirzebruch surface, and denoted* F_n. *We can see that* F_n *resembles* $\mathbb{P}^1 \times \mathbb{P}^1$, *except the second* $\mathbb{P}^1$ *intermingles with the first. In fact,* F_n *is a fibration of* $\mathbb{P}^1$ *over*

$\mathbb{P}^1$, *trivial when* $n = 0$. *We will return to explaining the caption in later sections.*

E.) Another interesting example (not pictured) is to take the diagram from (A) and shift it one unit from the origin in $\mathbb{R}^3$. *That is, take* $v_1 = (1,1,0)$, $v_2 = (1,0,1)$, $v_3 = (1,-1,-1)$, *and* $v_0 = (1,0,0)$ *(the origin becomes a vector after the shift). The single relation among these four vectors is* $(-3,1,1,1)$. *Let* ζ *be the coordinate associated to* v_0. *Then* U *is still* $\{z_1 = z_2 = z_3 = 0\}$, *as in (A), since* v_0 *is contained in all (three-dimensional) cones. The resulting space is* $(\mathbb{C}^4 \setminus U)/\mathbb{C}^*$ *and has something to do with* $\mathbb{P}^2$ *(and with the number 3). In fact, we recover* $\mathbb{P}^2$ *if we set* $\zeta = 0$. *Also, the space is not compact. We will see that this corresponds to a (complex) line bundle over* $\mathbb{P}^2$.

2.2.3. Some Line Bundles over $\mathbb{P}^n$. From the definition of $\mathbb{P}^n$ we see there is a natural line bundle over $\mathbb{P}^n$ whose fiber over a point l in $\mathbb{P}^n$ is the line it represents in $\mathbb{C}^{n+1}$. Define $J \subset \mathbb{P}^n \times \mathbb{C}^{n+1}$ to be $\{(l,v) : v \in l\}$. J is called the "tautological line bundle." Suppose we have coordinates X_k on $\mathbb{C}^{n+1}$ with which to describe the point v. Then X_k is a linear map from the fiber J_l to $\mathbb{C}$. In other words, X_k is a section of $\mathrm{Hom}(J,\mathbb{C})$, the line bundle *dual* to J. Let us call this H. Note that the equation $X_k = 0$ makes sense on $\mathbb{P}^n$, and its solution defines a hyperplane (hence the "H").

From J and its dual H we get lots of line bundles by considering $J^{\otimes d} = J \otimes J \otimes \cdots \otimes J$ and $H^{\otimes d}$. The transition functions of these bundles are respectively dth powers of the transition functions for J and H. $H^{\otimes d}$ is also written H^d or $\mathcal{O}_{\mathbb{P}^1}(d)$ or $\mathcal{O}(d)$. The trivial line bundle $\mathcal{O}(0)$ is also written $\mathcal{O}$. In fact, sometimes additive notation is used for line bundles, so it is not uncommon to see H^d as dH as well. We will try to be sensitive to these ambiguities. Note that the dual of a line bundle has inverse transition functions, so $J = H^{-1} = \mathcal{O}(-1)$.

EXAMPLE 2.2.5. *Consider the hyperplane bundle on* $\mathbb{P}^1$. *According to the paragraph above, the coordinate* X_0 *is a (global) section. Let us see how this works. On* $U = \{X_0 \neq 0\}$, *a coordinate* u *parametrizes the points* $[1,u]$ *with* $X_0 = 1$. *On* V *the coordinate* v *parametrizes* $[v,1]$ *with* $X_0 = v$. *Thus* $(X_0)_U = s_{UV}(X_0)_V \Rightarrow s_{UV} = v^{-1} = u$, *and therefore* H *has a transition function* u. *Furthermore,* H^n *has a transition function* u^n *on* $\mathbb{P}^1$.

EXAMPLE 2.2.6. *As another example, consider diagram (D) from Fig. 1, with* v_4 *(the downward pointing vector) and the two cones containing it removed. The resulting diagram has three vectors* ($v_1 = (1,0)$, $v_2 = (-1,-n)$, $v_3 = (0,1)$), *two cones (generated by* v_1 *&* v_3 *and by* v_2 *&* v_3*), and one relation,* $(1,1,n)$ — *i.e.,* $v_1 + v_2 + nv_3 = 0$. *To construct the corresponding toric variety, we start with* $\mathbb{C}^3$ *and remove* $U = \{z_1 = z_2 = 0\}$ *(as* v_1 *and* v_2 *do not share a cone), and quotient by* $\mathbb{C}^*$ *acting as* $\lambda : (z_1,z_2,z_3) \mapsto (\lambda^1 z_1, \lambda^1 z_2, \lambda^n z_3)$. *Define* Z *to be the resulting space* $Z = (\mathbb{C}^3 \setminus U)/\mathbb{C}^*$. *Let us now rename the coordinates* $X_0 \equiv z_1$; $X_1 \equiv z_2$; $\theta \equiv z_3$. *We can cover* Z *with two patches* $\mathcal{U} = \{X_0 \neq 0\}$ *and* $\mathcal{V} = \{X_1 \neq 0\}$. *Note* $\mathcal{U} \cong U \times \mathbb{C}$, *where* U *is the open set on* $\mathbb{P}^1$ *coordinatized by* $u = X_1/X_0$ *(invariant under the scaling), and we parametrize* $\mathbb{C}$ *by* ζ_U. *Thus* (u,ζ_U) *represents (uniquely) the point* $(1,u,\zeta_U)$. *Also,* $\mathcal{V} \cong V \times \mathbb{C}$, *with coordinates* $v = X_0/X_1$ *and* ζ_V *representing* $(v,1,\zeta_V)$. *Consider a point* (X_0,X_1,θ) *in* $\mathbb{C}^3$, *with* $X_0 \neq 0$ *and* $X_1 \neq 0$. *On* U *we would represent it by coordinates* (u,ζ_U), *with* $u = X_1/X_0$ *and* $\zeta_U = \theta/X_0^n$. *The reason for the denominator in* ζ_U *is that we must choose* $\lambda \in \mathbb{C}^*$ *to be* $1/X_0$ *to establish the* $\mathbb{C}^*$ *equivalence* $(X_0,X_1,\theta) \sim (1, X_1/X_0, \theta/X_0^n) = (1,u,\zeta_U)$. *On* V *we represent the point by coordinates* $v = 1/u$ *and* $\zeta_V = \theta/X_1^n$. *Note* $\zeta_U = u^n\zeta_V$. *We have thus established that the space* Z *represented by this toric fan is described by two open sets* $U \times \mathbb{C}$ *and* $V \times \mathbb{C}$, *with* U *and* V *glued together according to* $\mathbb{P}^1$ *and the fibers* $\mathbb{C}$ *glued by the transition function* $s_{UV} = u^n$. *Therefore* $Z = \mathcal{O}(n)$.

It is now not too hard to see that the first scaling in Example 2.2.4 (D) defines the direct sum $\mathcal{O}(n) \oplus \mathcal{O}$. *The second relation and the set subtraction effects a quotienting by an overall scale in the* $\mathbb{C}^2$ *fiber directions. This is the projectivization of the direct sum bundle described in the caption to Fig. 1. The individual fibers are converted into* $\mathbb{P}^1$*'s, but this quotienting has a base* $\mathbb{P}^1$ *dependence, so* F_n *is a non-trivial* $\mathbb{P}^1$ *bundle over* $\mathbb{P}^1$ *for* $n \neq 0$.

Any linear function s_a of the coordinates X_k will also be a section of H. Further, since we know how the operator $\frac{\partial}{\partial X_k}$ behaves under scaling, we easily see that $\sum_{k=0}^n s_k \frac{\partial}{\partial X_k}$ descends to a linear differential operator (i.e., a vector field) on $\mathbb{P}^n$. In other words, we have a map of sections of bundles $\Gamma(H^{\oplus(n+1)}) \to \Gamma(T\mathbb{P}^n)$. Note that multiples of the vector $v \equiv \sum_k X_k \partial X_k$ descend to *zero* on $\mathbb{P}^n$ since this generates the very scaling by which we quotient. All such multiples look like fv, with f a function on $\mathbb{P}^n$, i.e., a section of the trivial line bundle $\mathbb{C}$. We have described an exact sequence

$$0 \longrightarrow \mathbb{C} \longrightarrow H^{\oplus(n+1)} \longrightarrow T\mathbb{P}^n \longrightarrow 0 \tag{2.1}$$

(the Γ symbol has been suppressed). This sequence is called the "Euler sequence."

2.3. Sheaves

A sheaf is a generalization of the space of sections of a vector bundle. Over any open subset U, the sections $\Gamma(U)$ form a vector space with an action (multiplication) by the space (ring) of functions. We generalize this notion to the algebraic setting by saying that a sheaf assigns abelian groups to each open set, and we require these groups to be modules under the action of holomorphic functions on the open set. The power of this restriction is that if the abelian groups are free, then they look like sections of vector bundles $\Gamma(U)$, but, if not, we can talk about more general objects, such as vector bundles living on subvarieties.

Roughly speaking, a sheaf is the data of sections on open sets, with sections on unions of sets determined by their restrictions to the different components.

Let us restrict ourselves to a complex manifold X. A "sheaf" $\mathcal{F}$ consists of

- abelian groups $\mathcal{F}(U)$ of "sections" σ, one for every open set U;
- restrictions $\sigma|_V \in \mathcal{F}(V)$ for any $V \subset U$, $\sigma \in \mathcal{F}(U)$ with the compatibility relations $(\sigma|_V)|_W = \sigma|_W$ for $W \subset V \subset U$;
- if $\sigma|_{U_i} = 0$ on all sets U_i of an open covering of U, then $\sigma = 0$ in $\mathcal{F}(U)$;
- if $\sigma \in \mathcal{F}(U)$, $\tau \in \mathcal{F}(V)$ and $\sigma|_{U\cap V} = \tau|_{U\cap V}$, then there exists $\rho \in \mathcal{F}(U \cup V)$ which restricts to σ and τ on U and V respectively (ρ is unique by the property immediately above).

EXAMPLE 2.3.1. *A) $\mathbb{Z}$ is the sheaf of integer-valued functions. Over U, $\mathbb{Z}(U)$ are the locally constant, integer-valued functions on U. Then $\mathbb{Z}(X)$ is the group of globally-defined integer-valued functions. This is a vector space of dimension equal to the number of connected components of X.*

B) $\mathbb{R}$ and $\mathbb{C}$ are sheaves of real and complex constant functions.

C) $\mathcal{O}$ is the sheaf of holomorphic functions. $\mathcal{O}(U)$ is the set of holomorphic functions. Again, $\dim \mathcal{O}(X)$ *is the number of connected components of X if X is compact, since the only global holomorphic functions on a compact connected space are constants.*

D) $\mathcal{O}^*$ is the sheaf of nowhere zero holomorphic functions.

E) Ω^p is the sheaf of holomorphic $(p,0)$-forms. $\theta \in \Omega^p(U)$ looks like $\theta_{a_1\ldots a_p} dz^{a_1} \wedge \cdots \wedge dz^{a_p}$, where $\theta_{a_1\ldots a_p}$ are holomorphic functions on U; note that no $d\bar{z}$'s appear.

F) $\mathcal{O}(E)$ are holomorphic sections of a holomorphic bundle E.

Sheaves enjoy many properties from linear and homological algebra. A map between sheaves defines maps on the corresponding abelian groups, and its kernel defines the kernel sheaf.

In particular, we can have exact sequences of sheaves. Consider, for example, the sequence

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O} \longrightarrow \mathcal{O}^* \longrightarrow 0, \tag{2.2}$$

where the first map is inclusion as a holomorphic function and the second is exponentiation of functions (times $2\pi i$). Note that the sequence does not necessarily restrict to an exact sequence on every open set (for example, on $\mathbb{C} \setminus \{0\}$ the exponential map is not onto), but is exact for open sets that are "small" enough. From now on, we restrict ourselves to covers of manifolds that consist of open sets with trivial cohomology.

If a sheaf $\mathcal{S}$ is the sheaf of sections of a vector bundle, then the *stalk* over a point p is the closest thing to a fiber of a vector bundle and is defined as the intersection (direct limit) of $\mathcal{S}(U)$ over all U containing p. The stalk can be thought of as germs of sections, or, by local triviality of vector bundles, germs of vector-valued functions.

EXAMPLE 2.3.2. *As an example of how a sheaf differs from a vector bundle, consider $\mathbb{P}^n$ and the sheaf $\mathcal{O}_{\mathbb{P}^n}$, the sheaf of holomorphic functions. This sheaf is also the sheaf of holomorphic sections of the trivial bundle, and the stalk over any point is the additive group of germs of holomorphic functions at that point. Now consider a subvariety $V \subset \mathbb{P}^n$. We can consider $\mathcal{O}_V$, a sheaf over V, or we can consider a sheaf over $\mathbb{P}^n$ with* support *only along V. As a sheaf over $\mathbb{P}^n$, $\mathcal{O}_V$ can be defined as holomorphic functions modulo holomorphic functions vanishing along V. So $\mathcal{O}_V(U)$ is the zero group if U does not intersect V. In fact, the ideal sheaf $\mathcal{J}_V$ of holomorphic functions (on $\mathbb{P}^n$) that* vanish *along V is another sheaf not associated to sections of a bundle.*

For instance, consider $\mathcal{O}_p$, the structure sheaf of a point in $\mathbb{P}^2$, namely, let $V = \{p\} \subset \mathbb{P}^2$. We define $\mathcal{J}_p$ to be the sheaf of holomorphic functions vanishing at p. Then $\mathcal{O}_p = \mathcal{O}_{\mathbb{P}^2}/\mathcal{J}_p$, which can also be written as the cokernel in the exact sequence

$$0 \longrightarrow \mathcal{J}_p \longrightarrow \mathcal{O} \longrightarrow \mathcal{O}_p \longrightarrow 0.$$

Note that the stalk of $\mathcal{O}_p$ over p is just the vector space $\mathbb{C}$ of possible values of holomorphic functions at p. Now the sheaf $\mathcal{J}_p$ is not a sheaf of sections of a vector bundle either, and if we want to express $\mathcal{O}_p$ in terms of sheaves that locally look like sections of bundles, we can do so in the following way. Note that p can be described as the zero set of two linear functions f, g on $\mathbb{P}^2$ (e.g., if $p = [1, 0, 0]$ we can take $f = X_1$ and $g = X_2$), i.e., two sections of $\mathcal{O}(1)$. Then $\mathcal{J}_p$ looks like all things of the form $fs_1 - gs_2$, where, in order to be a function, we must have $s_1, s_2 \in \mathcal{O}(-1)$. So the map $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathcal{O}$, where $(s_1, s_2) \mapsto fs_1 - gs_2$, has image $\mathcal{J}_p$. The kernel is not locally free but that can be taken care of with another map. In all, we have

$$\mathcal{O}(-2) \longrightarrow \mathcal{O}(-1) \oplus \mathcal{O}(-1) \longrightarrow \mathcal{O},$$

where the first map is $s \mapsto (gs, fs)$.

EXERCISE 2.3.1. *Check exactness of this sequence.*

If we call this whole sequence $\mathcal{E}^\bullet$, then the sequence $0 \longrightarrow \mathcal{E}^\bullet \longrightarrow \mathcal{O}_p \longrightarrow 0$ is exact, and in many ways $\mathcal{E}^\bullet$ behaves precisely like $\mathcal{O}_p$ (as it would if this were an exact sequence of vector spaces or modules).

2.3.1. Cohomology of Sheaves. We now develop the appropriate cohomology theory for investigating global questions about sheaves. As a consequence, we will have a long exact sequence in cohomology, given an exact sequence of sheaves.

Čech cohomology is defined for a sheaf relative to a cover $\{U_\alpha\}$ of X. Our restriction to "good" covers allows us to ignore this possible uncertainty and work with a fixed good cover $\{U_\alpha\}$.

That said, we define the (co-)chain complex via

$$\begin{aligned} C^0(\mathcal{F}) &= \textstyle\prod_\alpha \mathcal{F}(U_\alpha), \\ C^1(\mathcal{F}) &= \textstyle\prod_{(\alpha,\beta)} \mathcal{F}(U_\alpha \cap U_\beta), \\ &\vdots \end{aligned}$$

where we require $\sigma_{U_\alpha,U_\beta} = -\sigma_{U_\beta,U_\alpha}$ for $\sigma \in C^1(\mathcal{F})$, with higher cochains totally anti-symmetric. The differential $\delta_n : C^n \to C^{n+1}$ is defined by $(\delta_0\sigma)_{U,V} = \sigma_V - \sigma_U$; $(\delta_1\rho)_{U,V,W} = \rho_{V,W} - \rho_{U,W} + \rho_{U,V}$. Higher δ's are defined by a similar anti-symmetrizing procedure. Note that $\delta^2 = 0$ (we often ignore the subscripts). Čech cohomology is defined by

$$H^p(\mathcal{F}) = \operatorname{Ker} \delta_p / \operatorname{Im} \delta_{p-1}.$$

A key point is that an exact sequence of sheaves,

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0,$$

leads to a long exact sequence in cohomology,

$$0 \longrightarrow H^0(A) \longrightarrow H^0(B) \longrightarrow H^0(C) \longrightarrow H^1(A) \longrightarrow H^1(B) \longrightarrow \cdots.$$

In particular, the exact sequence Eq. 2.2 leads to the sequence $\cdots \to H^1(X, \mathcal{O}^*) \to H^2(X, \mathbb{Z})$. As we will see in the next section, any line bundle defines a class in $H^1(X, \mathcal{O}^*)$, and the image under the map to $H^2(X, \mathbb{Z})$ is called the "first Chern class" of the line bundle, $c_1(L)$. The line bundle is determined up to C^∞ isomorphism by its first Chern class, although two line bundles with the same first Chern class may not be isomorphic as holomorphic line bundles.

Recall that a section is determined by its restriction to open subsets. Therefore a global section of any sheaf is defined by its values on elements U_α of a covera and must be compatible on overlaps. Thus a global section σ consists of data σ_α such that $\sigma_\alpha = \sigma_\beta$ on $U_\alpha \cap U_\beta$; i.e., $\delta_0\sigma = 0$, and we see that the global sections $\mathcal{F}(X)$ are equal to $H^0(\mathcal{F})$.

EXAMPLE 2.3.3. *On $\mathbb{P}^1$ we can use our two open sets as a cover (warning: not a "good" cover), and a little thought shows that $H^1(\mathbb{P}^1, \mathcal{O}^*)$ is classified by maps from an annulus to an annulus (or, equivalently, circle to circle), which are in turn classified by a winding number. This makes sense, because line bundles are determined by how we glue two copies of $\mathbb{C}$ (with a nonzero function) together along an equatorial strip. Clearly $H^1(\mathbb{P}^1, \mathcal{O}^*) = \mathbb{Z}$, and the generator is $\mathcal{O}_{\mathbb{P}^1}(1)$, or just $\mathcal{O}(1)$. If U is the set $X_0 \neq 0$ with coordinate $u = X_1/X_0$ and V is the set $X_1 \neq 0$ with coordinate $v = X_0/X_1$, then $\mathcal{O}(1)$ has transition function $s_{UV} = u$ (on the equator $u = e^{i\theta}$, $s_{UV} = e^{i\theta}$ represents a map from S^1 to S^1 of degree 1). Note that $\mathcal{O}(1)$ is a holomorphic line bundle.*

EXAMPLE 2.3.4. *What are the global sections of $\mathcal{O}(n) \equiv (\mathcal{O}(1))^n$ on $\mathbb{P}^1$ (denoted $\mathcal{O}_{\mathbb{P}^1}(n)$)? Let us first recall that $\mathcal{O}(1)$ has the transition function $s_{UV} = u$, so $\mathcal{O}(n)$ has transition function u^n. Consider the monomials $f_V = v^k$ on V. To construct a global section, we need $f_U = s_{UV} f_V = u^n v^k = u^{n-k}$, which will be holomorphic as long as $k \leq n$. Therefore $1, v, \ldots, v^n$ give rise to $n+1$ global sections, and there can be no others. Equivalently, we can think of the coordinate v as representing the homogeneous coordinates $[X_0, X_1] = [1, v]$. Then the global sections can be generated by the monomials $X_0^n, X_0^{n-1}X_1, \ldots, X_1^n$. In short, the global sections are homogeneous polynomials of degree n.*

The same is true on $\mathbb{P}^N$: $H^0\mathcal{O}_{\mathbb{P}^N}(n)$ = homogeneous polynomials of degree n in $X_0, \ldots, X_N$. So $\dim H^0(\mathcal{O}(n)) = \begin{pmatrix} N+n-1 \\ n-1 \end{pmatrix}$. *In particular, the sections of $\mathcal{O}_{\mathbb{P}^4}(5)$ are quintic polynomials in five variables, and there are* $9 \cdot 8 \cdot 7 \cdot 6/4! = 126$ *independent ones.*

2.3.2. The Čech–de Rham Isomorphism.

(These few paragraphs are merely a summary of the treatment in [**121**], pp. 43–44.)

Here we show that the cohomology $H^*_{\mathrm{dR}}(M)$ defined from the de Rham complex on M is equal to the Čech cohomology $H^*(\mathbb{R})$. The proof depends on the fact (Poincaré lemma) that if θ is a p-form with $p > 0$ on $\mathbb{R}^n$ and $d\theta = 0$, then $\theta = d\lambda$. In other words, closed forms are locally exact, meaning we can find open sets on which their restrictions are exact. At $p = 0$, the constant forms are closed but not exact. Therefore, the sequence of sheaves

$$0 \longrightarrow \mathbb{R} \longrightarrow C^0 \longrightarrow C^1 \longrightarrow C^2 \longrightarrow \ldots$$

(here C^k represents k-forms) is exact. (Recall that exactness of a sequence of sheaves means that the sequence is exact for a sufficiently fine — e.g., contractible — cover of open sets.)

From this sequence we can construct a series of exact sequences. Let $Z^k \subset A^k$ represent the closed k-forms. We then have

$$\begin{gathered} 0 \longrightarrow \mathbb{R} \longrightarrow A^0 \xrightarrow{d} Z^1 \longrightarrow 0, \\ 0 \longrightarrow Z^1 \longrightarrow A^1 \xrightarrow{d} Z^2 \longrightarrow 0, \\ \vdots \\ 0 \longrightarrow Z^{k-1} \longrightarrow A^{k-1} \xrightarrow{d} Z^k \longrightarrow 0. \end{gathered}$$

The next result we use is that $H^k(A^p) = 0$ for $k > 0$. (This can be shown by using a partition of unity, but we omit the proof.) Then from the first short exact sequence we get a long exact sequence yielding $H^k(\mathbb{R}) \cong H^{k-1}(Z^1)$. The next short exact sequence tells us $H^{k-1}(Z^1) \cong H^{k-2}(Z^2)$. We proceed until the long exact sequence from the last sequence above gives

$$H^0(A^{k-1}) \xrightarrow{d} H^0(Z^k) \longrightarrow H^1(Z^{k-1}) \longrightarrow 0,$$

where the last zero comes from $H^1(A^{k-1}) = 0$. This says nothing other than

$$H^1(Z^{k-1}) \cong \frac{H^0(Z^k)}{dH^0(A^{k-1})} \equiv H^k_{\mathrm{dR}}(M).$$

At this point, it is helpful, albeit somewhat premature, to mention a similar result that holds for complex manifolds. The usual exterior derivative d is expressed in real coordinates x_a as $d = \sum_a dx_a \wedge \frac{\partial}{\partial x_a}$. With (half as many) complex coordinates z_k we can break up d into two parts: $d = \partial + \overline{\partial}$, where $\partial = \sum_k dz_k \wedge \frac{\partial}{\partial z_k}$ and $\overline{\partial}$ is the complex conjugate. Note that since ∂ is not real, we must take it to act on (the tensor powers of) $T^*M \otimes \mathbb{C}$. (We will have more to say about these operators later.) We also have $\partial^2 = 0$ and $\overline{\partial}^2 = 0$. A form in $\operatorname{Ker}\overline{\partial}$ is called $\overline{\partial}$-closed.

Note that $\overline{\partial}$-closed forms are holomorphic. What's more, $\overline{\partial}$ acts on forms taking values in any holomorphic vector bundle! The reason is that $\overline{\partial}$ commutes with holomorphic transition functions: Holomorphic means holomorphic no matter the trivialization. Thus if E is a holomorphic bundle on M, or, more specifically, its sheaf of sections, we have the sequence

$$0 \longrightarrow E \xrightarrow{\overline{\partial}} E \otimes A^{0,1} \xrightarrow{\overline{\partial}} E \otimes A^{0,2} \xrightarrow{\overline{\partial}} \ldots,$$

where $A^{0,k}$ are forms $\theta_{a_1 \ldots a_k} d\bar{z}^{a_1} \wedge \cdots \wedge d\bar{z}^{a_k}$, and can form the associated cohomology groups $H^k_{\overline{\partial}}(E)$.

Now the Čech–Dolbeault isomorphism follows from the $\overline{\partial}$-Poincaré lemma ($\overline{\partial}$-closed implies locally $\overline{\partial}$-exact) and the fact that $H^{p>0}(A^{0,k}) = 0$. The proof is exactly analogous and states that

$$H^k(E) \cong H^k_{\overline{\partial}}(E).$$

Therefore, on a complex n-fold X, we can think about the Čech cohomology classes $H^k(E)$ as E-valued forms with k anti-holomorphic indices. We define the canonical bundle K_X to be the bundle of forms with n holomorphic indices. Then $H^{n-k}(E^* \otimes K_X)$ are E^*-valued $(n, n-k)$-forms. Wedging, using the pairing of E and E^* and integrating, gives a map

$\int_X : H^k(E) \times H^{n-k}(E^* \otimes K_X) \to \mathbb{C}$. Serre duality, discussed more in the next chapter, says this pairing is perfect: $H^{n-k}(E^* \otimes K_X) = H^k(E)^*$.

2.4. Divisors and Line Bundles

A "line bundle" (in algebraic geometry) is a complex vector bundle of rank 1, with holomorphic transition functions.

EXAMPLE 2.4.1. *Some examples are: the trivial bundle,* $\mathbb{C}$, *whose holomorphic sections (i.e., functions) comprise the sheaf* $\mathcal{O}$; *the tautological line bundle* J *over projective space; its dual* $H \equiv J^* = \mathrm{Hom}(J, \mathbb{C})$. *Note that the homogeneous coordinates* X_i *are global sections of* H, *and that the set of zeroes of any global section of* H *(also called* $\mathcal{O}(1)$*) defines a hyperplane.* H^n *is the line bundle* $\mathcal{O}(n)$, *and its global sections are homogeneous polynomials of degree* n. *The canonical bundle* $K_X = \Omega^n$ *of holomorphic* $(n,0)$*-forms over any complex* n*-fold* X *is a holomorphic line bundle. As a generalization, given any holomorphic vector bundle* E *of rank* r, *we can form the (holomorphic) line bundle* $\Lambda^r E$, *the "determinant line bundle," whose transition functions are the determinants of those for* E.

Recall that the data of a line bundle is a local trivialization $\varphi_\alpha : \pi^{-1}(U_\alpha) \cong U_\alpha \times \mathbb{C}$ or equivalently a set of holomorphic transition functions $s_{\alpha\beta} = \varphi_\alpha \circ \varphi_\beta^{-1}$ such that values in $\mathbb{C}^*$ with

$$s_{\alpha\beta}s_{\beta\alpha} = 1 \qquad \text{and} \qquad s_{\alpha\beta}s_{\beta\gamma}s_{\gamma\alpha} = 1.$$

Recalling sheaf cohomology, this data states *precisely* that the transition functions $s_{\alpha\beta}$ are closed one-chains in the Čech cohomology of the sheaf $\mathcal{O}^*$ (using multiplicative notation for the group of sections). Further, a different local trivialization corresponding to an isomorphic line bundle is defined by isomorphisms (of $\mathbb{C}$), $f_\alpha \in \mathcal{O}^*(U_\alpha)$; the transition functions $s'_{\alpha\beta} = \frac{f_\alpha}{f_\beta} s_{\alpha\beta}$ are then thought of as equivalent. Note that s' and s differ by a trivial (exact) Čech one-cycle, f_α/f_β. So line bundles up to isomorphism are classified by closed Čech one-chains modulo exact chains. We learn that $H^1(X, \mathcal{O}^*)$ is the group of isomorphism classes of line bundles on X. This is called the "Picard group" of X. The group multiplication is the tensor product of line bundles, corresponding to ordinary multiplication of transition functions: i.e., on $L \otimes \widetilde{L}$ we have the transition functions $s_{\alpha\beta}\widetilde{s}_{\alpha\beta}$.

The relationship

$$\text{locally defined functions} \longleftrightarrow \text{line bundles}$$

can be investigated more closely. Any analytic, codimension 1 subvariety V has (locally) defining functions: $V \cap U_\alpha = \{f_\alpha = 0\}$ (chosen such that f_α has a simple zero along V). On $V \cap U_\beta$ we have f_β, and on the intersection f_α/f_β is nonzero (zeroes of the same order cancel). Therefore the data $\{f_\alpha\}$ define a line bundle with transition functions $s_{\alpha\beta} = f_\alpha/f_\beta$.

EXAMPLE 2.4.2. *On* $\mathbb{P}^1$ *define* $D = N + S$, *where* N *and* S *are the north and south poles. On* $U = \mathbb{P}^1 \setminus N$ *with local coordinate* $u = X_1/X_0$, D *is written as the zeroes of* $f_U = u$. *On* $V = \mathbb{P}^1 \setminus S$ *with coordinate* $v = X_0/X_1$, $D = \{f_V = -v = 0\}$, *and on the overlap* $f_U/f_V = -u/v = -u^2$ *(the minus sign was chosen for convenience, as we will see, and doesn't affect anything). The chain rule says* $\frac{\partial}{\partial v} = -u^2\frac{\partial}{\partial u}$, *which means* $T\mathbb{P}^1$ *(a line bundle) also has transition function* $s_{UV} = -u^2$.

EXERCISE 2.4.1. *Try this for two other points.*

Thus $D \leftrightarrow T\mathbb{P}^1$. *We further see from the power of* u *that* $T\mathbb{P}^1 \cong \mathcal{O}(2) \cong K_{\mathbb{P}^1}$.

EXAMPLE 2.4.3. *The bundle defined by a hyperplane in this way is the hyperplane line bundle.*

Generalizing this, we can define a "divisor"

$$D = \sum_i n_i V_i$$

to be a formal sum of irreducible hypersurfaces[2] with integer coefficients n_i. Any given V_i can be described on U_α as the zero set of a holomorphic function f^i_α, where the f^i_α are defined up to multiplication by a nowhere-vanishing holomorphic function (section of $\mathcal{O}^*$). In U_α, we associate to $D \cap U_\alpha$ its defining function $f_\alpha = \prod_i (f^i_\alpha)^{n_i}$, so that if $n_i > 0$, then the zero of f_α has order n_i along $V_i \cap U_\alpha$, while if $n_i < 0$, then f_α has a pole of order n_i. The f_α are nonzero meromorphic functions, and since f_α and f_β must agree (up

[2] A "hypersurface" is a codimension 1 submanifold that can be written locally as the zeroes of a holomorphic function, and "irreducible" means it cannot be written as the union of two hypersurfaces. In the sum we require that an infinite number of hypersurfaces cannot meet near any point ("locally finite"). Our "divisor" will mean "Weil divisor."

to $\mathcal{O}^*$) on overlaps, they define a global section of the sheaf of meromorphic functions modulo non-vanishing holomorphic functions:

$$\mathrm{Div}\,(M) = H^0(\mathcal{M}^*/\mathcal{O}^*).$$

To summarize, a divisor D, with local defining functions f_α as above, defines a line bundle $\mathcal{O}(D)$ with transition functions $s_{\alpha\beta} = f_\alpha/f_\beta$. Note that the f_α define a (meromorphic) section of $\mathcal{O}(D)$, since $f_\alpha = s_{\alpha\beta}f_\beta$, whose zero locus is D.

In practice, it is very convenient to be able to think of hypersurfaces in terms of the line bundles they describe. A hypersurface defines an element of real codimension 2 homology, and we will explore the relationship between this homology class and a class in de Rham cohomology (or Čech cohomology) associated to any line bundle by the map $H^1(\mathcal{O}^*) \to H^2(\mathbb{Z})$ from the sequence in Eq. 2.2. These and other topological issues are the subject of the next chapter.

CHAPTER 3

Differential and Algebraic Topology

We try to convey just a hint of what various cohomology theories and characteristic classes are, and how they are used in applications essential for understanding mirror symmetry. Our scope is necessarily limited.

3.1. Introduction

Many physical questions are topological in nature, especially questions involving so-called BPS states in supersymmetric theories, as there are typically an integer number of these non-generic states. In this chapter, we develop some of the topological tools required to address such physical questions. Since analytical methods in physics typically involve derivatives and integrals, our approach to topology will be mainly differential and algebraic. Again, our focus will be on gaining a quick understanding of some of the constructions used in mirror symmetry — or at least how they are applied in practice.

3.2. Cohomology Theories

In Ch. 2 we discussed de Rham cohomology $d : \Omega^p(M) \to \Omega^{p+1}(M)$ and Dolbeault cohomology $\overline{\partial}$ for complex manifolds, $\overline{\partial} : A^{0,p}(M) \otimes E \to A^{0,p+1}(M) \otimes E$, where E is any holomorphic vector bundle. Particularly interesting is the example when $E = \Lambda^q T_{\mathrm{hol}} M$.

For a sheaf E we can also construct the Čech cohomology $H^k(E)$. When $E = \mathbb{R}$ we have the Čech–de Rham isomorphism, and when E is the sheaf of holomorphic sections of a holomorphic bundle, we have Čech–Dolbeault.

For completeness, let us recall singular homology and cohomology. We define singular p-chains to be linear combinations of maps from p-simplices to a topological space, X. For a map $f : \Delta_p \to X$, the restriction to the kth face of Δ_p is denoted by f_k, $k = 0, \ldots, p$. Let C_p denote the p-chains. Then the boundary operator $\partial : C_p \to C_{p-1}$ is given by $\partial f = \sum_{i=0}^{p} (-1)^k f_k$, extended to chains by linearity, and the associated homology cycles are in $H_*(X)$. Singular cohomology is formed from cochains $C^p = \mathrm{Hom}(C_p, \mathbb{Z})$, with $d\theta(f) = \theta(\partial f)$, for θ a cochain and f a chain. Then for manifolds X one has the result $H^*_{\text{sing}} = H^*(\mathbb{Z})$. If for singular cohomology we take $\mathrm{Hom}(C_p, G)$ with G an arbitrary abelian group, we get $H^*_{\text{sing}}(X;G)$. If $G = \mathbb{R}$ and X is a smooth manifold, then we also have an isomorphism between singular and de Rham cohomology.

3.3. Poincaré Duality and Intersections

Our aim here is to describe Poincaré duality, an intersection pairing of (co)homology classes. In this section, H^* denotes de Rham cohomology.

The wedge product of forms descends to a map on cohomology $H^k \otimes H^l \to H^{k+l}$, since $\theta \wedge d\eta = d(\pm\theta \wedge \eta)$ if θ is closed. This plus Stokes's theorem on a closed (oriented) manifold X implies that integration gives a map $\int_X : H^k(X) \otimes H^{n-k}(X) \to \mathbb{R}$ (we assume X is compact, or else one of these cohomology groups must be of forms with compact support). Poincaré duality says that this pairing is perfect, meaning H^k and H^{n-k} are duals: $H^{n-k} = (H^k)^*$.

Now consider a k-dimensional, closed submanifold ($C \subset X$ such that $\partial C = 0$). For any $\theta \in H^k(X)$ we can define $\int_C \theta$. Stokes's theorem ensures that this is independent of the representative of the cohomology class. Thus $\int_C$ is a linear map $H^k \to \mathbb{R}$, and Poincaré duality says that we can represent this map by an $(n-k)$-form $\eta_C \in H^{n-k}$: i.e.,

$$\int_C \theta = \int_X \theta \wedge \eta_C.$$

η_C is called the Poincaré dual class.

In fact, a rather explicit construction of η_C can be achieved.[1] The key lies in a construction for a general (oriented) vector bundle, E.[2] We define on the total space of E the "Thom form" Φ, which is a delta function top form *along each fiber*, i.e., $\int_{E_x} \Phi = 1$ for any $x \in M$, where pull-back of Φ to the fiber is implicit. Next, we prove that a tubular neighborhood of a submanifold C is diffeomorphic to its normal bundle $N_{C/M}$ of C in M (defined by $0 \to TC \to TM|_C \to N_{C/M} \to 0$). Extending the Thom form of $N_{C/M}$ by zero, we get a cohomology class Φ_C on M whose degree equals the rank of C's normal bundle, i.e., the codimension of C.

[1] We only describe the steps; the reader can find references in Ch. 40.

[2] "Oriented" means that the transition functions have positive determinants.

Now one computes $\int_M \theta \wedge \Phi_C$ by noting that 1) it restricts to a tubular neighborhood T of C (since Φ_C was an extension by zero from T); 2) T can be thought of as a vector bundle, on which we integrate in base and fiber directions; 3) Φ_C is a top form in the normal directions, so only the part of θ along the base C can matter; 4) since $\int \Phi_C = 1$ along each fiber, the final answer is $\int_C \theta$. We deduce that $\int_M \theta \wedge \Phi_C = \int_M \theta$, so Φ_C represents the Poincaré dual class η_C. This is a woeful derivation! However, if we only want a vague sense of the reasoning, it may be adequate.

In conclusion, the Thom class of the normal bundle is the Poincaré dual class, which can therefore be chosen to have support along (or within an arbitrarily small neighborhood of) C.

EXAMPLE 3.3.1. *On a torus $T^2 = S^1 \times S^1$, the total space of the normal bundle to one of the S^1's (defined, say, by $\theta_2 = 0$) is equal to $S^1 \times T_0 S^1$, where $T_0 S^1$ is the tangent space to S^1 at $\theta_2 = 0$. The Thom class of the normal bundle is $\Phi = \delta(\theta_2) d\theta_2$, where $\delta(\theta_2)$ is a Dirac delta function. Indeed, it has support on the first S^1 and (*EXERCISE*) it satisfies $\int_{T^2} \theta \wedge \Phi = \int_{S^1} \theta$.*

Given submanifolds C and D whose codimensions add up to n, the degree of $\eta_C \wedge \eta_D$ is n, so $C \cdot D \equiv \int_X \eta_C \wedge \eta_D$ is a number. Given the fact that η_C and η_D can be chosen to have support along C and D, $C \cdot D$ picks up contributions only from the intersection points $x \in C \cap D$. If we assume that the intersections are transverse, then the bump forms will wedge to a volume form for $TM|_x$, and the integration will produce ± 1 from each x, depending on the orientation. In total, $C \cdot D = \sum_x (-1)^{\epsilon_x}$. More generally, we have the following relation:

$$\eta_{C \cap D} = \eta_C \wedge \eta_D$$

(to compare with the case discussed, integrate). So the intersection and wedge products are Poincaré dual.

3.4. Morse Theory

Because there are points in the treatment of quantum field theory where Morse theory is a helpful tool (see, e.g., Sec. 10.4), we include here a short discussion.

Consider a smooth function $f : M \to \mathbb{R}$ with non-degenerate critical points. If no critical values of f occur between the numbers a and b (say

$a < b$), then the subspace on which f takes values less than a is a deformation retract of the subspace where f is less than b. To show this, one puts a metric on the space and flows by the vector field $-\nabla f/|\nabla f|^2$, for time $b-a$ (this obviously runs afoul at critical points). Furthermore, the Morse lemma states that one can choose coordinates around a critical point p such that f takes the form $-(x_1^2 + x_2^2 + \cdots + x_\mu^2) + x_{\mu+1}^2 + \cdots + x_n^2$, where p is at the origin in these coordinates and $f(p)$ is taken to be zero. The difference between $f^{-1}(\{x \leq -\epsilon\})$ and $f^{-1}(\{x \leq +\epsilon\})$ can therefore be determined by this local analysis, and only depends on μ (the "Morse index"), the number of negative eigenvalues of the Hessian of f at the critical point. The answer is that $f^{-1}(\{x \leq +\epsilon\})$ can be obtained from $f^{-1}(\{x \leq -\epsilon\})$ by attaching a μ-cell along the boundary $f^{-1}(0)$. By "attaching a μ-cell" to a space X, we mean taking the standard μ-ball $B_\mu = \{|x| \leq 1\}$ in μ-dimensional space and identifying the points on the boundary $S^{\mu-1}$ with points in the space through a continuous map $f : S^{\mu-1} \to X$. That is, we take $X \amalg B_\mu$ with the relation $x \sim f(x)$ for $x \in \partial B_\mu = S^{\mu-1}$. In this way, we recover the homotopy type of M through f alone.

In fact, we can find the homology of M through a related construction. f defines a chain complex C_f^* whose kth graded piece is $\mathbb{C}^{\alpha_k}$, where α_k is the number of critical points with index k. The boundary operator ∂ maps C_f^k to C_f^{k-1}, $\partial x_a = \sum_b \Delta_{a,b} x_b$, where $\Delta_{a,b}$ is the signed number of lines of gradient flow from x_a to x_b, where b labels points of index $k-1$. Such a gradient flow line is a path $x(t)$ satisfying $\dot{x} = \nabla(f)$, with $x(-\infty) = x_a$ and $x(+\infty) = x_b$. To define this number properly, one must construct a moduli space of such lines of flow by intersecting outward and inward flowing path spaces from each critical point and then show that this moduli space is an oriented, zero-dimensional manifold (points with signs). These constructions are similar to ones that we will encounter when discussing solitons in Ch. 18. The proof that $\partial^2 = 0$ comes from the fact that the boundary of the space of paths connecting critical points whose index differs by 2 is equal to a union over compositions of paths between critical points whose index differs by 1. Therefore, the coefficients of the ∂^2 operator are sums of signs of points in a zero-dimensional space which is the boundary of a one-dimensional space. These signs must therefore add to zero, so $\partial^2 = 0$.

EXERCISE 3.4.1. *Practice these two constructions when M is a tire standing upright and f is the height function. Practice the following construction of homology as well. Do the same for a basketball. Try deforming the ball so that more critical points are introduced. Verify that the Morse homologies are not affected.*

The main theorem is

THEOREM 3.4.1.
$$H_*(C_f) = H_*(M).$$

Cohomology can be defined through the dual complex. In fact, by looking at Y-shaped graphs of gradient flow (three separate paths meeting at a common point), one can define a "three-point function" to produce a product on Morse cohomology. We will not use this construction, but it is closely related to the Fukaya category (when M is taken to be the space of paths between Lagrangian submanifolds), discussed in Sec. 37.7.1.

3.5. Characteristic Classes

In this section, we focus on the Chern classes.

If the rank of a holomorphic vector bundle equals the complex dimension of the base manifold, then dimension counting says that a generic section should have a finite number of zeroes. For example, on any complex manifold we can consider the holomorphic tangent bundle and the number of zeroes of a generic holomorphic vector field is the Euler characteristic (for a non-holomorphic vector field we must count with signs). In general, the integral of the top Chern class, also called the "Euler class", encodes this number. Of course, not all sections are generic and one must account for multiplicities of certain zeroes. Here we will explore some generic and non-generic examples.

EXAMPLE 3.5.1. *On $\mathbb{P}^1$ consider the holomorphic vector field $u\frac{\partial}{\partial u}$. It has a zero at $u = 0$. On the patch with coordinate $v = 1/u$, we must transform $\frac{\partial}{\partial u} = -v^2\frac{\partial}{\partial v}$, so $u\frac{\partial}{\partial u} = -v\frac{\partial}{\partial v}$, which has a simple zero at $v = 0$. Of course, this vector field is just the generator of a rotation, which has fixed points at the north and south poles. In total, there are two zeroes, and $\chi(\mathbb{P}^1) = \chi(S^2) = 2$.*

EXAMPLE 3.5.2. *On $\mathbb{P}^1$ we can consider the vector field $\frac{\partial}{\partial z}$, which has no zeroes on the patch with coordinate z. However, on the other patch this* *vector field equals* $-\tilde{z}^2\frac{\partial}{\partial \tilde{z}}$, *which has a zero of multiplicity* 2 *at* $\tilde{z}=0$, *and the total number of zeroes, counted appropriately, is two.*

EXAMPLE 3.5.3. *On* $\mathbb{P}^2$ *we consider three patches to cover the manifold:* $U=\{X_0\neq 0\}$ *with coordinates* $u_1 = X_1/X_0$ *and* $u_2 = X_2/X_0$, $V = \{X_1\neq 0\}$ *with coordinates* $v_1 = X_0/X_1$ ($=1/u_1$ *on the overlap) and* $v_2 = X_2/X_1 = u_2/u_1$, *and* $W=\{X_2\neq 0\}$ *with coordinates* $w_1 = X_0/X_2 = 1/u_2$ *and* $w_2 = X_1/X_2 = u_1/u_2$. *Consider the holomorphic vector field*

$$s = u_1\frac{\partial}{\partial u_1} + Cu_2\frac{\partial}{\partial u_2}.$$

We consider two cases: $C=1$ *and* $C=2$.

- *If* $C=2$, *this vector field has a zero in* U *where* $u_1=u_2=0$, *i.e., the point* $[1,0,0]$ *in homogeneous coordinates. To look in the other patches, we transform* $\frac{\partial}{\partial u_1} = \frac{\partial v_1}{\partial u_1}\frac{\partial}{\partial v_1} + \frac{\partial v_2}{\partial u_1}\frac{\partial}{\partial v_2} = -v_1^2\frac{\partial}{\partial v_1} - v_1v_2\frac{\partial}{\partial v_2}$. *Proceeding this way and converting u's to v's (remember* $C=2$*), we find that* $s = -v_1\frac{\partial}{\partial v_1} + v_2\frac{\partial}{\partial v_2}$, *so it has a zero at* $v_1=v_2=0$, *i.e.,* $[0,1,0]$, *in this patch. In* W, $s = -2w_1\frac{\partial}{\partial w_1} - w_2\frac{\partial}{\partial w_2}$, *so the final zero is at* $[0,0,1]$ *(which does not intersect the other patches). There are three zeroes and* $\chi(\mathbb{P}^2)=3$.
- *Consider* $C=1$. *Now we have a zero at* $[1,0,0]$ *in* U, *but in* V *we see* $s=-v_1\frac{\partial}{\partial v_1}$, *which has a family of zeroes where* $v_1=0$. *In* W, $s = w_1\frac{\partial}{\partial w_1}$, *which is zero when* $w_1=0$. *This family of zeroes is the* $\mathbb{P}^1\subset\mathbb{P}^2$ *where* $X_0=0$ *(the complement of* U*). In order to compute the contribution to the Chern class integral, we use the "excess intersection formula" (cf. Sec. 4.4.1 and Theorem 26.1.2). This states that the contribution from a zero-locus* Y *(here* $Y=\{X_0=0\}\cong\mathbb{P}^1$*) of some section of a vector bundle* E *(of the same rank as the manifold* M, *here* $E=T\mathbb{P}^2$ *and* $M=\mathbb{P}^2$*) contributes*

$$\int_Y \frac{c_{top}(E)}{c_{top}(N_{Y/M})}$$

to the top Chern class, where $N_{Y/M}$ *is the normal bundle of* $Y\subset M$. *In this example, the exact sequence*

$$0\longrightarrow T\mathbb{P}^1\longrightarrow T\mathbb{P}^2\longrightarrow N_{\mathbb{P}^1/\mathbb{P}^2}\longrightarrow 0$$

tells us that $c_{top}(T\mathbb{P}^2) = c_{top}(T\mathbb{P}^1)c_{top}(N_{\mathbb{P}^1/\mathbb{P}^2})$, *so after cancelling we find that the contribution of the zero-locus is*

$$\int_{\mathbb{P}^1} c_{top}(T\mathbb{P}^1) = \chi(\mathbb{P}^1) = 2 \ \textit{(e.g., from the example above).}$$

Summing up the zero-loci, $\chi(\mathbb{P}^2)=1+2=3$. *The section with* $C=1$ *is not generic enough, but, as we will see in mirror symmetry, one cannot always obtain a generic section.*

EXERCISE 3.5.1. *Find a holomorphic vector field on* $\mathbb{P}^n$ *with* $n+1$ *isolated zeroes.*

We now give an account of Chern classes, before actually defining them. Poincaré duality says that cycles in H_{n-p} are dual to H^p, and cohomology H^p is dual to H_p as well. Therefore we can identify H^p with H_{n-p}. The Chern classes c_k will be given in the next section as classes in H^{2k}, but here we will discuss them as $(n-2k)$-cycles, i.e., cycles of codimension $2k$. The relation between forms and cycles is also seen by the fact that a cohomology p-form can be chosen (in the same cohomology class) to vanish everywhere outside of an $(n-p)$-cycle. For example, on the circle S^1, the delta-function 1-form $\delta(\theta)d\theta$ has support on a point.

The examples above demonstrate that the top Chern class is the cycle associated to a generic section. For a rank r bundle, this is represented by a codimension r cycle or by an r-form. (When $r=n$ we get a collection of points, possibly with multiplicities.) In fact, since the base manifold sits in the total space of a bundle as the zero section, the top Chern class represents the intersection of a generic section with the zero section. So it makes sense that intersection theory is needed to account for zero sets of non-generic sections.[3]

We now give an account of all the Chern classes c_k, for $k\leq r$. Let E be a rank r complex vector bundle on an n-fold, M. Let $s_1,\ldots,s_r$ be r global sections of E (C^∞ but not necessarily holomorphic, so they exist). Define D_k to be the locus of points where the first k sections develop a linear dependence (i.e., $s_1\wedge\cdots\wedge s_k = 0$ as a section of $\Lambda^k E$). Then the cycles D_k are Poincaré dual to the Chern classes c_{r+1-k}. For example, when $k=1$

[3]In general, intersection theory and the excess intersection formula account for non-generic cases of the type considered here. We will not be able to develop this interesting subject much further, however.

the top Chern class c_r is represented by D_1, the zeroes of a single section. When $k = r$, the first Chern class represents the zeros of a section of the determinant line bundle formed by wedging r sections. Indeed, $c_1(E) = c_1(\Lambda^r E)$.

3.5.1. Chern Classes from Topology. We would like to impart a sense of how Chern classes capture the topology of a bundle. This section is independent of the rest of the chapter.

Just as $\mathbb{P}^{n-1}$, the space of complex lines through the origin in $\mathbb{C}^n$, is equipped with a tautological ("universal") line bundle $H^{-1} = \mathcal{O}_{\mathbb{P}^{n-1}}(-1)$, similarly the space $G_k(n)$ of complex k-planes through the origin in $\mathbb{C}^n$ has a universal rank k vector bundle. Clearly, we can include $G_k(n) \subset G_k(n+1)$, since $\mathbb{C}^n \subset \mathbb{C}^{n+1}$. To accommodate general bundles, it is convenient to define the infinite-dimensional space G_k as the direct limit $G_k(k) \subset G_k(k+1) \subset \ldots$. It is the set of k-planes in $\mathbb{C}^\infty$. It, too, has a universal k-bundle,

$$E_k \subset G_k \times \mathbb{C}^\infty, \quad E_k = \{(p, v) : v \in p\},$$

where p is a k-plane.

We will show below that $E \equiv E_k$ is universal in the sense that 1) any $\mathbb{C}^k$-bundle $F \to X$ (over any topological space X) is the "restriction" of E to X via some map $\varphi : X \to G_k$, i.e., the bundle F is isomorphic to the pull-back $\varphi^* E_k$; and 2) any two such maps are homotopic. Then a calculation shows that the cohomology of G_k is a copy of the integers in each even degree; we call the generators $c_i(E_k) \in H^{2i}(G_k)$. Then we can define Chern classes via pull-back in cohomology; setting $c(E_k) = \oplus_i c_i(E_k)$, we define the total Chern class of F to be $c(F) := \varphi^* c(E_k)$. (Later we use cohomology isomorphisms to express $c(F)$ as a differential form.)

We first show how to construct φ. Cover X by open sets U_i (we assume X is compact, so $i = 1, \ldots, N$) on which F is trivial, and find open sets V_i, W_i such that $\overline{V}_i \subset U_i$ and $\overline{W}_i \subset V_i$, as in Fig. 1. Then we may choose bump functions λ_i on X equal to 1 on W_i and falling off to zero outside V_i, as illustrated. Now say $p \in F$ sits over x, so $\pi(p) = x$. Local triviality tells us there is an isomorphism $\pi^{-1}(U_i) \xrightarrow{\sim} U_i \times \mathbb{C}^k$, and if we take the projection to $\mathbb{C}^k$ we get maps $\mu_i : \pi^{-1}(U_i) \to \mathbb{C}^k$, linear on each fiber. We then map p to $\nu(p)$:

$$\nu(p) \equiv (\lambda_1(x)\mu_1(p), \lambda_2(x)\mu_2(p), \ldots, \lambda_N(x)\mu_N(p)) \in \mathbb{C}^{kN} \subset \mathbb{C}^\infty.$$

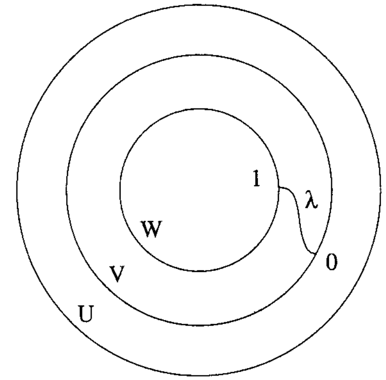


FIGURE 1. The open sets used to construct φ.

Each component makes sense even outside of the domains of μ_i, since the fall-off of the λ_i allows us smoothly to extend by zero. Note that $\nu(E_x)$ is a linear k-plane in $\mathbb{C}^{kN} \subset \mathbb{C}^\infty$, thus a point in G_k. Going from x to this point in G_k defines φ. Finally, we can map p to the pair (k-plane $\nu(E_x)$, $\nu(p)$)$\in G_k \times \mathbb{C}^\infty$. This map between total spaces of the bundles F and E, linear on the fibers, exhibits F as $\varphi^* E$.

The fact that any two such maps ν_0, ν_1 are homotopic comes from defining ν_t, $t \in [0,1]$ by linearly interpolating from $\nu_0(e)$ to $\nu_1(e)$ in $\mathbb{C}^\infty$. One needs to show that this can be done continuously and without hitting zero when e is nonzero.

Let $H^{-1} = \mathcal{O}(-1)$ be the universal line bundle over $\mathbb{CP}^\infty$. Then it turns out that the following axioms for Chern classes of rank k complex bundles $F \to X$ completely determine them:

- $c_i(F) \in H^{2i}(M, \mathbb{Z})$, $c_0(F) = 1$, $c_{i>k}(F) = 0$;
- $c(f^* F) = f^*(c(F))$;
- $c(F \oplus G) = c(F)c(G)$.
- $-c_1(H^{-1}) = e(H)$ is the generator of $H^2(G_k)$.

Topologically speaking, then, the set of Chern class of a given bundle determine the cohomology class of its classifying map φ, and so in simple cases determine the bundle, with its complex structure, up to homotopy (but not quite in general). Notice that knowing that two bundles are topologically isomorphic as complex $\mathbb{C}^k$-bundles does *not* mean that they are isomorphic as holomorphic bundles!

EXAMPLE 3.5.4. *Consider (complex) line bundles with vanishing Chern class on an elliptic curve $\mathbb{C}/\langle 1,\tau\rangle$. Any flat bundle has zero curvature, and therefore vanishing first Chern class. We can define a flat line bundle by specifying $U(1)$ holonomies around the two different cycles of the elliptic curve. Topologically, the space of such bundles forms a torus $S^1 \times S^1$. From a C^∞ point of view, all such bundles are homotopic, though they are different as holomorphic bundles. This can be seen by studying the kernel of the map from $H^1(\mathcal{O}^*)$ to $H^2(\mathbb{Z})$, whose image is the first Chern class. The kernel can be seen, from the long exact sequence of the exponential sequence, to be $H^1(\mathcal{O})/H^1(\mathbb{Z})$, which is $\mathbb{C}/\mathbb{Z}^2$.*

3.5.2. Chern Classes from Differential Geometry. To a physicist, the most "hands on" definition of a Chern class of a differentiable vector bundle is in terms of the curvature of a connection. While Chern classes can be defined in a more general context, the definition agrees with the definition given below when it is valid (when things are differentiable).

Let E be a differentiable complex vector bundle of rank r over a differentiable manifold M, and let $F = dA + A\wedge A$ be the curvature of a connection A on E. We define $c(E)$, the "total Chern class" of E, by

$$\begin{aligned} c(E) =& \det\left(1 + \frac{i}{2\pi}F\right) \\ =& 1 + \frac{i}{2\pi}\mathrm{Tr}F + \dots \\ =& 1 + c_1(E) + c_2(E) + \dots \in H^0(M,\mathbb{R}) \oplus H^2(M,\mathbb{R}) \oplus \dots . \end{aligned}$$

The form $c(E)$ is independent of the choice of trivialization (by conjugation invariance of the determinant) and is closed, by the Bianchi identity $DF = 0$. In fact, this definition is *independent of the choice of connection.* This follows (not immediately) from the fact that the difference of two connections is a well-defined $End(E)$-valued one-form. Different connections will yield different representatives of the cohomology classes c_k.

We see that the total Chern class is expressed in terms of the Chern classes $c_k(E) \in H^{2k}(M,\mathbb{R})$. Note that $c(E \oplus F) = c(E)c(F)$, which follows from properties of the determinant. In fact (though we will not prove it), if $0 \to A \to B \to C \to 0$ is a short exact sequence of sheaves, then $c(B) = c(A)c(C)$ (the "splitting principle").

EXAMPLE 3.5.5. *Let us compute the first Chern class of the line bundle defined by the $U(1)$ gauge field surrounding a magnetic monopole, integrated over the sphere at infinity. A magnetic monopole is a magnetic version of an electron, i.e., a source of divergence of magnetic (instead of electric) fields. We shall give the connection A explicitly. The curvature is just $F = dA$, since the $A\wedge A$ terms vanish for an abelian connection. (F is a combination of electric and magnetic fields, which can be determined by equating A_0 to the electric potential and $\vec{A}$ [the spatial components] to the magnetic vector potential, up to normalization constants.) The Dirac monopole centered at the origin of $\mathbb{R}^3$ is defined by*

$$A = i\frac{1}{2r}\frac{1}{z-r}(xdy - ydx).$$

One computes (**check**) $F = i\frac{1}{2r^3}(x\,dy\wedge dz + y\,dz\wedge dx + z\,dx\wedge dy)$. *In spherical coordinates, we can write $c_1 = \frac{i}{2\pi}F = \frac{1}{4\pi r^2}(r^2\sin\theta\, d\theta\wedge d\phi)$, and it is clear that the integral $\int_{S^2} c_1 = 1$ for any two-sphere around the origin.*

EXAMPLE 3.5.6. *Note that $\mathrm{Tr}F$ is the diagonal part of F, meaning it represents the $U(1) \subset GL(n,\mathbb{C})$ piece of the holonomy, at the level of Lie algebras. The first Chern class $c_1 = \frac{i}{2\pi}\mathrm{Tr}F$ is also the first Chern class of the determinant line bundle $\Lambda^r E$, which is evidenced by the fact that the trace measures how the logarithm of the determinant behaves under $GL(n,\mathbb{C})$ transformations. Therefore, if we are in a situation where the Levi–Civita connection on a complex manifold gives a connection on $T_{\mathrm{hol}}X$ and find that $c_1(T_{\mathrm{hol}}X) \equiv 0$* as a differential form, *then the holonomy must sit in $SU(n)$. Of course it is a necessary condition that $c_1 = 0$ as a cohomology class. Manifolds for which $c_1 = 0$ are called Calabi–Yau manifolds.*

The Chern Character. Suppose one defines x_i such that $c(E) = \prod_{i=1}^r (1 + x_i)$ (here $r \equiv rk(E)$). Then the Chern character class $ch(E)$ is defined by $ch(E) = \sum_i e^{x_i}$ (defined by expanding the exponential). Let us denote $c_k \equiv c_k(E)$. Then we find

$$ch(E) = r + c_1 + \frac{1}{2}(c_1^2 - 2c_2) + \frac{1}{6}(c_1^3 - 3c_1c_2 + 3c_3) + \dots .$$

Note $ch(E \oplus F) = ch(E) + ch(F)$ and $ch(E \otimes F) = ch(E)ch(F)$.

The Todd Class. With definitions as above, we define

$$td(E) = \prod_{i=1}^{r} \frac{x_i}{1-e^{-x_i}} = 1 + \frac{1}{2}c_1 + \frac{1}{12}(c_1^2 + c_2) + \frac{1}{24}c_1c_2 + \ldots.$$

Note that $td(E \oplus F) = td(E)td(F)$.

3.5.3. The Grothendieck–Riemann–Roch Formula. Very often, one wants to compute the dimension of a Čech cohomology group of some sheaf or vector bundle E over some variety X. These are typically difficult to count and may even jump in families. As an example, an elliptic curve has a family of holomorphic line bundles of degree 0, roughly parametrized by the dual elliptic curve or the Jacobian. However, only the trivial bundle $\mathcal{O}$ has a section (the constant function). A quantity that does not jump in families is the alternating sum $\chi(E) = \sum_k (-1)^k \dim H^k(E)$. The Grothendieck–Riemann–Roch formula calculates

$$\chi(E) = \int_X ch(E) \wedge td(X).$$

If we have other information telling us that some of the cohomology classes vanish or can otherwise determine their dimensions, the Grothendieck–Riemann–Roch theorem may suffice to determine the dimension of the desired cohomology group.

3.5.4. Serre Duality. One way of relating Čech classes among different sheaves is via Serre duality, which we motivate here but do not prove. If one recalls the Čech–Dolbeault isomorphism on a complex n-fold X, we can think of $H^k(E)$ as $H^k_{\bar\partial}(E)$. Therefore, there is a natural pairing $H^k(E) \otimes H^{n-k}(E^* \otimes K_M) \to \mathbb{C}$ defined by wedging together a $(0,k)$-form and a $(0, n-k)$-form and using the map $E \otimes E^* \to \mathbb{C}$, then combining with the canonical bundle to get an (n,n) form that is then integrated over X. Basically, then, Serre duality is just wedging and integrating. The statement is that this pairing is perfect, so

$$H^k(E) \cong H^{n-k}(E^* \otimes K_X)^*.$$

In the special case where X is Calabi–Yau, K_X is trivial and can be neglected in the formula above.

3.6. Some Practice Calculations

3.6.1. The Chern Class of $\mathbb{P}^n$ and the Euler Sequence. To compute the Chern class of $\mathbb{P}^n$ we recall that the homogeneous coordinates, being maps from the tautological bundle to $\mathbb{C}$, are sections of its dual, the hyperplane bundle H. To make a vector field invariant under $\mathbb{C}^*$, we can take $s_i \frac{\partial}{\partial X_i}$, where s_i, $i = 1, \ldots, n+1$ are any sections of H. We thus have a map from $H^{\oplus(n+1)}$ to $T\mathbb{P}^n$ (here $T\mathbb{P}^n$ represents the holomorphic tangent bundle), with the kernel sheaf being the trivial line bundle $\mathbb{C}$ of multiples of a nowhere-vanishing generator $(X_0, \ldots, X_n) \mapsto X_i \frac{\partial}{\partial X_i} \cong 0$ in $\mathbb{P}^n$. This is the Euler sequence:

$$0 \longrightarrow \mathbb{C} \longrightarrow H^{\oplus(n+1)} \longrightarrow T\mathbb{P}^n \longrightarrow 0.$$

Since $c(\mathbb{C}) = 1$, it follows from properties of the Chern class that $c(\mathbb{P}^n) \equiv c(T\mathbb{P}^n) = c(H^{\oplus(n+1)}) = [c(H)]^{n+1}$. Let $x = c_1(H)$. Then $c(\mathbb{P}^n) = (1+x)^{n+1}$.

Let us recall that a hyperplane represents the zeroes of a global section of the hyperplane bundle. In fact, this means that x is Poincaré dual to a hyperplane ($\cong \mathbb{P}^{n-1}$). It follows that $\int_{\mathbb{P}^n} x^n = 1$, since n hyperplanes intersect at a point (all hyperplanes are isomorphic, under $PGL(n+1)$, to setting one coordinate to zero). Further $\int_{\mathbb{P}^1 \subset \mathbb{P}^n} x = 1$, since a generic hyperplane intersects a $\mathbb{P}^1 \subset \mathbb{P}^n$ in a point. The Euler class of $\mathbb{P}^n$ is the top Chern class, so $c_n(\mathbb{P}^n) = \binom{n+1}{n} x^n$ and

$$\int_{\mathbb{P}^n} c_n(\mathbb{P}^n) = n+1.$$

This agrees with our previous observation that the Euler class or top Chern class (of a bundle of the same rank as the dimension of the manifold) counts the number of zeroes of a holomorphic section. The integral calculation above is also the Euler characteristic $\chi(\mathbb{P}^n) = n+1$.

3.6.2. Adjunction Formulas. Let X be a smooth hypersurface in $\mathbb{P}^n$ defined as the zero-locus of a degree d polynomial, p (so p is a section of $\mathcal{O}_{\mathbb{P}^n}(d)$, or H^d). Roughly speaking, since p serves as a coordinate near X, the normal bundle N_X of X in $\mathbb{P}^n$ is just $\mathcal{O}(d)|_X$. As a result, the exact sequence $0 \to TX \to T\mathbb{P}^n|_X \to N_X \to 0$ takes the form

$$0 \longrightarrow TX \to T\mathbb{P}^n|_X \longrightarrow \mathcal{O}(d)|_X \longrightarrow 0.$$

Now $ch(H) = e^x \Rightarrow ch(H^d) = e^{dx} = 1 + c_1(H^d) + \ldots$, so $c(\mathcal{O}(d)) = 1 + c_1 = 1 + dx$, and

$$c(X) = \frac{(1+x)^{n+1}}{(1+dx)}.$$

It is useful in what follows to note that the Euler class $e(X)$ of the normal bundle of a subvariety $X \subset \mathbb{P}^n$ is equal to its Thom class, namely its Poincaré dual cohomology cycle. This means $\int_X \theta = \int_{\mathbb{P}^n} \theta e(X)$. In the case of hypersurfaces, the normal bundle is one-dimensional and the Euler class (top Chern class) of the normal bundle is the first Chern class. In the case of $\mathcal{O}(d)$, the Poincaré dual class is the first Chern class dx (not "dx" the differential).

Curves in $\mathbb{P}^2$. A degree d curve X in $\mathbb{P}^2$ has Chern class $1 + (3-d)x$. Then $\chi(X) = \int_X c_1(X) = \int_{\mathbb{P}^2} c_1(X)(dx) = \int_{\mathbb{P}^2} d(3-d)x^2 = d(3-d)$. Setting the Euler characteristic $\chi(X) = 2 - 2g$, where g is the genus of the Riemann surface (number of handles), we find $g = (d-1)(d-2)/2 = \begin{pmatrix} d-1 \\ 2 \end{pmatrix}$.

The Quintic Hypersurface in $\mathbb{P}^4$. A quintic hypersurface Q in $\mathbb{P}^4$ has $c(Q) = (1+x)^5/(1+5x) = 1 + 10x^2 - 40x^3$ (recall $x^4 = 0$). Note that $c_1(Q) = 0$, so Q is a Calabi–Yau manifold. Its Euler characteristic is

$$\int_Q -40x^3 = \int_{\mathbb{P}^4} -40x^3(5x) = -200.$$

We saw in previous chapters that we could find 101 complex deformations, which (as we will see in later chapters) is the dimension of $H^1(TQ) \cong H^{2,1}(M)$, i.e., the Hodge number $h^{2,1} = h^{1,2}$. Since $h^{3,0} = h^{0,3} = 1$, the unique generators being the Calabi–Yau form and its complex conjugate, we learn that $b_3 = 204$ ($b_1 = b_5 = 0$ by simple connectivity). Now, since the Kähler form and its powers descend from $\mathbb{P}^n$ to a hypersurface, we have $h^{k,k}(Q) \geq 1$, and in fact there are no other forms ($h^{k,k} = 1$). For simply-connected Calabi-Yau's, $h^{1,0} = h^{2,0} = 0$, so the Hodge diamond has only $h^{1,1}$ and $h^{2,1}$ as undetermined, independent quantities. It is easy to see that $\chi(Q) = 2(h^{1,1} - h^{2,1})$, and we have found the compatible results that $h^{1,1} = 1$, $h^{2,1} = 101$, and $\chi(Q) = -200$.

3.6.3. The Moduli Space of Curves, $\mathcal{M}_{g,n}$.

(This section is only a prelude to the treatment given in Ch. 23.) A Riemann surface is a one-dimensional complex manifold, which means a differentiable, real two-dimensional manifold with choice of complex coordinates and holomorphic

transition functions across coordinate charts. The choice of holomorphic coordinate is often not unique, and the space of such choices (the moduli space of complex structures or "moduli space (of curves)") for a genus g closed surface is denoted $\mathcal{M}_g$. Infinitesimal changes of the complex structure (yet to be discussed) of a complex manifold X are classified by the Čech cohomology group $H^1(TX)$. This vector space therefore is the *tangent space* to $\mathcal{M}_g$ at the point defined by a genus g Riemann surface, X. We would like to compute the dimension of $\mathcal{M}_g = \dim_{\mathbb{C}} H^1(TX)$.

The Grothendieck–Riemann–Roch formula tells us

$$\begin{aligned} \dim_{\mathbb{C}} H^0(TX) - \dim_{\mathbb{C}} H^1(TX) &= \int_X ch(TX)td(TX) \\ &= \int_X (1 + c_1(TX))(1 + \tfrac{1}{2}c_1(TX)) \\ &= \tfrac{3}{2}\int_X c_1(TX) = 3 - 3g, \end{aligned}$$

where the last equality comes from the fact that the Euler class or top Chern class is just $c_1(TX)$ for a one-dimensional complex manifold. When $g \geq 2$ there are no nonzero vector fields of X and $H^0(TX) = 0$. We conclude that $\dim \mathcal{M}_g = 3g - 3$, $g \geq 2$. We commented on this fact when we discussed the constant curvature metric on the upper half-plane in the first chapter.

When $g = 1$, $H^0(TX) = \mathbb{C}$ and $\mathcal{M}_1$ is one-dimensional. The automorphism can be removed by selecting a distinguished point. When $g = 0$, $\dim H^0(TX) = 3$ (the generators of $PGL(2, \mathbb{C})$), and $\mathcal{M}_0$ is a point.

If we include n marked (ordered) points, we denote the space $\mathcal{M}_{g,n}$, and we require one additional complex dimension to describe the location of each marked point: $\dim \mathcal{M}_{g,n} = 3g - 3 + n$. When $g = 1$ and $n = 1$, the origin is marked as a distinguished point, and we have $\dim \mathcal{M}_{1,1} = 1$.

3.6.4. Holomorphic Maps into a Calabi–Yau.

An important space in mirror symmetry is the space of holomorphic maps from a Riemann surface Σ into a Calabi–Yau n-fold M (i.e., $c_1(M) = 0$). If $\phi : \Sigma \to M$ is a holomorphic map then, in local coordinates on M, ϕ obeys the equation $\overline{\partial}\phi^i = 0$. An infinitesimal deformation of ϕ can be generated by a vector field χ^i (think "$\phi \to \phi + \epsilon\chi$"), and the deformed map will still be holomorphic if $\overline{\partial}\chi^i = 0$. That is, χ defines an element of $H^0_{\overline{\partial}}(\phi^*TM) = H^0(\phi^*TM)$. ($\chi$ lives in ϕ^*TM since it need only be defined along the image curve.) We will assume here (not always justifiably) that $H^1(\phi^*TM) = 0$. Then

Grothendieck–Riemann–Roch gives

$$\begin{aligned}
\dim H^0(\phi^*TM) - \dim H^1(\phi^*TM) &= \int_\Sigma ch(\phi^*TM)td(\Sigma) \\
&= \int_\Sigma (n + \phi^*c_1(TM))(1 + \tfrac{1}{2}c_1(\Sigma)) \\
&= n(1-g),
\end{aligned}$$

where at the end we use $c_1(TM) = 0$ (Calabi–Yau). Note that the result is independent of the homology class of the image. Also note that when $n = 3$ and $g = 0$ the dimension is 3, which is also the dimension of the automorphism group of a genus 0 Riemann surface ($\mathbb{P}^1$). The automorphisms of the domain change the map pointwise, but do not move the image curve. Therefore the dimension of the genus 0 holomorphic curves inside a Calabi–Yau threefold is *zero*, so we may expect to be able to count them! Mirror symmetry will have a lot more to say on this subject.

Note that if M is not a Calabi–Yau manifold, the calculation holds up until the last line, and the index formula yields $n(1-g) + \int_\Sigma \phi^*c_1(TM)$. The second term is the pairing of the homology class $\phi(\Sigma)$ with $c_1(TM)$ and is the "degree" of the image.

CHAPTER 4

Equivariant Cohomology and Fixed-Point Theorems

Certain characteristic classes of bundles over manifolds are very simple to compute when the manifold and bundle carry an action of a group. This chapter contains a synopsis of various theorems concerning the localization of calculations to fixed points of diffeomorphisms, zeroes of vector fields or sections, or fixed points of group actions. (Some of these topics appear scattered in other chapters.) We try to motivate the results, but will not prove the theorems. The main example, Sec. 4.4, highlights our reason for exploring the subject: to calculate Gromov–Witten invariants.

In fact, we saw in Sec. 3.5 that the zeroes of a holomorphic vector field give the Euler class of a manifold and that the zeroes of holomorphic sections give Chern classes of vector bundles. In the case where we have a holomorphic S^1 (or $\mathbb{C}^*$) action on a manifold, the generator is a holomorphic vector field and its zeroes correspond to fixed points of the group action. Therefore, it is reasonable to expect that certain characteristic classes of bundles with group actions can be localized to the fixed-point sets of these actions. Given a bundle over the manifold, one can often lift the group action equivariantly to the total space of the bundle (so that it covers the original action); such a lift is automatic for the tangent bundle and other natural bundles on a manifold. The proper integrands to consider will turn out to be "equivariant cohomology classes."

For simplicity, we shall only consider actions by products of S^1 or $\mathbb{C}^*$.

4.1. A Brief Discussion of Fixed-Point Formulas

In our interpretation of Chern classes in Sec. 3.5, we saw that the zeroes of sections contain important topological, intersection-theoretic data. This allowed us to state generalizations of the Gauss–Bonnet theorem for surfaces.

Similarly, the fixed-point set F of an endomorphism $f : M \to M$ contains topological data defined by f, as F can be recast as the intersection in $M \times M$ of the diagonal with the graph of f, $\Gamma_f = \{(m, f(m)) \in M \times M\}$.

The two discussions merge when f is generated by a vector field, v. Then f is homotopic to the identity, and the intersection calculation gives the self-intersection of M in $M \times M$, that is, the Euler characteristic, $\chi(M)$. The formula, the Hopf index theorem, is

$$\chi(M) = \sum_{v(p)=0} \operatorname{sgn} \det \left(\frac{\partial v^i}{\partial x^j} - \delta^i{}_j \right).$$

Here $\frac{\partial v^i}{\partial x^j}$ is the explicit expression for the action of f_* on TM at a zero of v.

More generally, even if f is not generated by a vector field or homotopic to the identity, then f^* acts on cohomology and the Lefschetz fixed point theorem, which has a form similar to the equation above, gives the (signed) trace of the action of f^* on $H^*(M)$. (In the above, $f^* = \mathrm{id}$ and we get $\chi(M)$ from the trace.) These statements have refinements when M is a complex manifold and f is holomorphic, so that f^* can also act on Dolbeault cohomology.

Bott extended this kind of reasoning to a holomorphic vector field v acting on a manifold M with a holomorphic vector bundle $E \to M$. After assuming a lift of the action of v on functions to an action on sections of E, one is able to write characteristic classes of E as exact forms *outside the zero set of* v. The construction depends on a dual one-form to v, which only exists when $v \neq 0$. A unified understanding of these techniques led to the Atiyah–Bott fixed-point theorem, to which we will turn after discussing the necessary prerequisites.

4.2. Classifying Spaces, Group Cohomology, and Equivariant Cohomology

Equivariant cohomology is a way of capturing the topological data of a manifold with a group action in such a way that it enjoys the usual cohomological properties under pull-back and push-forward. (This is called "functoriality.")

EXAMPLE 4.2.1. *If M is a smooth manifold and G a group acting smoothly without fixed points on M, then M/G is a smooth manifold, and equivariant cohomology will be defined to agree with $H^*(M/G)$. However, if G has various fixed points with different stabilizers (subgroups leaving a point fixed), then we want the equivariant cohomology to "see" these stabilizers. This is demonstrated most sharply when M is a point $\{pt\}$ and any G action fixes pt. Then M/G is always a point, but the equivariant cohomology of a point should depend on which group is acting, and should give a cohomological invariant of the group. If $H \subset G$ is a subgroup, we should also have a pull-back map onto the H-equivariant cohomology of a point.*

The considerations above lead to the following definition of the equivariant cohomology, $H^*_G(M)$. First let us warm up with $M = \{pt\}$. We will define $H^*_G(\{pt\})$, also denoted $H^*(G)$ or H^*_G, to be $H^*(BG)$, where BG is the "classifying space of G." BG is defined by finding a contractible space EG — unique, up to homotopy — on which G acts freely (without fixed points) on the right, and setting

$$BG = EG/G. \tag{4.1}$$

When $M = \{pt\}$, $H^*(G)$ is also called the "group cohomology." Cohomology will be taken with coefficients in $\mathbb{Q}$.

EXAMPLE 4.2.2. *This definition comes to life in examples. If G is a finite group, then BG has fundamental group G and no other non-trivial homotopy groups. If $G = \mathbb{Z}$, then $EG = \mathbb{R}$ and $BG = EG/G = \mathbb{R}/\mathbb{Z} = S^1$. Note that $\pi_1(S^1) = \mathbb{Z}$.*

If $G = S^1$, then G is continuous and our intuition might lead us astray. We note, however, that $\mathbb{CP}^n$ is the quotient of S^{2n-1} by S^1. (This can be seen by taking the usual $\mathbb{C}^n \setminus 0$ and quotienting by $\mathbb{C}^*$ in two stages, first using the $\mathbb{R}$ freedom to solve $|z| = 1$, then quotienting by S^1.) If we blithely take the limit $n \to \infty$, then S^n becomes "contractible" and BS^1 is the quotient, $\mathbb{CP}^\infty$. Therefore,

$$H^*_{S^1}(\{pt\}) = \mathbb{Q}[t],$$

the polynomial algebra in one variable. For multiple S^1 or $\mathbb{C}^*$ actions, we get the polynomial algebra in several variables, so if $\mathbb{T} = (\mathbb{C}^*)^{m+1}$, then

$$H^*_{\mathbb{T}} = H^*((\mathbb{CP}^\infty)^{m+1}) = \mathbb{Q}[t_0, \ldots, t_m].$$

This will serve as our main example throughout this chapter.

In the example above, the "indeterminate" t is actually the generator of $H^*(\mathbb{CP}^\infty)$ and can be thought of as the first Chern class of the hyperplane line bundle, dual to the tautological line bundle over $\mathbb{CP}^\infty$. Equivalently, in what follows we can think of t as a complex indeterminate and use the group $\mathbb{C}^*$ instead of S^1. However, we will continue to consider S^1 in our discussion.

EXAMPLE 4.2.3. *Classifying spaces can be used to study isomorphism classes of bundles over compact spaces. A bundle of rank k is defined by giving a k-dimensional vector space at every point in M, that is, a point in the (infinite) real Grassmannian G_k of k-planes. One checks, as in Sec. 3.5.1, that isomorphic bundles give homotopic maps, and that* any *bundle can be pulled back from such a map. Therefore, isomorphism classes of bundles over* any *space are given by homotopy classes of maps into G_k. But G_k is precisely the classifying space of the structure group $GL(k)$, or equivalently $O(k)$. Complex bundles are classified by maps into the classifying space of $GL(k;\mathbb{C})$, i.e., the complex Grassmannian.*

Now note that G acts on EG on the right and M on the left, so we can set

$$M_G = EG \times_G M,$$

i.e., $(eg, m) \sim (e, gm)$. This space has some nice properties. M_G fibers over M/G with fiber over $[Gm]$ equal to $EG/\{g|gm = m\}$, which is itself BG_m, with $G_m \equiv \{g|gm = m\}$ the stabilizer of m. Therefore, if G_m is trivial for all M, then M_G is homotopic to M/G, as desired (they have the same cohomology). We define equivariant cohomology by the ordinary cohomology of M_G.

DEFINITION 4.2.4.

$$H_G^*(M) \equiv H^*(M_G).$$

Note that sending $(e, m) \mapsto e$ gives a map from M_G to BG with fiber M. The inclusion $M \hookrightarrow M_G$ as a fiber gives a map $H_G^*(M) \to H^*(M)$ by pull-back. We also have an equivariant map $M \to \{pt\}$, which gives $H_G^*(M)$ the structure of an $H^*(G)$ module. Equivariant cohomology classes pulled back from $H^*(G)$ are said to be "pure weight." In the case that $G = S^1$, we can think of $H_G^*(M)$ loosely, then, as being constructed out of polynomial-valued differential forms. (Soon we will allow denominators in these polynomials – this is called "localizing the ring.")

EXERCISE 4.2.1.

(a) *Show that if G acts trivially on M, then $H_G^*(M) = H^*(M) \times H_G^*$.*

(b) *Show that if G acts freely on M, then $H_G^*(M) = H^*(M/G)$, and is a torsion H_G^*-module. (For example, if $G = T$, then t_i acts on $H_G^*(M)$ by multiplication by 0.)*

As we will see in the next section, the essential insight of localization is that the non-torsion part of $H_G^*(M)$ is contributed by the G-fixed part of M. The proof involves little more than the previous exercise: Stratify M by the stabilizer type of points, apply the exercise to each stratum, and glue them together using Mayer–Vietoris.

When G is the torus $\mathbb{T}$, let $F \subset M$ be its fixed locus. A basic result in the subject is the following: If M is non-singular, then F is also non-singular. The vector bundle $T_M|_F$ on F carries a natural $\mathbb{T}$-action. The "fixed" part of the bundle (where the torus acts trivially, that is, with weight zero) is T_F, and the "moving" part of the bundle (where the torus acts non-trivially) is the normal bundle $N_{F/M}$. The inclusion $F \hookrightarrow M$ induces

$$H_{\mathbb{T}}^*(M) \to H_{\mathbb{T}}^*(F) = H^*(F) \otimes_{\mathbb{Q}} H_{\mathbb{T}}^*(pt) = H^*(F)[t_0, \ldots, t_m].$$

THEOREM 4.2.5 (Localization). *This is an isomorphism up to torsion (that is, an isomorphism once tensored with $\mathbb{Q}(\alpha_0, \ldots, \alpha_m)$).*

Note that the tensoring simply allows coefficients rational in the t_i. The localization theorem of the next section tells precisely which class in $H_{\mathbb{T}}^*(F)$ corresponds to a class $\phi \in H_{\mathbb{T}}^*(M)$.

4.2.1. De Rham Model. Not only can equivariant cohomology classes in this case be thought of as polynomial-valued (or rational-function-valued) differential forms, one can exploit this fact to build an explicit and simple de Rham-type construction for computing equivariant cohomology classes! Let X be a vector field generating the S^1 action. Let $i(X)$ denote the inner product by X and define $d_X = d + ui(X)$ acting on $\Omega^*(M)[u]$, with u an indeterminate to which we assign degree 2. Note that $d_X^2 \neq 0$. In fact, we must restrict ourselves to X-invariant forms, i.e., forms in the kernel of the Lie derivative $\mathcal{L}_X = di(X) + i(X)d$. Denoting this space of forms by $\Omega_X^*(M)$, we see that $d_X^2 = 0$ on $\Omega_X^*(M)[u]$, and in fact

$$H_{S^1}^*(M) = \operatorname{Ker} d_X / \operatorname{Im} d_X.$$

Therefore, ordinary closed differential forms that are killed by $i(X)$ represent equivariant cohomology classes. Even those that are not may have equivariant extensions involving cohomology classes of lower degrees (but higher powers of u).

4.3. The Atiyah–Bott Localization Formula

If $i : V \hookrightarrow M$ is a map of compact manifolds, then we can push forward cohomology (one can use Poincaré duality on V, push forward the homology cycle, then use Poincaré duality again on M), giving a map $i_* : H^*(V) \to H^{*+k}(M)$, where k is the codimension of V. This map makes sense even if i is not an inclusion. In this case k can be negative – e.g., if the map i is a fibering, then i_* is integration over the fibers.

A tubular neighborhood of V inside M can be identified with the normal bundle of V. On the total space of the normal bundle lives the form with compact support in the fibers that integrates to one in each fiber: the Thom form, Φ_V. Clearly, the degree of this form is equal to the codimension of V. Extending this form by zero gives a form in M, and in fact multiplying by Φ_V provides an isomorphism $H^*(V) \cong H^{*+k}(M, M \setminus V)$, which then maps to $H^{*+k}(M)$. As a result, we see that the cohomology class $1 \in H^0(V)$ is sent to the Thom class in $H^k(M)$ coming from the normal bundle of V. This class restricts (to V by pull-back under inclusion, i^*) to be the Euler class $\mathbf{e}$ of the normal bundle of V in M, $N_{V/M}$. Therefore, we see that

$$i^* i_* 1 = \mathbf{e}(N_{V/M}). \tag{4.2}$$

This natural structure can be shown to hold in equivariant cohomology by applying the same argument to the appropriate spaces M_G, V_G, etc.

What makes the localization theorem possible is the ability to invert the Euler class in equivariant cohomology. Normally, of course, one cannot invert a top form, as there is no form that would give the zero form, 1, as the result of wedging.

For example, suppose that V above is a point. Then the formula of Eq. (4.2) says that pushing forward and pulling back 1 gives the Euler characteristic of the tangent space of M at V. Of course, this space is a trivial bundle, but if M carries a group action and V is a fixed point of this action, then the tangent space at V is an equivariant vector space, which splits into a sum of non-trivial irreducible representations, V_i. Note that we need V to be a fixed point in order to have such a structure. When G is S^1 the irreducible representations are two-dimensional and labeled by real numbers

a_i (or more precisely exponentials of the dual of the Lie algebra of S^1 – and we can take complex coefficients, if we like). Then $\mathbf{e}(N_{V/M}) = \prod_i a_i$, which is invertible if we allow denominators, i.e., if we work over rational functions instead of polynomials. (We need not extend our ring all the way to rational numbers, but our discussion is rather coarse.)

The theorem of Atiyah and Bott says that such an inverse of the Euler class of the normal bundle always exists along the fixed locus of a group action. In such a case, $i^*/\mathbf{e}(N_{V/M})$ will be inverse to i_* in equivariant cohomology (not just for 1 but for any equivariant cohomology class). Let F run over the fixed locus. Then, for any equivariant class ϕ,

$$\phi = \sum_F \frac{i_* i^* \phi}{\mathbf{e}(N_{F/M})}. \tag{4.3}$$

We noted that pushing forward was accomplished through Poincaré duality, so for the map $\pi^M : M \to \{pt\}$, pushing forward is the same (for non-equivariant classes) as integration over M. Note, too, that the map from F to $\{pt\}$ factors through the map to M : $\pi_*^F i_* = \pi_*^F$. Applying π_*^M to (4.3) then gives the integrated version of the localization formula:

$$\int_M \phi = \sum_F \int_F \frac{i^* \phi}{\mathbf{e}(N_{F/M})}. \tag{4.4}$$

What makes this formula useful is that, as we have seen, computations in equivariant cohomology – at least for $G = S^1$ – are easy to carry out explicitly.

As an example, we prove that if M has a finite number n of fixed points under $\mathbb{T}$, then $\chi(M) = n$. Note that $\chi(M) = \mathbf{e}(T_M)$. There is a natural $\mathbb{T}$-action on T_M, inducing a bundle on $M_{\mathbb{T}}$, which we also call T_M by abuse of notation (adding the adjective "equivariant" to indicate that we are working on $M_{\mathbb{T}}$). By the localization formula (4.4),

$$\int_M \mathbf{e}(T_M) = \sum_F \int_F \left(\frac{\mathbf{e}(N_{F/M})}{\mathbf{e}(N_{F/M})} \right) = \sum_{j=1}^n 1 = n.$$

Exercise 4.3.1. *Find the Euler characteristic of (a) $\mathbb{P}^m$, (b) the Grassmannian of k-planes in $\mathbb{C}^m$, and (c) the flag manifold parametrizing complete flags in $\mathbb{C}^m$.*

In this case, where there are a finite number of fixed points, we are even given a cell decomposition, as follows. Take $\alpha_i = i\alpha_1$ (so now the torus acting is one-dimensional, with coordinate t, say). Then associate to the fixed point F_j the set of points p whose limit under the torus action is F_j: $\lim_{t\to\infty} tp = F_j$. Each cell has even (real) dimension (and is isomorphic to some $\mathbb{C}^k$), so M has no odd cohomology, and the ith Betti number of M is just the number of i-cells in the stratification.

EXERCISE 4.3.2. *Check that the Betti numbers of projective space are what you would expect, and describe these "Schubert cells". Find the Betti numbers of the Grassmannian parametrizing planes in* $\mathbb{C}^4$ *and describe its Schubert cells.*

4.4. Main Example

The main purpose for introducing equivariant cohomology and the localization theorem in this text is that computations of Gromov–Witten invariants are often done in the toric setting, where S^1 actions abound and the localization formula is the main computational tool.

Here we outline the approach about which we will learn much more in Part 4, especially Ch. 27. We focus on the genus 0 case.

We saw that Calabi–Yau manifolds can be described as hypersurfaces or complete intersections of hypersurfaces in ambient toric varieties. The simplest example is the quintic, the threefold described by a homogeneous polynomial s of degree 5 in $\mathbb{P}^4$. There is a nice space of "stable" holomorphic maps from genus 0 curves to $\mathbb{P}^4$. "Stability" is a technical term which we will not go into now (stable maps will be discused in detail in Ch. 24), but an open set inside this space of stable maps looks exactly like what you might expect: The genus 0 curve is $\mathbb{P}^1$ with coordinates U and V and maps of degree d look like five-tuples of degree-d polynomials in U and V. One must quotient this space by automorphisms of the source curve, that is, five-tuples of polynomials related by $PSL(2;\mathbb{C})$ transformations on U and V should be equated.

Now S^1 (or $\mathbb{C}^*$) can act on $\mathbb{P}^4$ in a number of ways, and we consider an action defined by weights $\lambda_1, \ldots, \lambda_5$, with $\mu \in \mathbb{C}^*$ acting by $(X_1, \ldots, X_5) \mapsto (\mu^{\lambda_1}X_1, \ldots, \mu^{\lambda_5}X_5)$. Then $\mathbb{C}^*$ also acts on our space of maps to $\mathbb{P}^4$ by composing the map with this action. The fixed points of this $\mathbb{C}^*$ action are maps that send $\mathbb{P}^1$ to an invariant $\mathbb{P}^1 \subset \mathbb{P}^4$ such that the action on the invariant $\mathbb{P}^1$ can be "undone" by a $PSL(2;\mathbb{C})$ transformation of U and V. An example of such a map, say of degree d, is $(U,V) \mapsto (U^d, V^d, 0, 0, 0)$. Note that if we

assign U and V the weights λ_1/d and λ_2/d, then this map is equivariant under $\mathbb{C}^*$, hence represents a fixed point. In the space of maps of degree d, for each pair (i,j), $1 \le i < j \le 5$, there is a fixed point of the type described above. There are other fixed-point loci corresponding to holomorphic maps from a genus 0 curve that consists of two $\mathbb{P}^1$'s meeting in a point (node), mapped as above with degrees d_1 and d_2 such that $d_1 + d_2 = d$. There are other types of degenerations as well, but we leave such discussions to Chs. 23 and 24.

What kind of computation on this space concerns us? We know that the quintic polynomial s defining the Calabi–Yau is a global section of $\mathcal{O}_{\mathbb{P}^4}(5)$. In fact, the bundle $\mathcal{O}_{\mathbb{P}^4}(5)$ can be used to define a bundle E over the space of stable maps of degree d, say, where the fiber over a map is the space of global sections of $\mathcal{O}_{\mathbb{P}^4}(5)$ pulled back to the genus 0 curve, C, by the map $f: C \to \mathbb{P}^4$. Since s is a global section of $\mathcal{O}_{\mathbb{P}^4}(5)$, it certainly pulls back to a global section of $f^*\mathcal{O}_{\mathbb{P}^4}(5)$. Therefore we induce a natural section $\widetilde{s}$ of E.

EXERCISE 4.4.1. *Using the G–R–R theorem (Sec. 3.5.3), calculate the rank of the bundle E. Show that it is equal to the dimension of the space of quintuples of polynomials discussed, taking into account* $PSL(2;\mathbb{C})$ *equivalence.*

A zero of $\widetilde{s}$ looks like a map $f: C \to \mathbb{P}^4$ whose image is wholly contained in the zero set of s. But this zero set is precisely the quintic Calabi–Yau threefold, so zeroes of $\widetilde{s}$ count maps to the Calabi–Yau! Therefore, we want to count the zeroes of the section $\widetilde{s}$. We know from our discussion of Chern classes that the number of zeroes of a section of a bundle whose rank is equal to the dimension of a manifold (see Exercise 4.4.1) gives the Euler class of the bundle. Therefore, we want to calculate the Euler class of E, and the Atiyah–Bott theorem is just what we need.

4.4.1. A Note on Excess Intersection. A subtlety arises due to the fact that $\widetilde{s}$ is not only zero when our holomorphic map is an embedding into a degree d curve in the quintic. Indeed, a degree d map from $\mathbb{P}^1$ to the quintic can be a composition of a degree d/k map to the quintic with a k-fold cover of $\mathbb{P}^1$ by $\mathbb{P}^1$ (when k divides d): the image will still lie in the zero set of s. Such contributions can be accounted for through "multiple cover formulas." These multiple cover formulas concern the case where a section has more zeroes than expected. In our example, $\widetilde{s}$ is a section of a bundle whose rank equals the dimension of the manifold, but as there are many such self-covers of $\mathbb{P}^1$ by $\mathbb{P}^1$, we see that $\widetilde{s}$ has a non-isolated zero set, larger than expected. The excess intersection formula (Theorem 26.1.2) accounts for such a non-generic situation. For further details, see the discussion following the statement of that theorem.

Readers noting that the formula looks a lot like the localization formula will be assured that the reason is again that we are interested in representing a class on M by a class on a submanifold (the zero set of a section).

CHAPTER 5

Complex and Kähler Geometry

In this chapter we discuss the basics of complex geometry and Kähler metrics, which play an important role in string theory. As we will see in Ch. 13, manifolds with Kähler metric admit the $N = 2$ supersymmetric sigma models crucial for formulating mirror symmetry. We also discuss the Calabi–Yau condition.

5.1. Introduction

Here we review the basics of complex geometry. We will focus on Kähler metrics, i.e., those for which the parallel transport of a holomorphic vector remains holomorphic. This property means that the connection splits into holomorphic plus anti-holomorphic connections on those two summands in the decomposition of the tangent bundle.

Another consequence of this property is that the metric (in complex coordinates) is a Hermitian matrix at every point, and is completely determined in a neighborhood by a (non-holomorphic) function, Φ, called the Kähler potential. In fact, Φ is not uniquely defined, and corresponds to a section of a line bundle.

Yet another hallmark of Kähler geometry is a closed two-form ω determined by the metric, or equivalently its Kähler potential. This "Kähler form" is non-degenerate, and from its definition can be seen to satisfy $\omega^n/n! = dV$. Also, $\omega^k/k!$ has the property that it restricts to the induced volume form on any holomorphic submanifold of dimension k ($k = n$ was just noted). Since ω is a closed two-form on our space X, its cohomology class is determined by its values on $H_2(X, \mathbb{Z})$, namely the real numbers $t_i = \int_{C_i} \omega$, where C_i are a basis for $H_2(X, \mathbb{Z})$. The t_i are called "Kähler parameters."

5.2. Complex Structure

We have already defined a complex n-manifold as a topological space covered by charts isomorphic to open sets in $\mathbb{C}^n$, with holomorphic transition functions. Given a real $2n$-manifold, one might ask when it can be endowed with coordinates and transition functions satisfying the requirements of a complex manifold, and, if so, is this choice unique?

The differential at some point of a path in $\mathbb{C}^n$ has a real and an imaginary part, and multiplication by $i = \sqrt{-1}$ sends $dx \mapsto -dy$ and $dy \mapsto dx$, where x and y are the real and imaginary parts. Such a structure must exist for any manifold that might be a complex manifold. An "almost complex structure" J is a map on tangent spaces that squares to $-\mathbf{1}$: that is, $J \in End(T)$, $J^2 = -\mathbf{1}$. With an almost complex structure, we have a pointwise notion of holomorphic and anti-holomorphic tangent vectors (with complex values), depending on whether the eigenvalue under J is $\pm i$. In local (real) coordinates we can write J in terms of a matrix $J^a{}_b$, where $J(\frac{\partial}{\partial x^a}) = J^c{}_a \frac{\partial}{\partial x^c}$.

The theorem of Newlander and Nirenberg makes the following argument. If the Lie bracket[1] of two holomorphic vectors is always a holomorphic vector ("integrability"), then coordinates can be found whose derivatives are always holomorphic, i.e., we can find suitable complex coordinates. (Clearly, since the Lie bracket of coordinate vectors vanishes, the integrability condition is necessary.) Since $P = (1 - iJ)/2$ is a projection onto the holomorphic sub-bundle of the tangent bundle (tensored with $\mathbb{C}$) and $\overline{P} = (1 + iJ)/2$ is the anti-holomorphic projection, the condition of integrability for finding complex coordinates is

$$\overline{P}[PX, PY] = 0.$$

EXERCISE 5.2.1. *Define the Nijenhuis tensor by* $N(X,Y) = [JX, JY] - J[X, JY] - J[JX, Y] - [X, Y]$. *Given two vector fields, N returns a vector. Show that in local coordinates x^a,* $N^a{}_{bc} = J^d{}_b(\partial_d J^a{}_c - \partial_c J^a{}_d) - J^d{}_c(\partial_d J^a{}_b - \partial_b J^a{}_d)$. *Show that the integrability condition is equivalent to $N \equiv 0$. It is also equivalent to $\overline{\partial}^2 = 0$, where $\overline{\partial}$ is the part of d which adds one anti-holomorphic form degree (see below). Hint: Use the relation you get from* $J^2 = -\mathbf{1}$, *i.e.,* $\partial_a(J^b{}_c J^c{}_e) = 0$.

[1] The Lie bracket $[X, Y]$ of two vector fields, $X = X^a \frac{\partial}{\partial x^a}$ and $Y = Y^b \frac{\partial}{\partial x^b}$, is the "commutator" $(X^a \partial_a Y^b - Y^a \partial_a X^b)\frac{\partial}{\partial x^b}$, where $\partial_a = \frac{\partial}{\partial x^a}$, etc.

Of course, the eigenvalues of J are $\pm i$, but we can only find eigenvectors if we complexify our space, so we work with $TM \otimes \mathbb{C}$. Let $T'M$ and $T''M$ represent the eigenspaces with respective eigenvalues $+i$ and $-i$.[2] We call $T'M$ the *holomorphic tangent bundle*, and $T''M$ the *anti-holomorphic tangent bundle*. If $z^k = x^k + iy^k$ are holomorphic coordinates, then $\partial_k \equiv \frac{\partial}{\partial z^k} = \frac{1}{2}(\frac{\partial}{\partial x^k} - i\frac{\partial}{\partial y^k})$ generate $T'M$ and $\partial_{\bar{k}} \equiv \frac{\partial}{\partial \bar{z}^k} = \frac{1}{2}(\frac{\partial}{\partial x^k} + i\frac{\partial}{\partial y^k})$ generate $T''M$. If the context is clear, we sometimes abuse notation and write $TM \otimes \mathbb{C} = TM \oplus \overline{T}M$, i.e., we write TM for $T'M$ and $\overline{T}M$ for $T''M$. This is because the real TM, with its complex structure J, is isomorphic as a complex vector bundle to $T'M$ (whose complex structure is by multiplication by i) via $v \mapsto \frac{1}{2}(v - iJv)$. Similarly, therefore, T^*M represents the *holomorphic* cotangent bundle.

The decomposition into holomorphic and anti-holomorphic pieces carries through to cotangent vectors and p-forms in general. Thus, a (p,q)-form θ is a complex-valued differential form with p holomorphic pieces and q anti-holomorphic pieces, i.e., $\theta \in \Gamma(\Lambda^p T^*M \otimes \Lambda^q \overline{T^*}M)$. We can write $\theta = \theta_{a_1 \ldots a_p \bar{b}_1 \ldots \bar{b}_q} dz^{a_1} \ldots dz^{a_p} d\bar{z}^{b_1} \ldots d\bar{z}^{b_q}$. The functions $\theta_{a_1 \ldots a_p \bar{b}_1 \ldots \bar{b}_q}$ are in general neither holomorphic nor anti-holomorphic. Note that this decomposition can be written

$$\Omega^n(M) = \bigoplus_{p+q=n} \Lambda^p T^*M \otimes \Lambda^q \overline{T^*}M = \bigoplus_{p+q=n} \Omega^{p,q}(M),$$

where we have defined $\Omega^{p,q}(M)$ as the (p,q)-forms.

On a complex manifold, the operator $d : \Omega^p \to \Omega^{p+1}$ has a decomposition as well:

$$d = \partial + \overline{\partial},$$

where

$$\partial : \Omega^{p,q}(M) \to \Omega^{p+1,q}(M), \qquad \overline{\partial} : \Omega^{p,q}(M) \to \Omega^{p,q+1}(M)$$

are defined by $\partial\theta = \sum_k \partial_{z^k} \theta_{I,\bar{J}} dz^k \wedge dz^I d\bar{z}^{\bar{J}}$, if $\theta = \theta_{I,\bar{J}} dz^I d\bar{z}^{\bar{J}}$ is a (p,q)-form.[3] $\overline{\partial}$ is defined similarly. Then matching form degrees in $d^2 = 0$ gives

$$\partial^2 = 0, \qquad \overline{\partial}^2 = 0, \qquad \partial\overline{\partial} + \overline{\partial}\partial = 0.$$

In particular, we can define $H^{p,q}_{\overline{\partial}}(M)$ as those (p,q)-forms which are killed by $\overline{\partial}$ modulo those which are $\overline{\partial}$ of a $(p, q-1)$-form. The Čech–Dolbeault

[2] We have also called these T_{hol} and $T_{\mathrm{anti-hol}}$.

[3] We have used multi-indices. Here, for example, I represents a p-element subset of $\{1, \ldots, n\}$ and $dz^I = dz^{i_1} \wedge \cdots \wedge dz^{i_p}$.

isomorphism says $H^{p,q}_{\overline{\partial}}(M) \cong H^q(\Lambda^p T^*M)$. On an almost complex manifold, $d = \partial + \overline{\partial} + \ldots$, where, on (p,q)-forms, say, $\overline{\partial}$ is the projection of d onto $(p, q+1)$-forms. The integrability condition of Exercise 5.2.1 is equivalent to $\overline{\partial}^2 = 0$.

EXERCISE 5.2.2. *Show that on the complex plane* $\partial\overline{\partial}f = (i/2)\Delta f\, dV$, *where* Δ *is the flat Laplacian* $-(\partial_x^2 + \partial_y^2)$ *and* dV *is the volume form.*

The operators J and $\overline{\partial}$ are related. In the next chapter, we will see that deformations of the complex structure can be phrased in terms of deformations of either of these operators that preserve the defining properties.

5.2.1. Hermitian Metrics and Connections. A Hermitian metric is a positive-definite inner product $TM \otimes \overline{T}M \to \mathbb{C}$ at every point of a complex manifold M. In local coordinates z^i we can write $g_{i\bar{j}}dz^i d\bar{z}^{\bar{j}}$. Then $g_{i\bar{j}}(z)$ is a Hermitian matrix for all z.

As a real manifold with complex structure $J : T_{\mathbb{R}}M \to T_{\mathbb{R}}M$, the Hermitian condition is

$$g(X,Y) = g(JX, JY).$$

In terms of the components $J_m{}^n$, this condition says that $J_{ab} = -J_{ba}$, where $J_{ab} = J_a{}^c g_{cb}$. Therefore, we can define a two-form $\omega = \frac{1}{2}J_{ab}dx^a \wedge dx^b$.

In complex coordinates, this can be written $\omega = i g_{i\bar{j}}dz^i \wedge d\bar{z}^{\bar{j}}$. More invariantly, we can write the action of ω on vectors as

$$\omega(X,Y) = g(X, JY).$$

Now consider a rank r *complex* vector bundle with metric h_{ab}, $a, b = 1, \ldots, r$. The metric is said to be Hermitian if $h_{ab}(x)$ is Hermitian for all x.

Any Hermitian metric on a *holomorphic* vector bundle defines a Hermitian connection as follows. In a local frame with sections $e_a(x)$ generating the fibers and metric $h_{ab}(x)$, let z^k be local coordinates. Then we take the connection one-form to be

$$A_k = (\partial_k h)h^{-1}, \qquad A_{\bar{k}} = 0.$$

This can be shown to be the unique connection compatible with the Hermitian metric (like the Levi–Civita connection for the real tangent bundle) and trivial in the anti-holomorphic directions (this means it is compatible with the complex structure). A Kähler metric is a Hermitian metric on the

tangent bundle for which the holomorphic part of the Levi–Civita connection agrees with the Hermitian connection. We now turn to the study of these metrics.

5.3. Kähler Metrics

As we just saw, the data of a Hermitian metric allow us to define a $(1,1)$-form $\omega = \frac{i}{2}g_{i\bar{j}}dz^i \wedge d\bar{z}^{\bar{j}}$. We say the metric is "Kähler" if $d\omega = 0$.

EXERCISE 5.3.1. *Show that Kählerity is equivalent (in a coordinate patch) to* $\partial_i g_{j\bar{k}} = \partial_j g_{i\bar{k}}$. *Compute the Levi–Civita connection for a Kähler manifold and show that it has pure indices, either all holomorphic or all anti-holomorphic. Show that its holomorphic piece agrees with the unique Hermitian connection on the tangent bundle compatible with the complex structure, as claimed above.*

An important consequence of Kählerity is found by calculating Laplacians. In addition to the usual Laplacian, on a complex manifold a Hermitian metric determines adjoint operators $\partial^\dagger$ and $\overline{\partial}^\dagger$ for ∂ and $\overline{\partial}$, respectively (so $(\theta, \overline{\partial}\psi) = (\overline{\partial}^\dagger\theta, \psi)$, etc.):

$$\overline{\partial}^\dagger : \Omega^{p,q} \to \Omega^{p,q-1}, \qquad \partial^\dagger : \Omega^{p,q} \to \Omega^{p-1,q}.$$

From these we can form the Laplacians $\Delta_\partial = \partial\partial^\dagger + \partial^\dagger\partial$ and $\Delta_{\overline{\partial}} = \overline{\partial}\overline{\partial}^\dagger + \overline{\partial}^\dagger\overline{\partial}$. We can represent $\overline{\partial}$ cohomology classes $H^{p,q}_{\overline{\partial}}(M)$ with $\overline{\partial}$-harmonic forms $\mathcal{H}^{p,q}_{\overline{\partial}}(M)$, as we did with d and Δ_d. But now an important result is that for a *Kähler* metric,

$$\Delta_d = 2\Delta_{\overline{\partial}} = 2\Delta_\partial,$$

and so all the operators have the same harmonic forms. As a result, and since Δ_d preserves (p,q)-form degree, we have $\mathcal{H}^r(M) = \bigoplus_{p+q=r}\mathcal{H}^{p,q}(M)$, and therefore the de Rham cohomology decomposes into $\overline{\partial}$ cohomology. Define $b_r(M) = \dim H^r(M)$ and $h^{p,q}(M) = \dim H^{p,q}(M) = \dim H^q(\Lambda^p T^*M)$ (Čech–Dolbeault). Then

$$b_r(M) = \sum_{p+q=r} h^{p,q}(M).$$

Further, Hodge $*$ says that $h^{p,q} = h^{n-p,n-q}$ while $h^{p,q} = h^{q,p}$ by complex conjugation. For example, $h^{0,1} = \dim H^1(\mathcal{O})$.

EXAMPLE 5.3.1. *The Hodge numbers of* $T^2 = \mathbb{C}/\mathbb{Z}^2$ *are* $h^{0,0} = h^{0,1} = h^{1,0} = h^{1,1} = 1$. *The generators are* $1, d\bar{z}, dz$, *and* $dz \wedge d\bar{z}$, *respectively.*

EXAMPLE 5.3.2. *A Calabi–Yau manifold can be defined as a complex n-manifold M whose bundle of $(n,0)$-forms is trivial. This bundle $\Lambda^n T^*M$ is called the "canonical bundle" and is often denoted K_M. Triviality of this bundle means that we can identify the total space of K_M as $M \times \mathbb{C}$. So, corresponding to the unit section $M \times \{1\}$ (i.e., the section is the constant function 1) must be a* nowhere vanishing *global* holomorphic *$(n,0)$-form, Ω. Further, every global $(n,0)$-form can be written as $f\Omega$, for f some* function *on M. If M is compact and the form is holomorphic, f must be holomorphic and therefore constant, and the space of holomorphic $(n,0)$-forms is one-dimensional: $h^{n,0}(M) = 1$. If M is further a simply connected Calabi–Yau threefold, as we often assume, then $b_1 = 0$, which implies $h^{1,0}(M) = h^{0,1}(M) = 0$. Serre duality relates $H^1(\mathcal{O})$ with $H^2(\mathcal{O} \otimes K_M)^* = H^2(\mathcal{O})^*$ for a Calabi–Yau threefold, and so $\dim H^{0,2}(M) = 0$ as well (we have used the Dolbeault theorem). In total, Calabi–Yau threefolds have a Hodge diamond with $h^{0,0} = h^{3,3} = h^{3,0} = h^{0,3} = 1$, leaving $h^{1,1}$ and $h^{2,1}$ ($= h^{2,2}$ and $h^{1,2}$, respectively) undetermined (see Fig. 1).*

	1	0	0	1
	0	h^{21}	h^{11}	0
	0	h^{11}	h^{21}	0
q	1	0	0	1
	p			

$h^{p,q}$

FIGURE 1. Hodge diamond of a simply connected Calabi–Yau threefold.

$h^{1,1}$ is the number of possible Kähler forms. We will interpret $h^{2,1}$ in the following chapter.

Another important consequence of Kählerity is that the Levi–Civita connection has no mixed indices, meaning vectors with holomorphic indices remain holomorphic under parallel translation (a real vector can be written as the sum of a vector with holomorphic indices and its conjugate). This says that holonomy maps TM to TM and $\overline{T}M$ to $\overline{T}M$. Since $T_{\mathbb{R}}M \otimes \mathbb{C} = TM \oplus \overline{T}M$, this says that the holonomy sits in a $U(n)$ subgroup of $SO(2n,\mathbb{R})$, where $n = \dim_{\mathbb{C}} M$.

5.3.1. Kähler Potential. The Kähler condition $\partial_i g_{j\bar{k}} = \partial_j g_{i\bar{k}}$ and its conjugate equation $\partial_{\bar{m}} g_{j\bar{k}} = \partial_{\bar{k}} g_{j\bar{m}}$ means that locally we can find a function Φ such that $g_{j\bar{k}} = \partial_j \partial_{\bar{k}} \Phi$. The function Φ is not uniquely determined: Φ and $\Phi + hol + \overline{hol}$ define the same metric, if *hol* is any holomorphic function.

EXAMPLE 5.3.3. *We return to the sphere $S^2 \cong \mathbb{P}^1$ from the first chapter. Recall that the round metric on the unit sphere is given by $g_{\theta\theta} = 1$, $g_{\theta\phi} = 0$, and $g_{\phi\phi} = \sin^2(\theta)$. We mapped the sphere onto the plane by stereographic projection from the two open sets (complements of the poles) and checked that the transition functions were holomorphic. The map was $x = \cot(\theta/2)\cos\phi$, $y = \cot(\theta/2)\sin\phi$. Changing to these coordinates (e.g., $g_{xx} = 1 \cdot (\frac{\partial\theta}{\partial x})^2 + \sin^2(\theta)(\frac{\partial\phi}{\partial x})^2$, etc.) gives $g_{xx} = g_{yy} = 4(x^2+y^2+1)^{-2}$, $g_{xy} = 0$ (**Show this**). In terms of $z = x + iy$, we find $g_{z\bar{z}} = 2(1+|z|^2)^{-2}$. We can write $g_{z\bar{z}} = \partial_z \overline{\partial}_{\bar{z}}[2\log(1+|z|^2)]$, so we find that $\Phi = 2\log(1+|z|^2)$ is a Kähler potential in this patch.*

On the patch with coordinate $\tilde{z} = 1/z$, the metric is $g_{\tilde{z}\bar{\tilde{z}}} = g_{z\bar{z}}/|\tilde{z}|^4 = 2(|\tilde{z}|^2+1)^{-2}$ and $\tilde{\Phi} = 2\log(1+|\tilde{z}|^2)$. On the overlap, $\Phi = \tilde{\Phi} - 2\log z - 2\log\bar{z}$. Note that in this case, $e^{-\Phi}$ has transition function $z^2\bar{z}^2$. This means that it can be written as the single component of a 1×1 Hermitian metric for a holomorphic line bundle with transition function z^{-2}, i.e., $\mathcal{O}(-2)$, or the cotangent bundle! The Chern class of this tangent bundle is simply $\omega/2\pi$, and $\int_{S^2} \omega/2\pi = \chi(S^2) = 2$.

We will encounter another line bundle formed from the Kähler potential in later chapters.

We note some properties and examples of Kähler manifolds (i.e., manifolds equipped with a Kähler metric).

- There exists Φ defined locally such that $g_{i\bar{j}} = \partial_i \overline{\partial}_{\bar{j}} \Phi$.
- $\omega \equiv (i/2) g_{i\bar{j}} dz^i d\bar{z}^{\bar{j}}$ is a closed (real) $(1,1)$-form, called the "Kähler form." On a compact manifold, ω defines an element of $H^{1,1}(M)$, and ω^p defines a non-trivial element of $H^{p,p}(M)$. In particular $h^{p,p} \geq 1$.
- $H^r(M) = \bigoplus_{p+q=r} H^{p,q}(M)$. $\Delta_d = 2\Delta_{\partial} = 2\Delta_{\overline{\partial}}$.
- $\mathbb{P}^n$ is Kähler. Consider the function $\Phi = \log(Z\overline{Z})$ on $\mathbb{C}^{n+1}$. On any coordinate patch, this defines a Kähler potential (Fubini–Study metric).
- The holonomy is in $U(n) \subset SO(2n)$. Γ has pure indices.

- $\mathrm{Tr}_{\mathbb{C}} R_{i\bar{j}} dz^i \wedge d\bar{z}^{\bar{j}} = \partial\bar{\partial} \ln g$.[4]

5.3.2. The Kähler Cone. Since a Kähler metric determines a class in $H^{1,1}(M)$, we can ask which classes could possibly come from Kähler metrics. From the construction via components $g_{i\bar{j}}$, the data describing the metric and the class are virtually the same. However, a non-degenerate Riemannian metric must imply positive volumes for all submanifolds. We can therefore anticipate that the set of possible classes will be closed under arbitrary positive rescalings, with boundary walls where certain submanifolds are assigned zero volume. In other words, we have a "Kähler cone" inside $H^{1,1}(M)$, of the same dimension. At the boundary of the cone, some submanifold has zero volume and we have a singular metric.

EXAMPLE 5.3.4. *We meet one such singularity in Ch. 6, the "conifold." The resolution involves a blow-up procedure that puts a $\mathbb{P}^1 \cong S^2$ where the singularity was. The total resolved space is Kähler and is given as a subspace of $\mathbb{C}^4 \times \mathbb{P}^1$. When the two-sphere vanishes we recover the singularity. In order to look at near-singular metrics, one can simply pull back the metric from $\mathbb{C}^4 \times \mathbb{P}^1$, with $\mathbb{P}^1$ assigned an area of ϵ (e.g., $\omega_{\mathbb{P}^1} = \epsilon i(1+|z|^2)^{-2} dz \wedge d\bar{z})$) and let $\epsilon \to 0$.*

Note that in the interior of the Kähler cone, any class in $H^{1,1}(M)$ can be used to deform the metric slightly. Thus $h^{1,1}(M)$ classifies infinitesimal deformations of the metric that preserves Kählerity. Note that complex form degrees are also preserved, as the Kähler class is still $(1,1)$ in the original metric. In the next chapter we will encounter variations that do not preserve the complex structure.

5.4. The Calabi–Yau Condition

Let us re-examine the Calabi–Yau condition that the canonical bundle is trivial. Since the canonical bundle is the determinant line bundle (highest antisymmetric tensor product) of the holomorphic cotangent bundle, its first Chern class equals minus the first Chern class of the holomorphic tangent bundle, TM. Triviality of the canonical bundle is therefore precisely expressed by the equation $c_1(TM) = 0$. Recalling the definition of the first Chern class from the curvature of a connection, this tells us that the class of

[4] The notation $\mathrm{Tr}_{\mathbb{C}}$ is explained in Sec. 5.4.

$\mathrm{Tr}_{\mathbb{C}} R$ is zero *as a cohomology class*, but not necessarily identically zero as a two-form. $\mathrm{Tr}_{\mathbb{C}} R$ depends on the connection, and if we are using a Kähler metric and its associated connection, then $\mathrm{Tr}_{\mathbb{C}} R$ depends only on the metric (as in the last bullet above). In fact, since Kähler implies $U(n)$ holonomy, $\mathrm{Tr}_{\mathbb{C}} R = 0$ means the vanishing of the trace part of the connection — which implies $SU(n)$ holonomy.

EXERCISE 5.4.1. *It is amusing and illustrative to see how $SU(n)$ imbeds in $SO(2n)$. We imagine the following taking place at a fiber over a point in a complex n-fold. In coordinates x^k, y^k adapted to a complex structure ($z^k = x^k + iy^k$), we may write J as the matrix with 2×2 block diagonal components $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. (The ordering of the basis is $x^1, y^1, x^2, y^2, \ldots$.) Complex conjugation takes the block diagonal form with blocks $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Let Q be the block anti-diagonal matrix with blocks $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Finally, let T be the totally anti-diagonal matrix ($T_{ij} = \delta_{i+j,2n+1}$). Now suppose $A \in SO(2n)$. If A respects the complex structure, i.e., if $AJ = JA$, then $A \in U(n) \subset SO(2n)$.* (**Verify. Hint:** *Use the defining relations of the groups. For example, the transpose A^T is given by $A^T = TAT$.) This condition also characterizes $GL(n,\mathbb{C}) \subset GL(2n,\mathbb{R})$. In this case, A has 2×2 block (not diagonal) form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ and the complex $n \times n$ matrix corresponding to A can be found by replacing such a block with the complex number $a + ib$.* (**Verify.**) *We call such a matrix $A_{\mathbb{C}}$. Then the complex trace $\mathrm{Tr}_{\mathbb{C}} A_{\mathbb{C}}$ and the real trace $\mathrm{Tr}_{\mathbb{R}} A$ are related by $\mathrm{Tr}_{\mathbb{C}} A_{\mathbb{C}} = \frac{1}{2}(\mathrm{Tr}_{\mathbb{R}} A - i \mathrm{Tr}_{\mathbb{R}} AJ)$. At the infinitesimal level, then, $su(n) \subset u(n)$ is given by the vanishing of the complex trace. Note that $so(2n)$ already requires that the real part of the complex trace vanishes. The curvature measures infinitesimal holonomy, so we can state the $SU(n)$ holonomy condition in real coordinates as $-\frac{1}{2}(R_{ab})_e{}^c J_c{}^e = 0$. We must recall now that the Chern classes were defined using complex traces.*

The relation to the Ricci tensor is as follows. The Ricci tensor is defined in real coordinates s^a as $(R_{ac})_b{}^c ds^a \otimes ds^b$ (one verifies that this motley-indexed object is indeed a tensor). On a Hermitian manifold we have the relation $(R_{ac})_b{}^c = (R_{kc})_l{}^c J_a{}^k J_b{}^l$. This fact, along with $R_{[abc]}{}^d = 0$ (the "algebraic Bianchi identity"; the symbol "[...]" means to take the totally antisymmetric piece) and the fact that $J^2 = -1$ allows us to equate the condition of $SU(3)$ holonomy precisely to Ricci flatness. (**Verify. Hints:** *Start with $R_{abc}{}^e J_e{}^c = 0$, apply the algebraic Bianchi identity, rewrite curvature terms* *using the Hermitian property above, and verify that the remaining terms are zero if and only if the Ricci tensor vanishes. Use the fact that the Riemann tensor is anti-symmetric in its first two [as well as last two] indices.)*

It is therefore natural to ask for a manifold with trivial first Chern class ($c_1 = 0$) if, for a given complex structure there exists a Kähler metric (the $(1,1)$-condition depends on the choice of complex structure) such that $\mathrm{Tr}_{\mathbb{C}} R = 0$ *pointwise*. In 1957, Calabi conjectured the existence of such a metric and proved that uniqueness (up to scaling) would follow. In 1977, Yau proved existence. This deep theorem tells us that the moduli space of complex structures is equivalent to the moduli space of Ricci-flat, Kähler metrics. Metrics of $SU(n)$ holonomy are important because they (imply the existence of covariant constant spinors and therefore) allow for superstring compactification (typically, $n = 3$). Without Yau's theorem, describing the space of possible solutions to the coupled, non-linear differential equations would be nearly impossible. The moduli space of complex structures, on the other hand, can be studied with algebro-geometric techniques and is therefore tractable. We will discuss Calabi–Yau moduli in the next chapter.

CHAPTER 6

Calabi–Yau Manifolds and Their Moduli

We discuss deformations of complex structure and the moduli space of complex structures of a Calabi–Yau manifold. Our main example of the quintic threefold and its mirror is developed in detail. Singularities and their smoothings are also discussed.

6.1. Introduction

In this chapter we describe the geometry and structure of the moduli space of complex structures of a Calabi–Yau manifold, with the express goal of investigating these in the example of the (mirror of the) quintic hypersurface in $\mathbb{P}^4$. It may be instructive to refer to the main example Sec. 6.5 of this chapter periodically while reading it.

From physics, one wants solutions to Einstein's equation $R_{\mu\nu} = 0$, where $R_{\mu\nu}$ is the Ricci tensor derived from the metric g. On a Calabi–Yau manifold with a complex structure, we have a unique solution given by the Ricci-flat metric in that complex structure. Let us look at the space of all possible solutions. It turns out that we can deform a solution without changing the complex structure, and we can deform a solution by changing the complex structure. To see these two types of solutions, let us look at a nearby metric $g \to g + h$, and linearize $R_{\mu\nu}$ in this new metric.

EXERCISE 6.1.1. *Assuming $R_{\mu\nu} = 0$ and $\nabla^\mu h_{\mu\nu} = 0$, perform this linearization to find the following equation ("Lichnerowicz equation") for h: $\Delta h_{\mu\nu} + 2R_\mu{}^\alpha{}_\nu{}^\beta h_{\alpha\beta} = 0$. (Here $\Delta = \nabla_\alpha \nabla^\alpha$. The exercise is particularly difficult, since it requires figuring out how to differentiate tensors covariantly, which we have not explicitly discussed.)*

It turns out that on a complex manifold, because the projection to holomorphic and anti-holomorphic degree commutes with the Laplacian, we can separate the solutions to the Lichnerowitz equation into two types. In complex coordinates z^a the solutions (ignoring their conjugates) look like h_{ab} or

$h_{a\bar{b}}$. The $h_{a\bar{b}}$ represent different choices of the Kähler class. The h_{ab} are a new type of deformation.

As we have mentioned, a complex manifold has a notion of holomorphicity furnished by the charts. Two manifolds are isomorphic as complex manifolds if there is a holomorphic diffeomorphism between them. With different charts and different transition functions, the same underlying differentiable manifold may have several complex structures. The h_{ab} represent deformations of the complex structure. In this chapter, we investigate the space of complex structures of a Calabi–Yau manifold. This is called "Calabi–Yau moduli space."

More generally, we can consider any complex manifold and try to vary the complex structure.

EXAMPLE 6.1.1 (T^2). *The prototypical example of a manifold with a moduli space of complex structures is the complex torus or "elliptic curve," $\mathbb{C}/\mathbb{Z}^2$, formed under the identifications $z \sim z + m\lambda_1 + n\lambda_2$ for fixed nonzero (and non-proportional over $\mathbb{R}$) λ_1, λ_2, with $m, n \in \mathbb{Z}$. Let us note immediately that a lattice is not uniquely determined by λ_1 and λ_2, two vectors in $\mathbb{R}^2$. In fact,*

$$\begin{pmatrix} \lambda_1' \\ \lambda_2' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \equiv A \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$(a, b, c, d \in \mathbb{Z})$ generate the same lattice if and only if we can write $\lambda = U\lambda'$ for some integral matrix, U. These equations say $AU = \mathbf{1}$, which means that A must be invertible. So the lattice is defined only up to $GL(2, \mathbb{Z})$ transformation. By taking λ_2 to $-\lambda_2$ if necessary, the complex number $\tau = \lambda_2/\lambda_1$ can be chosen to have positive imaginary part, so that only $PSL(2, \mathbb{Z})$ acts on this ratio ("P" since -1 acts trivially). Now every elliptic curve is isomorphic to one with $\lambda_1 = 1$, since we can define a complex-analytic isomorphism $z \mapsto w = z/\lambda_1$. Then w lives on an elliptic curve with $\lambda_1 = 1$. From here on, we take $\lambda_1 = 1$ and set $\tau = \lambda_2/\lambda_1$. τ is therefore well defined only up to $PSL(2, \mathbb{Z})$ transformation $\tau \to \frac{a\tau+b}{c\tau+d}$, with $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z})$. (This group is generated by $T : \tau \mapsto \tau + 1$ and $S : \tau \mapsto -1/\tau$.) Are there complex-analytic maps between elliptic curves with different nearby values of τ (not related by $PSL(2, \mathbb{Z})$? The fact that there are not will follow from our general discussion. We denote the elliptic curve by E_τ.

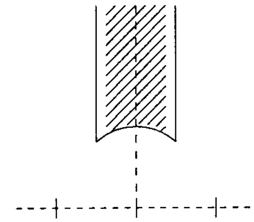


FIGURE 1. The moduli space of an elliptic curve.

The parameter τ is a coordinate for the moduli space of complex structures (see Fig. 1). The elliptic curve admits a flat metric (which descends from the flat metric on $\mathbb{C}$, invariant under the quotiented translations), so the tangent bundle is trivial. E_τ is therefore a Calabi–Yau one-fold, and it is instructive to treat it as such. Note that $b_1(E_\tau) = 2$. A basis for the homology one-cycles can be taken to be the circles a and b, which are the respective images from $\mathbb{C}$ of the line segments connecting z and $z + 1$ (resp. $z + \tau$). The Calabi–Yau holomorphic $(1,0)$-form is simply dz, which we recall is not *exact. The pairing between dz and the cycles a and b looks like*

$$\pi_a \equiv \oint_a dz = 1 = \lambda_1,$$

$$\pi_b \equiv \oint_b dz = \tau = \lambda_2.$$

These integrals are called "periods." Note that we can recover τ from π_b/π_a. We learn that periods can determine the complex structure. This might seem obvious, but elliptic curves are not always presented in such a tidy form. A degree 3 polynomial f in $\mathbb{P}^2$ determines a curve of genus $g = \binom{3-1}{2} = 1$ that has the structure (induced from $\mathbb{P}^2$) of a complex manifold. Therefore, it is holomorphically isomorphic to E_τ, for some τ. τ must be determined by the ten coefficients a_i of f, and one can calculate the periods to find it.

We will follow a similar procedure for the (mirror of the) Calabi–Yau quintic in $\mathbb{P}^4$.

6.2. Deformations of Complex Structure

For a higher-dimensional Calabi–Yau, the situation is more difficult, and one typically can't describe the moduli space globally. Locally, however, we can look at what an infinitesimal deformation of the complex structure would look like (this deforms the very notion of holomorphicity, since the holomorphic coordinates are chosen subordinate to some complex structure).

Infinitesimal deformations of the complex structure form the would-be tangent space to the moduli space of complex structures.[1] There are several ways of doing this. First, we can note that a complex structure is defined by an almost complex structure (an endomorphism $J : T_{\mathbb{R}}M \to T_{\mathbb{R}}M$ such that $J^2 = -1$) whose Nijenhuis tensor N vanishes. We can look at first-order deformations of these equations, modulo changes of the local form of the complex structure associated to coordinate redefinitions. This already has the appearance of a cohomology class.

It will be convenient to switch first to complex coordinates. Let us fix a complex structure and compatible complex coordinates $z^1, \ldots, z^n$. J is diagonalized in these coordinates, so that $J^a{}_b = i\delta^a{}_b$ and $J^{\bar{a}}{}_{\bar{b}} = -i\delta^{\bar{a}}{}_{\bar{b}}$, with mixed components zero. (Note that $J^{\bar{a}}{}_{\bar{b}}$ must be the complex conjugate of $J^a{}_b$ since J is a real tensor.) Now send $J \to J + \epsilon$.

EXERCISE 6.2.1. *Linearize the equation $(J+\epsilon)^2 = 0$ to get $J\epsilon + \epsilon J = 0$, and conclude that the pure indices of ϵ vanish.*

One can linearize the equation $N = 0$, where N is the Nijenhuis tensor associated to $J + \epsilon$, to conclude $\bar{\partial}\epsilon = 0$. In this equation, $\epsilon_{\text{hol}} = (\epsilon^a{}_{\bar{b}}\partial_a)d\bar{z}^{\bar{b}}$ is interpreted as a $(0,1)$-form with values in the holomorphic tangent bundle, so its action as a one-form on a (anti-holomorphic) tangent vector produces a (holomorphic) tangent vector. There is a conjugate equation for $\epsilon_{\text{anti-hol}}$ as well.

EXERCISE 6.2.2. *Perform the linearization mentioned above. Hint: It is convenient to take the two input vectors X and Y for the Nijenhuis tensor to be the holomorphic vectors ∂_a and ∂_b.*

If x' represents new (not necessarily complex) coordinates and $M = \left(\frac{\partial x}{\partial x'}\right)$ is the Jacobian matrix, then $J' = M^{-1}JM$, where we have used matrix notation. Infinitesimally, if x' is close to x then it is generated by a vector field $v^i\frac{\partial}{\partial x^i}$ and $M^i{}_j = \delta^i{}_j + \frac{\partial v^i}{\partial x^j}$. In complex coordinates, this means $J' = J + \bar{\partial}v_{\text{hol}} + \partial v_{\text{anti-hol}}$.

EXERCISE 6.2.3. *Check this.*

[1]It could happen that an infinitesimal deformation makes sense but that no finite deformation can be formed from it. For Calabi–Yau manifolds, this will not be the case.

So (focussing on the upper holomorphic index, for example), coordinate transformations change J by $\bar{\partial}v$. We conclude that infinitesimal deformations of the complex structure are classified by the cohomology group

$$H^1_{\bar{\partial}}(TM).$$

By the Čech–Dolbeault isomorphism, this vector space has an interpretation in Čech cohomology as $H^1(TM)$. This gives vector fields over overlaps along which we infinitesimally twist the overlap functions to produce a deformation of the original complex manifold.

EXAMPLE 6.2.1. *If M is a Riemann surface with no infinitesimal automorphisms (so, no holomorphic vector fields, $H^0(TM) = 0$, which is true for $g \geq 2$) then the Grothendieck–Riemann–Roch formula tells us (see Ch. 3) that $\dim H^0(TM) - \dim H^1(TM) = 3 - 3g$, so $\dim H^1(TM) = 3g - 3$. The moduli space of genus $g > 2$ curves, $\mathcal{M}_g$, has dimension $3g-3$. When $g = 1$, $\dim H^0(TM) = 1$ (it is generated by the global holomorphic vector field ∂_z), so $3 - 3g = 0 \Rightarrow \dim H^1(TM) = 1$. Indeed, we saw that the moduli space was one-dimensional, coordinatized by τ. When $g = 0$, $\dim H^0(TM) = \dim H^0(\mathcal{O}(2)) = 3$, so $3 - 3g = 3 \Rightarrow \dim H^1(TM) = 0$, i.e., $\mathbb{P}^1$ is "rigid" as a complex manifold. The moduli space is a point.*

EXAMPLE 6.2.2. *If M is a Calabi–Yau three-manifold, the canonical bundle (the bundle Λ^3T^*M of holomorphic $(3,0)$-forms) is trivial; hence so is its dual Λ^3TM. Since we have wedge pairing $\wedge : TM \otimes \Lambda^2TM \to \Lambda^3TM = 1$, we learn that $TM \cong (\Lambda^2TM)^* = \Lambda^2T^*M$. So $H^1(TM) = H^1(\Lambda^2T^*M) = H^{2,1}(M)$, and the Hodge number $h^{2,1}$ therefore counts the dimension of the moduli space of complex structures of a Calabi–Yau.*

There is a more hands-on way of seeing these isomorphisms. Let $\Omega = \Omega_{abc}dz^a dz^b dz^c$ be the holomorphic three-form (in some patch). Then we can map $(\epsilon^a{}_{\bar{e}}\partial_a)d\bar{z}^{\bar{e}}$ to a holomorphic $(2,1)$-form $\epsilon^a{}_{\bar{e}}\Omega_{abc}d\bar{z}^{\bar{e}}dz^b dz^c$.

We now have a complete understanding of the Hodge diamond of a Calabi–Yau threefold.

Another way of seeing the space $H^1_{\bar{\partial}}(TM)$ arise is by considering deformations of the $\bar{\partial}$ operator by a vector-valued one-form: $\bar{\partial} \to \bar{\partial} + A$. Linearizing $(\bar{\partial} + A)^2 = 0$ in A gives $\bar{\partial}A = 0$, and the same arguments involving coordinate transformations can be made.

6.3. Calabi–Yau Moduli Space

6.3.1. Unobstructedness. So far, we have constructed the space of infinitesimal deformations. In doing so, we neglected quadratic terms in our deformation parameter. To be sure that a *finite* deformation exists, we must solve the equations without truncation and show that the solution, if written as a power series of solutions at each finite order, converges. This is the content of the theorem of Tian and Todorov.

If we look for finite deformations of $\bar\partial$ we need to solve $(\bar\partial + A)^2 = 0$ for finite A. This amounts to the equation $\bar\partial A + \frac{1}{2}[A, A] = 0$. If we write A as an expansion in a formal parameter, $A = A_1 t + A_2 t^2 + \dots$, then equating powers of t gives the equation (above) $\bar\partial A_1 = 0$ for $n = 1$ and

$$\bar\partial A_n + \frac{1}{2}\sum_{i=1}^{n-1}[A_i, A_{n-i}] = 0$$

for $n \geq 2$. It is possible to show that the sequence of equations can be solved inductively (i.e., $A_n \Rightarrow A_{n+1}$) in a given gauge choice, using the $\partial\bar\partial$-lemma that comes from the Kähler form. We refer the reader to the literature (see Ch. 40) for more details.

EXAMPLE 6.3.1. *The zero set Q of a degree 5 polynomial p in $\mathbb{P}^4$ is a Calabi–Yau manifold, since $c_1 = 0$ follows from the adjunction formula $c(Q) = (1+x)^5/(1+5x)$. We discussed early on that the coefficients of p can be thought of as complex structure parameters. Indeed, the exact sequence of bundles over Q,*

$$0 \longrightarrow TQ \longrightarrow T\mathbb{P}^4 \longrightarrow \mathcal{O}(5)|_Q \longrightarrow 0$$

(recall $N_{Q/\mathbb{P}^4} = \mathcal{O}(5)|_Q$), leads to the long exact sequence (on Q)

$$H^0(TQ) \to H^0(T\mathbb{P}^4) \to H^0(\mathcal{O}(5)|Q) \to H^1(TQ) \to H^1(T\mathbb{P}^4).$$

The ends of this sequence are zero, since Q does not have automorphisms if smooth and since $H^1(\mathbb{P}^4) = 0$.[2] As a result, we can express $H^1(TQ)$ as $H^0(\mathcal{O}(5)|_Q)/H^0(T\mathbb{P}^4)$. Now $H^0(\mathcal{O}(5)|_Q)$ are precisely degree 5 polynomials not vanishing on Q — so p is excluded, and there are $126 - 1 = 125$ of them — and $H^0(T\mathbb{P}^4) = 5^2 - 1$ is the $5^2 - 1 = 24$-dimensional space of

[2] This can be shown to follow from the long exact sequence associated to the Euler sequence restricted to Q: $0 \to \mathcal{O} \to \mathcal{O}(1)^{\oplus 5} \to T\mathbb{P}^4 \to 0$.

automorphisms of $\mathbb{P}^4$ that must be subtracted. In total, we learn $h^{2,1} = 101$, as previously claimed.

Mirror symmetry associates to the quintic a "mirror" $\widetilde{Q}$, whose Hodge diamond is "flipped": $h^{1,1}(\widetilde{Q}) = h^{2,1}(Q) = 101$, and $h^{2,1}(\widetilde{Q}) = h^{1,1}(Q) = 1$. We will construct the family of $\widetilde{Q}$ by quotienting a one-parameter subfamily of the different Q's by a discrete group and then taking care of singularities coming from fixed points.

We therefore expect an honest moduli space $\mathcal{M}_M$ of complex structures of M, of dimension $h^{2,1}(M)$. A natural set of questions now emerges. Can we find coordinates on moduli space? Is there a natural metric? Is it Kähler? Can we find the Kähler potential? Is the Kähler potential associated to a line bundle? Does this line bundle have a natural interpretation, and can we find its metric? The answer to all of these questions is Yes, as we presently learn.

6.3.2. The Hodge Bundle. In different complex structures, the decompositions of the tangent (or cotangent) bundle into holomorphic and anti-holomorphic parts are different. Therefore, what was a closed, holomorphic $(n, 0)$-form in one complex structure will no longer be of type $(n, 0)$ (nor holomorphic) in another complex structure. However, the form will still be closed, as the exterior derivative d is independent of complex structure. In fact, in this description it is easy to see that, to linear order, a $(3, 0)$-form can only change into a linear combination of $(3, 0)$- and $(2, 1)$-forms. The change can be measured by $H^{2,1}$, which is what we already know.

We learn that the cohomology class in H^3 representing the holomorphic $(n, 0)$ form must change over the moduli space of complex structures, $\mathcal{M}_M$. In fact, H^3 forms a bundle over the moduli space, and the Calabi–Yau form is a section of this bundle, its multiples thus determining a line sub-bundle. The bundle of H^3 can be given a flat connection, since we can use integer cohomology, which does not change locally, to define a local trivialization of covariant constant sections. (Specifying the covariant constant sections is enough to define a connection.)

EXAMPLE 6.3.2. *Consider the family M_t of zero loci of the polynomials $P_t = X^2 - t = 0$ in $\mathbb{C}$, i.e., $M_t = \{X = \pm\sqrt{t}\}$. Note that when $t = 0$, $P_t = X^2$, and P_t and dP_t are both zero at $X = 0$, so this is a singular "submanifold." We therefore restrict our "moduli space" to $\mathbb{C}_t \setminus \{0\}$. Over t*

we have the "cohomology bundle" with fiber $H^0(M_t)$ generated by functions equal to 1 on one point and 0 on the other. These functions (sections) are flat in the connection described above. A flat bundle has no curvature, but the vectors can be rotated when transported around a non-trivial loop. Such a rotation for a flat bundle is called a "monodromy." (In physics it is known as a "Wilson line.") On $\mathbb{C}_t\backslash\{0\}$ there is a non-trivial loop $t \mapsto e^{2\pi i x}t$, $x \in [0,1]$, which induces an automorphism of homology and cohomology following from $\pm\sqrt{t} \leftrightarrow \mp\sqrt{t}$. Therefore the total space of the cohomology bundle can be described as $\mathbb{R}^+ \times \mathbb{R} \times \mathbb{C}^2/\left\{((r,\theta);(v_1,v_2)) \sim ((r,\theta+2\pi);(v_2,v_1))\right\}$, where $t = re^{i\theta}$.

We call such a cohomology bundle a "Hodge bundle," and such a connection the "Gauss–Manin connection." The Hodge decomposition (at weight three, for threefolds) of M will change over Calabi–Yau moduli space; we study, therefore "variations of Hodge structure." In our studies, we will find that the line bundle determined by the Calabi–Yau form is the Kähler line bundle, and a natural metric on this bundle will give rise to the Kähler potential on moduli space, from which physical quantities are determined. (In physics this line bundle is called the "vacuum line bundle.")

Now the Calabi–Yau form, defined up to scale, wanders through H^3 as we vary the complex structure. In fact, its position as a line in H^3 (or point in $\mathbb{P}^{h^3-1}$) can be used to describe the complex structure. Note that this description will be redundant, since we know we only need $h^{2,1} = \frac{1}{2}h^3 - 1$ parameters. (Here $h^3 = b_3$.)

Now let M be a Calabi–Yau threefold (or any odd-dimensional Calabi–Yau manifold), and let $\mathcal{H}$ denote the Hodge bundle over $\mathcal{M}_M$ with fibers $H^3(M;\mathbb{C})$. There is a natural Hermitian metric on $\mathcal{H}$ derived from the intersection pairing of three-cycles. Let $\theta,\eta \in H^3(M;\mathbb{C})$. Define $(\theta,\eta) = i\int\theta\wedge\overline{\eta}$. Note this is Hermitian since $(\eta,\theta) = (\theta,\eta)^*$. In fact, the anti-symmetry of the intersection pairing on $H_3(M;\mathbb{Z})$ means that we can find a "symplectic basis" of *real* integer three-forms α_a, β^b, $a,b = 1,\ldots,h^3/2$, such that $(\alpha_a,\alpha_b) = (\beta^a,\beta^b) = 0$, with $(\alpha_a,\beta^b) = i\delta_a{}^b$ (this is akin to finding real and imaginary parts of complex coordinates). This basis is unique up to a $Sp(h^3;\mathbb{Z})$ transformation (i.e., up to preservation of the intersection form). Dual to this basis we have a basis (A^a, B_b) for $H_3(M,\mathbb{Z})$ such that $\int_{A^a}\alpha_b = \delta^a{}_b$, $\int_{B_a}\beta^b = \delta_a{}^b$, all others zero.

6.3.3. Periods and Coordinates on Moduli Space. Since we have a basis for cohomology, we can expand the Calabi–Yau form Ω as $\Omega = z^a\alpha_a - w_b\beta^b$, for some z^a, w_b, $a,b = 1,\ldots,h^3(M)/2 = h^{2,1}(M)+1$. The coordinates z^a and w_b will change as we move in Calabi–Yau moduli space, since Ω will change. In fact, as we have mentioned, since the location of Ω in $H^3(M)$ determines the complex structure, the z^a and w_b *determine* the point in moduli space – even *over-determine* it, as can be seen by counting parameters (moduli space is $h^{2,1}(M)$-dimensional).

It is immediate from the dual basis relations that z^a and w_b can be expressed in terms of the "period integrals"

$$z^a = \int_{A^a}\Omega, \qquad w_b = \int_{B_b}\Omega.$$

Therefore we can express the complex structure (redundantly) in terms of periods $\int_C\Omega$ of the Calabi–Yau form. This is exactly what we did in describing the elliptic curve earlier in this chapter.

In fact, it can be shown that the z^a alone locally determine the complex structure (see references in Ch. 40). We can therefore imagine solving for the w_b in terms of the z^a. Then the z^a are only redundant by one extra variable, but there is also an overall scale of Ω that is arbitrary, and it is often convenient to keep the z^a as homogeneous coordinates on $\mathcal{M}_M$.

6.3.4. The Vacuum Line Bundle. Since the Calabi–Yau form is unique only up to scale, it defines a complex line in the Hodge bundle, i.e., a line sub-bundle. We can define a natural metric on this line bundle

$$h = \|\Omega\|^2 = (\Omega,\Omega) = i\int\Omega\wedge\overline{\Omega}.$$

If z is a coordinate on moduli space and $f(z)$ is a holomorphic function, then $\Omega \to e^{f(z)}\Omega$ defines the same projective section, but $h \to h\,e^f\,e^{\overline{f}}$. We see that h indeed transforms like a Hermitian metric on a line bundle in a new trivialization defined by e^f (never zero). We saw such a phenomenon previously, where we noted that

$$K = -\log\|\Omega\|^2 = -\log\int\Omega\wedge\overline{\Omega}$$

(up to an irrelevant constant) transforms as a Kähler potential, $K \to K - f - \overline{f}$. We therefore can define a metric on moduli space by

$$g_{a\bar{b}} = \partial_a\overline{\partial}_{\bar{b}}K,$$

and this is well defined no matter the gauge choice of Ω, since $f+\overline{f}$ is killed by $\partial\overline{\partial}$.

We are not done yet. We already decided that the tangent space to moduli space is $H^{2,1}(M)$ and there is a natural Hermitian pairing given by the intersection form (or integration). As well, we can choose harmonic (= $\overline{\partial}$-harmonic) representatives θ, η, and compute their inner product as forms using the unique Ricci-flat metric in that complex structure. This metric is called the Weil–Petersson metric. Or, one can look at the variation of the Ricci-flat metric corresponding to the chosen directions and compute the inner product using the inner product on metrics as sections of $\mathrm{Sym}^2(T^*_{\mathbb{R}}M)$. Fortunately, as we will show, these metrics and the one defined from K above are the same!

To see the metric in more detail, let us write the variation of Ω with respect to a coordinate direction z^a as

$$\begin{aligned}\partial_a\Omega &= (3,0)\text{ piece} + (2,1)\text{ piece}\\ &= k_a\Omega + \chi_a,\end{aligned} \tag{6.1}$$

where there are no other terms since the variation of a holomorphic (1,0)-form dx has a $(1,0)$ and a $(0,1)$ piece. Then, keeping track of form degrees and using Eq. (6.1), one finds,

$$\begin{aligned}\partial_a\overline{\partial}_{\overline{b}}K &= \partial_a\left[\frac{-1}{\int\Omega\wedge\overline{\Omega}}\int\Omega\wedge\overline{\partial}_{\overline{b}}\overline{\Omega}\right]\\ &= -\frac{1}{(\int\Omega\wedge\overline{\Omega})^2}\int\partial_a\Omega\wedge\overline{\Omega}\int\Omega\wedge\overline{\partial}_{\overline{b}}\overline{\Omega}+\frac{1}{\int\Omega\wedge\overline{\Omega}}\int\partial_a\Omega\wedge\overline{\partial}_{\overline{b}}\overline{\Omega}\\ &= \frac{\int\chi_a\wedge\chi_{\overline{b}}}{\int\Omega\wedge\overline{\Omega}}.\end{aligned}$$

EXERCISE 6.3.1. *To check the claim, write the variation of the Ricci-flat metric corresponding to the ath direction as*

$$(\delta_a g)_{\overline{\mu}\overline{\nu}} = \left(\frac{\partial g}{\partial z^a}\right)_{\overline{\mu}\overline{\nu}} = -\frac{1}{\|\Omega\|^2}\overline{\Omega}_{\overline{\mu}}{}^{\rho\sigma}\left(\chi_{a\,\rho\sigma\overline{\nu}}\right)$$

(or $\chi_{a\,\alpha\beta\overline{\mu}} = -\frac{1}{2}\Omega_{\alpha\beta}{}^{\overline{\nu}}(\frac{\partial g}{\partial z^a})_{\overline{\mu}\overline{\nu}}$*).*

We have answered all of our questions about moduli space. It is Kähler, with Kähler potential associated to the metric on the vacuum sub-(line) bundle of the Hodge bundle. It is easy to write down explicitly.

Eq. (6.1) is useful in deriving identities by comparing form degrees. Consider: $\int\Omega\frac{\partial\Omega}{\partial z^c}=0$, since there are no $(3,3)$ pieces. This means

$$(z^a\alpha_a - w_b\beta^b, \alpha_c - \partial_c w_d\beta^d) = w_c - z^a\partial_c w_a = 0,$$

where $\partial_c \equiv \frac{\partial}{\partial z^c}$. This says that $w_c = z^a\partial_c w_a = \partial_c(z^a w_a) - w_c$. Define

$$\mathcal{G} \equiv z^a w_a.$$

Then we see $2w_c = \partial_c\mathcal{G}$, which means w_c can be derived from $\mathcal{G}$. Summing with z^c on both sides, we get $z^c\partial_c\mathcal{G} = 2\mathcal{G}$, so *$\mathcal{G}$ is homogeneous of degree 2 in the z^a.*

EXERCISE 6.3.2. *Show that $h = e^{-K} = i\int\Omega\wedge\overline{\Omega}$ is given by*

$$i\int\Omega\wedge\overline{\Omega} = i(\overline{z}^{\overline{a}}\partial_a\mathcal{G} - z^a\overline{\partial}_{\overline{a}}\overline{\mathcal{G}}).$$

6.4. A Note on Rings and Frobenius Manifolds

We learn from the study of topological field theories that physical operators correspond to tangent vectors on the moduli space of theories, since we can use them to perturb the Lagrangian. Since these operators form a ring, this says that there is a product structure on the tangent space to the moduli of topological theories. Such a structure, with a few more requirements such as compatibility with the metric and a direction corresponding to the identity operator, defines a "Frobenius manifold." In the case of Calabi–Yau manifolds, we saw two types of deformations, hence two types of moduli space (and two Frobenius manifolds), Kähler and complex. The Kähler deformations form a ring defined by the "Gromov–Witten invariants," which will be discussed later in the text (the "A-model"). The complex deformations (the "B-model") form another ring, which we now discuss. When M is the quintic threefold, mirror symmetry relates the Kähler ring/Frobenius manifold (A-model) of M (a quintic in $\mathbb{P}^4$) with the complex ring/Frobenius manifold (B-model) of $\widetilde{M}$, another Calabi–Yau. For this case, both rings are commutative.

The ring structure on the B-model can be defined with a symmetric three-tensor Φ_{abc} on moduli space. Using a metric to raise the last index, such a tensor defines a map $TX\otimes TX \to TX$, i.e., the indices are the structure constants of the ring. Thus, given three tangent vectors or elements in

$H^{2,1}(X) \cong H^1(TX)$, we need to produce a number. Let χ_a be a basis for $H^1(TX)$, $a = 1, \ldots, h^{2,1}(X)$. Then

$$\kappa_{a,b,c} = \int_X (\Omega_{\mu\nu\rho}\chi_a^\mu \wedge \chi_b^\nu \wedge \chi_c^\rho) \wedge \Omega,$$

which can be explained simply as follows. χ_a is a $(0,1)$-form with values in the tangent bundle. The wedge product of three χ's is therefore a $(0,3)$-form with values in $\Lambda^3 TX \cong 1$, where in the formula, the holomorphic three-form (with indices) was used to map $\Lambda^3 TX$ to the trivial bundle 1, by contraction. After doing so, we are left with a $(0,3)$-form, which we wedge with Ω to get a $(3,3)$-form to be integrated.

We now show that the Frobenius structure can also be derived from $\mathcal{G}$. This function, the "prepotential," encodes all the data of the topological theory, and mirror symmetry is most often shown by demonstrating the equivalence of prepotentials.

EXERCISE 6.4.1. *Let χ_a be the $(2,1)$ piece of $\partial_a\Omega$ (see Eq. (6.1)),* considered as an element of $H^1(TM)$. *Show*

$$\kappa_{a,b,c} = \partial_a\partial_b\partial_c\mathcal{G}.$$

We learn that every geometric structure on moduli space is encoded in the function $\mathcal{G}$, which is itself determined by the period integrals.

6.5. Main Example: Mirror Symmetry for the Quintic

In this section, we apply our knowledge of moduli space geometry to gain a complete understanding of the moduli space in the simplest threefold example. The differential equations, along with the mirror program, lead to striking mathematical predictions whose verification occupies much of this text.

While we shall only study this one example, it should be mentioned that all of the techniques we use can be generalized to arbitrary Calabi–Yaus inside toric varieties. Though the level of complexity grows in general, the crux of mirror symmetry is well captured by the quintic. (The quintic threefold will be revisited in Sec. 7.10.)

6.5.1. The Mirror Quintic. Let M be a quintic hypersurface in $\mathbb{P}^4$, meaning the zero locus of a homogeneous, degree five polynomial, in other words the zero-set of a section of $\mathcal{O}_{\mathbb{P}^4}(5)$. We saw in Example 6.3.1 that

there were 101 independent (up to $PGL(5,\mathbb{C})$) parameters describing the polynomial, which we can interpret as $h^{2,1}(M) = 101$ complex structure parameters. The $(1,1)$-form on $\mathbb{P}^4$ (e.g., from the Fubini–Study Kähler metric) Poincaré-dual to a hyperplane descends to the single non-trivial generator of $H^{1,1}(M)$.[3]

The "mirror quintic" is another Calabi–Yau manifold $\widetilde{M}$ with reversed Hodge numbers, i.e., $h^{1,1}(\widetilde{M}) = 101$ and $h^{2,1}(\widetilde{M}) = 1$. It can be constructed as follows. Consider a one-dimensional sub-family of quintics defined by the equation $\sum_i a_i X_i^5 - 5\psi\prod_i X_i = 0$ for some coefficients a_i, $i = 1, \ldots, 5$ and ψ. Note that each member of this family has the property that it is preserved under $X_i \to \lambda^{k_i} X_i$, where λ is a fifth root of unity and $\sum_i k_i = 0 \pmod 5$. In fact it is the largest sub-family on which this group G of transformations acts. In fact, when one remembers the scale invariance of $\mathbb{P}^4$ one sees that $G = (\mathbb{Z}_5)^3$. We will define $\widetilde{M}$ by considering the quotient

$$\widetilde{M} = \left(\sum_i a_i X_i^5 - 5\psi\prod_i X_i\right) \Big/ (\mathbb{Z}_5)^3.$$

Note that the a_i can be absorbed by a diagonal $PGL(5,\mathbb{C})$ action, so we momentarily set $a_i = 1$. In the next sections, it will be convenient to reinstate the a_i as parameters, albeit redundant ones. As we will see, $G = (\mathbb{Z}_5)^3$ has fixed points, which means $\widetilde{M}$ is singular unless we resolve the singularities somehow. (We will defer doing so, however, until the end of the chapter.)

Consider $g_1 \in G$, $g_1 : (X_1, X_2, X_3, X_4, X_5) \mapsto (\lambda X_1, X_2, X_3, X_4, \lambda^4 X_5)$, $\lambda^5 = 1$. g_1 generates a $\mathbb{Z}_5$ subgroup of G and clearly fixes the points in $\widetilde{M}$ where $X_1 = X_5 = 0$. The fixed curve C defined by

$$X_1 = X_5 = 0, \qquad X_2^5 + X_3^5 + X_4^5 = 0,$$

is a degree 5 curve in $\mathbb{P}^2 \cong \{X_1 = X_5 = 0\}$ and therefore has genus $\binom{5-1}{2} = 6$, $\chi(C) = -10$. There are other fixed curves and points in $\widetilde{M}$ as well, and their resolution produces new $H^{1,1}$ classes, as we shall see. All told, after resolving to get a smooth manifold, $h^{1,1}(\widetilde{M}) = 101$ and $h^{2,1}(\widetilde{M}) = 1$. Thus ψ is the only parameter describing complex variations of $\widetilde{M}$.

In fact, ψ is slightly redundant, since the holomorphic motion $X_1 \to \lambda X_1$ maps $\widetilde{M}_\psi$ to $\widetilde{M}_{\lambda\psi}$. We learn that only ψ^5 is a good coordinate for the (complex structure) moduli space of $\widetilde{M}$.

[3]Note here that we use M to denote any manifold in the family of manifolds. We will add a label if a particular member of a family of manifolds is needed.

$\widetilde{M}$ can have another type of singularity, namely, $\widetilde{M}$ is singular if $P = 0$ and $dP = 0$, where $P = \sum_i X_i^5 - 5\psi \prod_i X_i$. Setting $\frac{\partial P}{\partial X_4} = 0$ gives $X_4^4 = \psi X_1 X_2 X_3 X_5$. Multiplying by X_4 gives $X_4^5 = \psi \prod X_i$, and the same is true for the other X_k. Thus $\sum_i X_i^5 - 5\psi \prod_i X_i = 0$ and all X_k must be equal. This means, modulo action by G, that all $X_k = 1$, and then $X_i^5 = \psi \prod X_i$ implies $\psi = 1$ (or really $\psi^5 = 1$).

EXERCISE 6.5.1. *Investigate the neighborhood of* $(1, 1, 1, 1, 1)$ *by expanding nonhomogeneous coordinates around* 1 *when* $\psi = 1$ *(remember scale invariance) and conclude that the singularity point is a conifold singularity. (See Sec. 6.6 later in this chapter before attempting.)*

Finally, $\psi \to \infty$ is the singular variety $X_1 \ldots X_5 = 0$, which is the union of five $\mathbb{P}^3$'s ($\{X_i = 0\}$), meeting along lower-dimensional projective spaces defined by common zero sets of the coordinates. The neighborhood of this singularity (ψ large) will be important in the sequel.

Now consider the Hodge bundle $\mathcal{H}$ for $\widetilde{M}$ and its associated Gauss-Manin connection and Hermitian metric. A symplectic basis can be written $(\alpha_1, \alpha_2, \beta^1, \beta^2)$, with dual basis (A^1, A^2, B_1, B_2). Since the α, β form a basis for $H^3(\widetilde{M})$, we can express the Calabi–Yau form Ω at a point in moduli space as a linear combination:

$$\Omega = z^1\alpha_1 + z^2\alpha_2 - w_1\beta^1 - w_2\beta^2$$

for some z^a, w_a. It is immediate from the dual basis relations that

$$z^a = \int_{A^a} \Omega, \qquad w_b = \int_{B_b} \Omega.$$

Therefore we can express the complex structure (redundantly) in terms of periods $\int_C \Omega$ of the Calabi–Yau form.

EXAMPLE 6.5.1. *It is instructive to recall the elliptic curve,* $E_\tau = \mathbb{C}/\langle 1, \tau\rangle$. *The Calabi–Yau form is* $\Omega = dz$, *and a symplectic basis of cycles is* a, *the horizontal circle from* 0 *to* 1, *and* b, *the circle from* 0 *to* τ. *Dual to these we have* $\alpha = dx - (\tau_1/\tau_2)dy$ *and* $\beta = (1/\tau_2)dy$ *(***Check***). Note the orientation is such that* $a \cap b = \int \alpha \wedge \beta = +1$. *Now we can reconstruct the coordinates for moduli space from* $\oint_a \Omega = 1$ *and* $\oint_b \Omega = \tau$, *whose ratio is* τ. *Consider the family of elliptic curves* $X^3 + Y^3 + Z^3 - 3\psi XYZ = 0$ *parametrized by* ψ. *In this case,* Ω *and* α *and* β *are* ψ*-dependent, and* τ *can be recovered from the quotient. In fact, one can write down differential equations in* ψ *governing the periods* $\int_C \Omega$, *and* τ *can be recovered from the solutions. We will do exactly the analogue of this for the mirror quintic.*

6.5.2. The Calabi–Yau Form. First let us write down the Calabi–Yau form explicitly. Define the form Ξ on $\mathbb{C}^5$ by $\Xi = \sum_k (-1)^k dX_1 \wedge \cdots \wedge X_k \wedge \cdots \wedge dX_5$ (note that we replace dX_k by X_k). This form is not invariant under scalings, but $\frac{1}{P}\Xi$, where P is some degree 5 polynomial, is invariant and therefore is well defined on $\mathbb{P}^4$ (though singular along the quintic $P = 0$).

EXERCISE 6.5.2. *(easy)* $\frac{1}{2\pi i}\oint_\gamma \frac{1}{z}dz = 1$, *where* γ *is a circle around the origin. Compute* $\frac{1}{2\pi i}\oint_{\gamma_u} \frac{1}{v} du \wedge dv$ *where* γ_u *is a contour around the plane* $u = 0$ *in* $\mathbb{C}^2$. *(Answer:* dv*)*

Now let γ_P be a small loop around $P = 0$ in $\mathbb{P}^4$. Then

$$\Omega = \int_{\gamma_P} \frac{\Xi}{P}$$

is a well-defined holomorphic $(3, 0)$-form on $P = 0$. The reasoning is simple from the exercise above. Since P can be considered as a coordinate in a direction normal to $P = 0$ (as long as this variety is non-singular), we can rewrite dX_4, say, as $\left(\frac{\partial X_4}{\partial P}\right) dP$, and the dP/P gets integrated to a constant. Therefore,

$$\Omega = \frac{X_5 dX_1 \wedge dX_2 \wedge dX_3}{\left(\frac{\partial X_4}{\partial P}\right)}.$$

In nonhomogeneous coordinates, one can set $X_5 = 1$ above and replace X_i by x_i, $i = 1, \ldots, 3$, as coordinates on $P = 0$ (X_4 is determined by $P = 0$.)

Let Γ_i be a basis for $H_3(\widetilde{M})$. Define the periods

$$\Omega_i \equiv \int_{\Gamma_i} \int_{\gamma_P} \frac{\Xi}{P}. \tag{6.2}$$

We will find differential equations for Ω_i in terms of the a_i and ψ.

6.5.3. The Picard–Fuchs Differential Equations. By using simple scaling arguments, we will be able to derive differential equations obeyed by the Ω_i. It will turn out that these are enough to determine all of the periods in the neighborhood of a singular point in complex structure moduli space ($\psi \to \infty$). Such differential equations for the periods are called "Picard–Fuchs" equations.

Recall that $\widetilde{M}$ is defined by quotienting the zero set of the polynomial $\sum_i a_i X_i^5 - 5\psi \prod_i X_i$ by a $(\mathbb{Z}_5)^3$ action (then resolving the singularities). By the explicit form of the Ω_i given in Eq. (6.2), we have the following relations:

(1) $\Omega_i(\lambda a_1, \ldots, \lambda a_5, \lambda\psi) = \lambda^{-1}\Omega_i(a_1, \ldots, a_5, \psi)$. Let $(s_1, \ldots, s_6) = (a_1, \ldots, a_5, \psi)$. Taking $\frac{\partial}{\partial\lambda}$ gives (at $\lambda = 1$)

$$\left(\sum_j s_j \frac{\partial}{\partial s_j} + 1\right) \Omega_i = 0.$$

This says that the Ω_i are homogeneous of weight -1 in the coordinates.

(2) $\Omega_i(a_1, \ldots, \lambda^5 a_j, \ldots, \lambda^{-5} a_5, \psi) = \Omega_i(a_1, \ldots, a_5, \psi)$, as the change can be absorbed by the $PGL(5, \mathbb{C})$ transformation $X_j \mapsto \lambda X_j$, $X_5 \mapsto \lambda^{-1} X_5$. Now $\frac{\partial}{\partial\lambda}$ at $\lambda = 1$ gives

$$\left(a_i \frac{\partial}{\partial a_i} - a_5 \frac{\partial}{\partial a_5}\right) \Omega_i = 0.$$

This means Ω_i is a function of $a_1 \ldots a_5$.

(3) The relation $(X_1)^5 \ldots (X_5)^5 = (X_1 \ldots X_5)^5$ gives the equation

$$\left(\prod_i \frac{\partial}{\partial a_i} - \left(\frac{1}{5}\frac{\partial}{\partial\psi}\right)^5\right) \Omega_i = 0.$$

Note that the toric nature of $\mathbb{P}^4$ was crucial here, as we used scalings in our argument. The five powers of 1 in the product of the $\frac{\partial}{\partial a_i}$ are ultimately due to the weights of the $\mathbb{C}^*$ quotienting action. In fact, the Picard–Fuchs equations for Calabi–Yaus in toric varieties can be derived from the toric data and provide many interesting examples of mirror symmetry calculations. We will not pursue such generalities here, however.

The first two equations say that $\Omega_i = \frac{1}{5\psi}\omega_i(\frac{a_1 \ldots a_5}{(5\psi)^5})$. Therefore, we put $z = \frac{a_1 \ldots a_5}{(5\psi)^5}$ and rewrite the last equation.

$$\left[\partial_{a_1} \ldots \partial_{a_5} - (\tfrac{1}{5}\partial_\psi)^5\right] \frac{1}{5\psi}\omega(z) = 0.$$

Now *on a function of* z, we have $\partial_{a_i} = \frac{z}{a_i}\partial_z = \frac{1}{a_i}\Theta$, where $\Theta \equiv z\frac{d}{dz}$. So we can replace $\partial_{a_1} \ldots \partial_{a_5}$ by $\frac{1}{a_1 \ldots a_5}\Theta^5$.

EXERCISE 6.5.3. *Show* $\frac{1}{5}\partial_\psi \frac{1}{(5\psi)^N} f(z) = -\frac{1}{(5\psi)^{N+1}}(5\Theta + N) f(z)$. *Show, using this commutation relation, that*

$$\left(\frac{1}{5}\partial_\psi\right)^5 \frac{1}{5\psi}\omega = -\frac{1}{(5\psi)^6}(5\Theta + 5) \ldots (5\Theta + 1)\omega.$$

Putting things together, we get

$$\left[\Theta^5 - z(5\Theta + 5) \ldots (5\Theta + 1)\right] \omega = 0.$$

Using $z(\Theta + 1) = \Theta z$, we get

$$\Theta\left[\Theta^4 - 5z(5\Theta + 4) \ldots (5\Theta + 1)\right] \omega = 0.$$

We now focus on the equation

$$\left[\Theta^4 - 5z(5\Theta + 4) \ldots (5\Theta + 1)\right] f = 0.$$

Define $\mathcal{L}$ to be the differential operator in brackets. Then $\mathcal{L} f = 0$. It can be shown that the periods obey this equation, factored from the fifth-order equation that precedes it. The reason the periods obey a fourth-order equation is as follows. The first derivative of Ω lives in $H^{3,0} \oplus H^{2,1}$; the second mixes with $H^{1,2}$ as well. Clearly, the fourth is expressible in terms of lower derivatives.

Due to the logarithmic derivatives in $\mathcal{L}$, the solutions have singularities.

EXAMPLE 6.5.2. *Consider* $\Theta^3 f = 0$. *A basis for solutions is*

$$f_0 = 1,\ f_1 = \frac{1}{2\pi i}\ln z,\ f_2 = \frac{1}{2}(\frac{1}{2\pi i}\ln z)^2,$$

where f_0 *is a basis for* $\operatorname{Ker}\Theta$, f_1 *for* $\operatorname{Ker}\Theta^2/\operatorname{Ker}\Theta$, *etc. These solutions undergo a monodromy transformation, due to the branch cut:* $f_1(e^{2\pi i}z) = f_1(z) + f_0(z)$, *etc. The monodromy matrix* $M = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ *is maximally unipotent, meaning* $(M - \mathbf{1})^k$ *does not vanish until* $k = 3$, *the order of the differential equation.*

At $z = 0$ our equation looks like $\Theta^4 = 0$, and we expect our monodromy structure to be maximally unipotent, with one invariant holomorphic solution, as in the example. Let us look for a holomorphic solution by power series methods. Write $f_0 = \sum_{n=0}^\infty c_n z^n$. Noting $\Theta z^n = n z^n$, then $\mathcal{L} f_0 = 0$ leads to the recursion

$$n^4 c_n = 5\,(5(n-1) + 4) \ldots (5(n-1) + 1)\, c_{n-1}.$$

We get $c_n/c_{n-1} = (5n)!/[n^5(5n-5)!]$, whence

$$c_n = \frac{(5n)!}{(n!)^5} = \frac{\Gamma(5n+1)}{(\Gamma(n+1))^5}.$$

In fact, we can write a family of solutions $f = \sum_n c(n,\rho)z^{n+\rho}$, where $c(n,\rho) = \Gamma(5(n+\rho)+1)/\Gamma(n+\rho+1)^5$. We put $f_p = \frac{1}{p!}\left(\frac{1}{2\pi i}\frac{\partial}{\partial\rho}\right)^p f|_{\rho=0}$. Then f_0 is our holomorphic solution, and the $f_{k\leq 3}$ have $(\ln z)^k$ singularities (f_4, of course, is a linear combination of $f_0,\dots,f_3$). Note that the f_k are not themselves periods of integral cycles. The cycle not vanishing as $z\to 0$ must correspond to the holomorphic solution. Then, Poincaré duality tells us about the leading singularities of the periods of three other cycles, so the three other periods look like f_k + *less singular solutions.* Finally, these additional terms are fixed by requiring the periods to have integral monodromies around the singular points of moduli space.

6.5.4. Mirror Symmetry. The beauty of mirror symmetry comes from the interpretation of our function $\mathcal{G}$ of the coordinate z (we haven't yet said how to relate the solutions f_p to the periods z^a and w_a). The philosophy is that $\widetilde{M}$ and M define the same physical theory (for why, see the physics chapters!). The measurable quantities of the physical theory are the triple pairings $\kappa_{a,b,c}$, defined through $\mathcal{G}$ by its derivatives (in our example, there is only one coordinate for moduli space).

The interpretation of the $\kappa_{a,b,c}$ for M is in terms of holomorphic maps (from genus 0 curves) into M, which meet the three divisors dual to the $H^{1,1}$ classes corresponding to the differentiated directions in moduli space. The first approximation to this quantity is by degree 0 maps, or points in M.[4] The number of points intersecting three divisors is equal to the triple intersection. Higher-degree maps correct this "classical" intersection, which is why the ring defined by the $\kappa_{a,b,c}$ is called the "quantum cohomology ring." Roughly speaking, the higher-degree maps are weighted by $e^{-\text{Area}}$, so the expansion we derive is valid near where M has large radius, which corresponds to being on moduli space near where $\widetilde{M}$ is maximally unipotent ($z=0$), also called "large complex structure." Mirror symmetry allows us to compute this ring with the equivalent, mirror model on $\widetilde{M}$, and extract these numbers of curves ("Gromov–Witten invariants").

[4] "Degree," here, is the class of the image curve, written as $d[\mathbb{P}^1]$, where $[\mathbb{P}^1]$ generates the one-dimensional $H^{1,1}(M)$.

One writes $F = \mathcal{G}/(z^0)^2$ as a function of $t = t(z) = z^1/z^0$ (recall that the z^a were homogeneous coordinates on moduli space). It has the form (up to some factors of $e^{2\pi i}$) $F = \frac{5}{6}t^3$ + lower order + $F_{\text{inst}}(q)$, where $q = e^{2\pi i t}$ and F_{inst} represents the degree $d>0$ curves. Then

$$F_{\text{inst}} = \sum_{d>0} K_d q^d.$$

A decade of developments in mathematics has been geared toward the proper formulation and computation of the K_d. Many of the remaining chapters of this text will describe these calculations.

As for the approach via differential equations, we note only that the manipulations we have performed can be done (with varying computational ease) in any toric variety in which a Calabi-Yau can be expressed as a hypersurface or a complete intersection of such. A version of mirror symmetry can be performed for non-compact Calabi-Yaus as well ("local mirror symmetry"). Some of these non-compact Calabi–Yaus are local models of resolutions of singularities. We conclude this chapter with a discussion of several such examples, as well as the conifold singularity (at $z=1$) of $\widetilde{M}$, which we encountered earlier.

6.6. Singularities

We turn now to a brief discussion of singularities in Calabi–Yau manifolds. Singularities and their smoothings are not just important for understanding the mirror quintic; their local geometries often have interesting physical interpretations as well.

There are many different types of singularities and ways of smoothing them. In this section, we will consider just a few. In the case of a Calabi–Yau singularity, we are directed somewhat in our smoothing by the condition that we want the smooth manifold to have trivial canonical bundle (hence no "discrepancy" in the canonical bundle — such resolutions are thus called "crepant"). The conifold singularity appears frequently and with import in string theory, so we turn now to a discussion.

6.6.1. The Conifold Singularity. The conifold singularity refers to a singular point in a threefold that locally (in some coordinates) looks like

$$XY - UV = 0$$

in $\mathbb{C}^4$. Note that the polynomial $p = XY - UV$ is zero at the origin, and $dp = YdX + XdY - VdU - UdV = 0$ there too, so the origin is singular. This can take other guises. For example, if $A = (X+Y)/2$, $B = i(X-Y)/2$, $C = i(U+V)/2$, $D = (U-V)/2$, the conifold looks like

$$A^2 + B^2 + C^2 + D^2 = 0, \tag{6.3}$$

which is known in the mathematical literature as an "ordinary double point" or "node."

EXAMPLE 6.6.1. *Show that at $z = 1$ the mirror quintic $\widetilde{M}$ has a conifold singularity at the point* $(1,1,1,1,1)$.

Let us investigate the region around the singularity more closely. Set $\vec{x} = (\operatorname{Re} A, \operatorname{Re} B, \operatorname{Re} C, \operatorname{Re} D)$ and $\vec{y} = (\operatorname{Im} A, \operatorname{Im} B, \operatorname{Im} C, \operatorname{Im} D)$. Set $r^2 = \vec{x}^2 + \vec{y}^2$ and let us consider $r^2 > 0$, fixed. The real and imaginary parts of Eq. (6.3) say

$$\vec{x}^2 - \vec{y}^2 = 0, \qquad \vec{x} \cdot \vec{y} = 0.$$

The first says that $\vec{x}^2 = \frac{1}{2}r^2$, so $\vec{x}$ lives on an S^3, while the second says that $\vec{y}$ is perpendicular to $\vec{x}$ with $\vec{y}^2 = \frac{1}{2}r^2$. Thus for a fixed $r^2 > 0$ and given x, there is an S^2 of choices for $\vec{y}$. Thus we have an S^2 fibered over S^3. In fact, all such fibrations are trivial, and we get $S^2 \times S^3$. At $r^2 = 0$ we only have $\vec{x} = \vec{y} = 0$, a point. In total, a neighborhood of the conifold locus looks like a cone over $S^2 \times S^3$ (see Fig. 2).

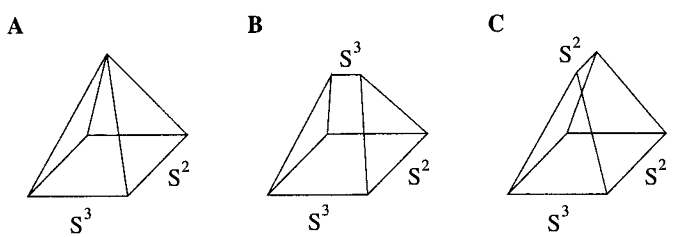


FIGURE 2. A. The conifold singularity. B. Its deformation. C. Its resolution.

Deformation: We can deform the defining equation of the conifold so that it is no longer singular at the origin. For example, put

$$XY - UV = \epsilon.$$

As this smoothing of the singularity results from changing the polynomial, it corresponds to the desingularization arising from deforming the complex structure (e.g., $z \neq 1$ for $\widetilde{M}$). For simplicity, let us use the form of Eq. (6.3) for the conifold and change the right-hand side from 0 to R^2, with R a positive real number. Then the analysis proceeds as before, only now we have

$$\vec{x}^2 - \vec{y}^2 = R^2, \qquad \vec{x} \cdot \vec{y} = 0.$$

Again, we have a family of $S^2 \times S^3$, only this time the minimum radius S^3 is R, when $\vec{y} = 0$.

In fact, if we write $\vec{\tilde{x}} \equiv \vec{x}/\sqrt{R^2 + \vec{y}^2}$, then the defining equations become

$$\vec{\tilde{x}}^2 = 1, \qquad \vec{\tilde{x}} \cdot \vec{y} = 0.$$

In fact, this is the equation for the total space of T^*S^3, with $\pi : T^*S^3 \to S^3$ given by $(\vec{\tilde{x}}, \vec{y}) \mapsto \vec{\tilde{x}}$. To see the relation, replace dx_i by y_i in d of the equation $f = R^2$ (i.e., $df = 0$), where $f = x_1^2 + \cdots + x_4^2 - R^2$.

Resolution: Another way to remove a singularity on a space X is to construct a smooth space $\widetilde{X}$ which looks exactly like X away from the singular points.

EXAMPLE 6.6.2. The Blow-up of a Point. *Consider $\mathbb{C}^2$. We can consider a new space $\widetilde{\mathbb{C}}^2$ where the origin is replaced by a new set as follows. Any smooth path toward the origin contains an extra piece of data in addition to its endpoint, namely the line tangent to the path at the origin. This line defines a point in $[\lambda_1, \lambda_2]$ in $\mathbb{P}^1$. (You can consider the same construction in $\mathbb{R}^2$, where you remember the slope of the path at the origin — the resulting space sort of looks like a spiral staircase.) Formally, define $\widetilde{\mathbb{C}}^2 \subset \mathbb{C}^2 \times \mathbb{P}^1$ by the equations*

$$X_1\lambda_2 = X_2\lambda_1.$$

Note that when $(X_1, X_2) \neq (0,0)$, $[\lambda_1, \lambda_2]$ is completely determined (remember scaling), but when $X_1 = X_2 = 0$, the λ_i can range over all of $\mathbb{P}^1$. The map $\pi : (X_1, X_2; [\lambda_1, \lambda_2]) \mapsto (X_1, X_2)$ from $\widetilde{\mathbb{C}}^2$ to $\mathbb{C}^2$ is therefore an isomorphism outside the origin, while $\pi^{-1}((0,0)) \cong \mathbb{P}^1$. This set is called the "exceptional divisor."

This procedure generalizes to $\mathbb{C}^n$ with $\mathbb{P}^{n-1}$ as an exceptional divisor, where we use the equations $X_i\lambda_j = X_j\lambda_i$. In addition, we can blow up along a subvariety by considering slices in the normal direction, in which case the variety appears as a point.

To resolve the conifold, first note that the first form of the conifold can be presented as $\det \begin{pmatrix} X & U \\ V & Y \end{pmatrix} = 0$. We now resolve the singular point by considering a new space $Z \subset \mathbb{C}^4 \times \mathbb{P}^1$ defined by

$$\begin{pmatrix} X & U \\ V & Y \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0; \qquad \text{i.e.,}$$

$$X\lambda_1 + U\lambda_2 = 0, \qquad V\lambda_1 + Y\lambda_2 = 0. \tag{6.4}$$

Note that by sending $(X, Y, U, V; \lambda_1, \lambda_2) \mapsto (X, Y, U, V)$ we have a map from Z to the conifold.

EXERCISE 6.6.1. *Show that this map is an isomorphism outside of the origin.*

The singular point at the origin 0 has been replaced by $\pi^{-1}(0) = \mathbb{P}^1 \cong S^2$. In this new space, therefore, we have an extra element in the homology class H_2, and since it is defined by algebraic equations, we in fact get new classes in $h^{2,2}$ and therefore $h^{1,1}$ (by Poincaré duality) as well. If we vary the size of the blow-up $\mathbb{P}^1$ and let it shrink to zero, we recover the conifold singularity.

The space Z has another description. Let us cover Z by two sets, $A = \{\lambda_1 \neq 0\}$ and $B = \{\lambda_2 \neq 0\}$. On A let $u = \lambda_2/\lambda_1$. Then Eq. (6.4) implies $X = -Uu$, so (u, U) are coordinates on A. On B we have $v = \lambda_1/\lambda_2$ and $U = -vX$, so (v, X) are coordinates on B, and on the overlap $U = -u^{-1}X$ tells us that these coordinates form $\mathcal{O}_{\mathbb{P}^1}(-1)$. Including V and Y shows us that Z is the total space of $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$. This is perhaps the most basic "local" (non-compact) Calabi–Yau threefold.

The process of varying a complex structure from a smooth Calabi–Yau so that a conifold singularity appears, and then resolving that conifold so that a new S^2 appears is called a "conifold transition."

6.6.2. Calabi–Yau Surface Singularities. Singularities within a Calabi–Yau surface (two–fold) are classified by finite subgroups Γ of $SU(2)$, and have a local description as $\mathbb{C}^2/\Gamma$.

EXAMPLE 6.6.3. *Let $\Gamma = \mathbb{Z}_{n+1}$ be generated by $\begin{pmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{pmatrix}$, $\xi = e^{2\pi i/(n+1)}$. We can coordinatize $\mathbb{C}^2/\Gamma$ by invariant polynomials $u = X^{n+1}$, $v = Y^{n+1}$, and $t = XY$. These obey the relation $uv - t^{n+1} = 0$, so these singularities can be described by the equations $uv = t^{n+1}$ in $\mathbb{C}^3$. These are called the A_n singularities.*

The McKay correspondence says that there is a relationship,

finite subgroups of $SU(2) \longleftrightarrow$ simply laced (i.e., ADE) Lie algebras,

which is described as follows.[5] Let V_i be the irreducible representations of $\Gamma \subset SU(2)$. Let R be the representation induced by the fundamental representation of $SU(2)$. Decompose

$$V_i \otimes R \cong \bigoplus_j C_{ij} \cdot V_j.$$

Then the McKay correspondence states that C_{ij} is the adjacency matrix of the affine version of the associated Lie algebra. Further, the resolution of $\mathbb{C}^2/\Gamma$ has, in its middle homology, spheres intersecting in the pattern of the Dynkin diagram of Γ, with one sphere for each vertex and an intersection for each edge.

This correspondence has a physical interpretation in terms of "geometric engineering," to be discussed in Sec. 36.1.

EXAMPLE 6.6.4. *For $\Gamma = \mathbb{Z}_N$, the irreducible representation $\mathbf{k}$ is given by ξ^k, where $\xi = e^{2\pi i/N}$ (clearly $k \sim k+N$). Then $R = \mathbf{1} \oplus \mathbf{-1}$ and $\mathbf{k} \otimes R = (\mathbf{k+1}) \oplus (\mathbf{k-1})$. So $C_{ij} = \delta_{|i-j|,1}$, which is the adjacency matrix of a cycle of $N+1$ vertices. This is the Dynkin diagram of the affine Lie algebra $\hat{A}_N$. Not all the spheres are linearly independent, and if we excise a dependent one, we recover A_N.*

To "see" the spheres, consider an ordinary double point inside a surface: $x^2 + y^2 + w^2 = \epsilon$ (we tacitly assume that deformation and resolution are equivalent for surfaces, as they both introduce two-spheres, and we work with the former). Write this as $x^2 + y^2 = \epsilon - w^2$, and let us assume ϵ is real and positive. The right-hand side has two solutions, at $w_\pm = \pm\sqrt{\epsilon}$, at which there is a single solution for x and y, i.e., $x = y = 0$. At a fixed real value of w between w_- and w_+, there is a real x-y circle of solutions. The family of circles forms a non-trivial two-cycle. For higher A_n singularities, we can replace the right-hand side by a polynomial $\epsilon - P_{n+1}(w)$, which has

[5] The D_n singularities are defined by the polynomial $u^2 + tv^2 + t^{n-1}$, E_6 by $u^2 + v^3 + t^4$, E_7 by $u^2 + v^3 + vt^3$, and E_8 by $u^2 + v^3 + t^5$.

$n+1$ *roots. A similar analysis yields the desired cycle of spheres between roots.*

6.6.3. Surfaces in a Calabi–Yau. If B is a surface in a Calabi–Yau M, then we have the sequence

$$0 \to TB \to TM|_B \to N_{B/M} \to 0.$$

Now taking Λ^3 tells us $\Lambda^2 TB \otimes N_{B/M} \to \Lambda^3 TM \cong 1$, by the Calabi–Yau condition. Thus $N_{B/M} \cong \Lambda^2 T^* B$, which is the canonical bundle of B. We learn that $N_{B/M} \cong K_B$, i.e., the local geometry of a surface inside a Calabi–Yau is its canonical bundle, which is intrinsic to the surface.

Toric descriptions of the canonical bundles of some Fano surfaces can be found in Ch. 7. These geometries are important for local mirror symmetry, which is similar to the compact version of mirror symmetry developed in this book. Though we do not describe the mathematics of local mirror symmetry here, the same physical proof applies (see Sec. 20.5).

CHAPTER 7

Toric Geometry for String Theory

7.1. Introduction

We saw a brief introduction to toric varieties in Sec. 2.2.2. In this chapter, we give a more thorough treatment.

Toric varieties have arisen in a wide range of contexts in mathematics during recent decades, and more recently in physics. We do not attempt completeness here, but instead focus on certain themes that recur in the interaction of toric geometry with string theory, providing many examples. Many topics that could have been covered here have been completely omitted.

To anchor the subject matter, here is a formal definition of a toric variety.

DEFINITION 7.1.1. *A* toric variety X *is a complex algebraic variety containing an algebraic torus* $T = (\mathbb{C}^*)^r$ *as a dense open set, together with an action of* T *on* X *whose restriction to* $T \subset X$ *is just the usual multiplication on* T.

EXAMPLE 7.1.2. *Consider* $\mathbb{CP}^r$ *with homogeneous coordinates expressed as* $(x_1, \ldots, x_{r+1})$. *The dense open subset*

$$T = \{x : x_i \neq 0, i = 1, \ldots, r+1\} \subset \mathbb{CP}^r$$

is isomorphic to $(\mathbb{C}^*)^r$ *and acts on* $\mathbb{CP}^r$ *by coordinatewise multiplication, giving* $\mathbb{CP}^r$ *the structure of a toric variety.*

As the utility of toric varieties came to be appreciated, two standard ways of characterizing them emerged. Normal toric varieties (meaning that all singularities are normal) can all be described by a *fan*, and projective toric varieties (with a specified ample line bundle) can all be described by lattice points in a *polytope*. Toric varieties that are both normal and projective can be described by either a fan or a polytope, which turn out to be related to each other. Reinterpretation of certain data for a fan as data for a polytope leads to a geometric construction of mirror manifolds. We develop both of these descriptions and their relationships.

We start by discussing fans of toric varieties. This description of toric varieties is given in Sec. 7.2, emphasizing the use of homogeneous coordinates. We explain how to construct toric varieties from fans and conversely. The gauged linear sigma model (GLSM), which is closely related to toric geometry, is studied in Sec. 7.3. In particular, in the absence of a superpotential, the set of supersymmetric ground states of the GLSM is a toric variety. Conversely, toric varieties can be described as the set of ground states of an appropriate gauged linear sigma model. This link is explored further in Sec. 7.4, where we explicitly identify the connection between intersection numbers in toric geometry and charges in the GLSM. We also develop the geometry of curves and divisors in that section. In Sec. 7.5 we discuss orbifolds in toric geometry and see how they arise naturally in a general context. Sec. 7.6 considers toric blow-ups, and Sec. 7.7 toric morphisms. In Sec. 7.8 we take a look at the application of toric geometry to $\mathcal{N}=2$ geometric engineering.

The final sections are devoted to polytopes and mirror symmetry. In Sec. 7.9, we explain how to construct toric varieties from polytopes and the converse. This section also relates the fan and polytope descriptions of toric varieties. Sec. 7.10 is devoted to mirror symmetry. We will formulate Batyrev's geometric construction of mirror symmetry for Calabi–Yau hypersurfaces in toric varieties as an interchange of the fan and polytope descriptions. Then we relate the toric language to the physical description of mirror symmetry given in Ch. 20.

7.2. Fans

Let N be a lattice, and set $N_{\mathbb{R}} = N \otimes \mathbb{R}$. We will denote the rank of N by r. At times, we will fix an isomorphism $N \simeq \mathbb{Z}^r$, which induces an isomorphism $N_{\mathbb{R}} \simeq \mathbb{R}^r$. At other times, there will be benefits to thinking of N as an abstract lattice.

DEFINITION 7.2.1. *A* strongly convex rational polyhedral cone $\sigma \subset N_{\mathbb{R}}$ *is a set*

$$\sigma = \{a_1 v_1 + a_2 v_2 + \cdots + a_k v_k \mid a_i \geq 0\}$$

generated by a finite set of vectors $v_1, \ldots, v_k$ *in* N *such that* $\sigma \cap (-\sigma) = \{0\}$.

Without further comment, strongly convex rational polyhedral cones will simply be referred to as cones in this chapter.

DEFINITION 7.2.2. *A collection Σ of strongly convex rational polyhedral cones in $N_{\mathbb{R}}$ is a called a* fan *if*

(1) *each face of a cone in Σ is also a cone in Σ, and*
(2) *the intersection of two cones in Σ is a face of each.*

There will also be a need for the dual lattice $M = \mathrm{Hom}(T, \mathbb{C}^*) \simeq \mathrm{Hom}(N, \mathbb{Z})$ of characters of T. The natural pairing between M and N will be written as $\langle\ ,\ \rangle : M \times N \to \mathbb{Z}$. We will also need the accompanying vector space $M_{\mathbb{R}} = M \otimes \mathbb{R}$.

7.2.1. Constructing Toric Varieties from Fans.

There are two standard ways to construct a toric variety X_Σ from a fan Σ yielding the same result. The original construction associates an affine toric variety $X_\sigma = \mathrm{Spec}\, \mathbb{C}[\check{\sigma} \cap M]$ to each cone in $\sigma \in \Sigma$, then glues them together in a natural way to obtain X_Σ. We will not discuss the details of this construction here, but will recover another description of X_σ later in this chapter.

Instead, it is more convenient for applications to mirror symmetry to construct toric varieties via homogeneous coordinates.

Let Σ be a fan in $N_{\mathbb{R}}$, and let $\Sigma(1)$ be the set of edges (one-dimensional cones) of Σ. For each $\rho \in \Sigma(1)$, let $v_\rho \in N$ be the unique generator of the semigroup $\rho \cap N$. This v_ρ is referred to as the *primitive generator* of ρ. Identifying ρ with v_ρ, the set $\Sigma(1)$ can be thought of as a subset of N.

For ease of exposition, we assume that the v_ρ span $N_{\mathbb{R}}$ as a vector space for the rest of this chapter.

Putting $n = |\Sigma(1)|$, the toric variety X_Σ is constructed as a quotient of an open subset in $\mathbb{C}^n$ as follows.

To each edge $\rho \in \Sigma(1)$ is associated a coordinate x_ρ. It is sometimes convenient to choose an ordering $\{v_1, \ldots, v_n\}$ of $\Sigma(1)$. Then the coordinates can be denoted by $(x_1, \ldots, x_n)$ if desired.

Let $\mathcal{S}$ denote any subset of $\Sigma(1)$ that does *not* span a cone of Σ. Let $V(\mathcal{S}) \subset \mathbb{C}^n$ be the linear subspace defined by setting $x_\rho = 0$ for all $\rho \in \mathcal{S}$. Now let $Z(\Sigma) \subset \mathbb{C}^n$ be the union of all of the $V(\mathcal{S})$. The toric variety will be constructed as a quotient of $\mathbb{C}^n - Z(\Sigma)$ by a group G.

To define G, consider the map $\phi : \mathrm{Hom}(\Sigma(1), \mathbb{C}^*) \to \mathrm{Hom}(M, \mathbb{C}^*)$ defined by sending a map (of sets) $f : \Sigma(1) \to \mathbb{C}^*$ to the map (of groups) $m \mapsto \prod_{v\in\Sigma(1)} f(v)^{\langle m,v\rangle}$.

In coordinates, ϕ is very easy to write down. If v_j has coordinates $(v_{j1}, \ldots, v_{jr})$ relative to a convenient basis for M, then ϕ can be expressed as the map

$$\phi : (\mathbb{C}^*)^n \to (\mathbb{C}^*)^r, \qquad (t_1, \ldots, t_n) \mapsto (\prod_{j=1}^n t_j^{v_{j1}}, \ldots, \prod_{j=1}^n t_j^{v_{jr}}). \tag{7.1}$$

The group G is defined as the kernel of ϕ:

$$G = \mathrm{Ker}\Big(\mathrm{Hom}(\Sigma(1), \mathbb{C}^*) \xrightarrow{\phi} \mathrm{Hom}(M, \mathbb{C}^*)\Big). \tag{7.2}$$

Since $G \subset \mathrm{Hom}(\Sigma(1), \mathbb{C}^*)$, we have $g(v_\rho) \in \mathbb{C}^*$ for each $g \in G$ and $\rho \in \Sigma(1)$. This gives an action of G on $\mathbb{C}^n$ by

$$g \cdot (x_1, \ldots, x_n) = (g(v_1)x_1, \ldots, g(v_n)x_n).$$

It is easy to see that G preserves $\mathbb{C}^n - Z(\Sigma)$. Then set

$$X_\Sigma = (\mathbb{C}^n - Z(\Sigma))/G. \tag{7.3}$$

X_Σ contains the dense open torus $T = (\mathbb{C}^*)^n/G$, which acts on X_Σ by coordinatewise multiplication. It is easy to see that this torus has rank r, so that X_Σ is an r-dimensional toric variety. In fact, there are natural identifications

$$T \simeq N \otimes \mathbb{C}^* \simeq \mathrm{Hom}(M, \mathbb{C}^*).$$

With this identification, Eq. (7.2) can be expressed as $T = (\mathbb{C}^*)^n/G$, and the identification of $T \hookrightarrow X_\Sigma$ is obvious from comparison with Eq. (7.3).

It is not hard to see that X_Σ is compact if and only if the union of the cones $\sigma \in \Sigma$ is equal to all of $N_\mathbb{R}$. This point will be amplified in Sec. 7.2.2.

One of the nice features of toric varieties is that it is easy to describe T-invariant subvarieties. Let $\sigma \in \Sigma$ be a cone generated by edges $\rho_1, \ldots, \rho_k$. To this cone is associated the codimension k subvariety

$$Z_\sigma = \{x \in X_\Sigma \mid x_{\rho_1} = \cdots = x_{\rho_k} = 0\},$$

where the x_i are the homogeneous coordinates of x. Clearly Z_σ is T-invariant, and the assignment $\sigma \mapsto Z_\sigma$ clearly reverses the order of inclusions. It is not hard to see that these are all of the non-empty T-invariant subvarieties of X_Σ. Thus,

Classification of T-invariant subvarieties. The assignment $\sigma \mapsto Z_\sigma$ gives an order reversing correspondence (cones in fan)$\longleftrightarrow$(non-empty T-invariant subvarieties).

Note that, in particular, the edges of Σ are in one-to-one correspondence with the set of T-invariant divisors in X_Σ. In general, if σ is a k-dimensional cone, then Z_σ is an $(r-k)$-dimensional subvariety of X_Σ.

Note also that if a set of edges $\{\rho_1, \ldots, \rho_l\}$ does not span a cone in Σ, then the solution to the equations $x_{\rho_1} = \cdots = x_{\rho_l} = 0$, viewed as equations in $\mathbb{C}^n$, are contained in $Z(\Sigma)$. These equations define the empty set in X_Σ.

Each Z_σ is in fact a toric variety. To construct its fan, simply replace N with the quotient N' of N by the sublattice spanned by $\sigma \cap N$. Then project each cone in Σ, which contains σ as a face to N', to get a new fan in N'.

We now give some examples, some of which were briefly introduced in Sec. 2.2.2. The first two examples are two-dimensional. Note that for a compact two-dimensional toric variety, Σ is completely determined by its edges $\Sigma(1)$.

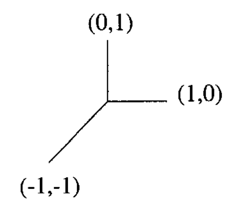


FIGURE 1. The fan for $\mathbb{CP}^2$

EXAMPLE 7.2.3. *We consider $\mathbb{CP}^2$ as a toric variety described by the fan Σ spanned by the three edges $\{(-1,-1), (1,0), (0,1)\}$ as shown in Fig. 1. We will fix this ordering of the edges throughout. This fan will be derived in Sec. 7.2.2, but for now we accept this as given.*

There are seven cones in Σ: the trivial cone $\{0\}$ of dimension 0, the three one-dimensional cones spanned by each of $\{(-1,-1)\}$ $\{(1,0)\}$, and $\{(0,1)\}$, *and the three two-dimensional cones spanned by the sets*

$$\{(1,0),(0,1)\},\ \{(-1,-1),(0,1)\},\{(-1,-1),(1,0)\}.$$

Thus the only set of edges that does not span a cone in Σ *is* $\mathcal{S}=\Sigma(1)=\{(1,0),(0,1),(-1,-1)\}$. *Hence* $Z(\Sigma)=Z(\mathcal{S})=\{(0,0,0)\}\subset\mathbb{C}^3$.

The group G *is defined as the kernel of*

$$\phi:(\mathbb{C}^*)^3\to(\mathbb{C}^*)^2,\qquad (t_1,t_2,t_3)\mapsto(t_1^{-1}t_2,t_1^{-1}t_3).$$

Thus G *is the diagonal group* $\{(t,t,t)\mid t\in\mathbb{C}^*\}\simeq\mathbb{C}^*$. *We immediately recover the usual definition of* $\mathbb{CP}^2$ *as* $(\mathbb{C}^3-\{(0,0,0)\})/\mathbb{C}^*$, *where the* $\mathbb{C}^*$ *acts diagonally on* $\mathbb{C}^3$.

The torus T *defined in Sec. 7.1 is recovered in this context as* $(\mathbb{C}^*)^3/\mathbb{C}^*$, *where* $\mathbb{C}^*$ *is embedded diagonally in* $(\mathbb{C}^*)^3$.

The only non-empty T*-invariant subvarieties are* $\mathbb{CP}^2$ *itself, the coordinate lines, and their pairwise intersections. This can also be seen from toric geometry. We summarize the calculations in Eq. (7.4), where cones are described in terms of generators.*

$$\begin{array}{cc} \sigma & Z_\sigma \\ \{0\} & \mathbb{CP}^2 \\ \{(-1,-1)\} & x_1=0 \\ \{(1,0)\} & x_2=0 \\ \{(0,1)\} & x_3=0 \\ \{(1,0),(0,1)\} & \{(1,0,0)\} \\ \{(-1,-1),(0,1)\} & \{(0,1,0)\} \\ \{(-1,-1),(1,0)\} & \{(0,0,1)\} \end{array} \tag{7.4}$$

The reader can easily check that this correspondence reverses the order of inclusion.

EXAMPLE 7.2.4. *We consider the compact toric variety associated with the fan* Σ *with edges* $\Sigma(1)=\{(1,0),(-1,-n),(0,1),(0,-1)\}$, *shown in Fig. 2. This is the Hirzebruch surface* F_n.

In this example, $\{v_1,v_2\}$ *and* $\{v_3,v_4\}$ *do not span a cone in* Σ, *and any set of edges that does not span a cone in* Σ *must contain at least one of these sets. From this, it follows that* $Z(\Sigma)=\{x_1=x_2=0\}\cup\{x_3=x_4=0\}$. *The*

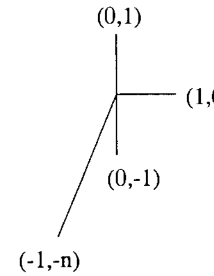


FIGURE 2. The fan for F_n

group G *is the kernel of the map* $\phi:(\mathbb{C}^*)^4\to(\mathbb{C}^*)^2$ *defined by*

$$\phi(t_1,t_2,t_3,t_4)=(t_1t_2^{-1},t_2^{-n}t_3t_4^{-1}).$$

Thus G *can be identified with* $(\mathbb{C}^*)^2$ *via the embedding*

$$(\lambda_1,\lambda_2)\mapsto(\lambda_1,\lambda_1,\lambda_1^n\lambda_2,\lambda_2).$$

There are four T*-invariant divisors* $D_1,\ldots,D_4$ *corresponding to the four edges. Since* $\{v_1,v_2\}$ *does not span a cone in* Σ, *it follows that* D_1 *and* D_2 *are disjoint. Similarly,* D_3 *and* D_4 *are disjoint. All other pairs of these divisors meet in a point, since the corresponding edges span a two-dimensional cone of* Σ.

It is easy to see from the above description that F_n *is a* $\mathbb{CP}^1$ *bundle over* $\mathbb{CP}^1$. *Simply define* $F_n\to\mathbb{CP}^1$ *by* $(t_1,t_2,t_3,t_4)\mapsto(t_1,t_2)$. *A glance at the action of* G *shows that this mapping is well defined, and the fibers of* ϕ *are immediately seen to be isomorphic to* $\mathbb{CP}^1$ *as well. The fibers over* $(1,0)$ *and* $(0,1)$ *are respectively* D_1 *and* D_2.

If $n=0$, *there is a well-defined projection onto the* $\mathbb{CP}^1$ *with coordinates* (t_3,t_4), *and it follows quickly that* F_0 *is simply* $\mathbb{CP}^1\times\mathbb{CP}^1$. *The divisors* D_3 *and* D_4 *are fibers of the second projection in this case.*

In Sec. 7.4 we will calculate $D_4^2=-n$ to see that the different F_n have different geometries. In Sec. 7.7 we will see how to recognize the map $F_n\to\mathbb{CP}^1$ directly from the fan. The more general toric construction of projective bundles is very useful in string theory, for example, in constructing F-theory compactifications.

EXAMPLE 7.2.5. *Consider the total space of the bundle* $\mathcal{O}(-3)$ *on* $\mathbb{CP}^2$. *We have already seen that* $\mathbb{CP}^2$ *contains the torus* $(\mathbb{C}^*)^2$. *Restricting* $\mathcal{O}(-3)$

over this $(\mathbb{C}^*)^2$ *subset, then removing the zero section, we get a torus* $T = (\mathbb{C}^*)^3$. *It is easy to define an action of* T *on* $\mathcal{O}(-3)$, *hence* $\mathcal{O}(-3)$ *is a three-dimensional toric variety.*

Let us construct $\mathcal{O}(-3)$ *from a fan* Σ *in* $\mathbb{R}^3$. *We put*

$$\Sigma(1) = \{(1,0,1),(0,1,1),(-1,-1,1),(0,0,1)\}.$$

The convex hull of $\Sigma(1)$ *is a triangle in the plane* $z = 1$ *with vertices* $\{v_1, v_2, v_3\}$; v_4 *lies in the interior of this triangle, subdividing it into three smaller triangles. The three-dimensional cones in* Σ *are the cones over these triangles. The remaining cones in* Σ *consist of the faces of these cones. Note that these cones do not span* $\mathbb{R}^3$, *which is consistent with the fact that* $\mathcal{O}(-3)$ *is not compact. (Non-compact toric varieties similar to this one are useful in geometric engineering, which is discussed in Sec. 7.8.)*

We compute that $Z(\Sigma) = \{(x_1 = x_2 = x_3 = 0\}$. *Also,* G *is the kernel of*

$$\phi : (\mathbb{C}^*)^4 \to (\mathbb{C}^*)^3, \qquad (t_1,t_2,t_3,t_4) \mapsto (t_1t_3^{-1}, t_2t_3^{-1}, t_1t_2t_3t_4),$$

so that

$$G = \{(t,t,t,t^{-3})\}, \tag{7.5}$$

which is isomorphic to $\mathbb{C}^*$.

There are four T*-invariant divisors* $D_1, \ldots, D_4$. *Since* $\{v_1, v_2, v_3\}$ *do not span a cone in* Σ, *the divisors* D_1, D_2, D_3 *have an empty intersection. All other triples of divisors meet in a point, since the corresponding edges span a three-dimensional cone. The toric description of the* D_i *(see the discussion preceding Example 7.2.3) shows immediately that* D_4 *is compact, while the other* D_i *are non-compact.*

Projection to the first three factors gives a map $X_\Sigma \to \mathbb{CP}^2$ *whose fibers are isomorphic to* $\mathbb{C}$, *so* X_Σ *is a line bundle over* $\mathbb{CP}^2$, *as claimed. The divisor* D_4 *is identified with the zero section of the bundle, and the divisors* D_1, D_2, D_3 *are the restrictions of these bundles over the corresponding coordinate lines* $x_1 = 0,\ x_2 = 0,\ x_3 = 0$.

In Sec. 7.4, we will be able to identify that the bundle is indeed $\mathcal{O}(-3)$.

At this point, we make contact with the construction $X_\Sigma = \cup_{\sigma\in\Sigma} X_\sigma$. Let $\sigma \in \Sigma$ be an r-dimensional cone. In our context, we can *define* $X_\sigma \subset X_\Sigma$ as the subset obtained by setting $x_\rho = 1$ for all $\rho \in \Sigma(1)$ that are not edges of σ. It can be seen that this agrees with the usual definition.

Now let $\Sigma_\sigma \subset \Sigma$ be the fan consisting of σ and all of its faces. Then it is straightforward to check that $X_{\Sigma_\sigma} \simeq X_\sigma$. This is a useful way to do local calculations.

We close this section by giving a useful criterion for smoothness.

Smoothness Criterion. A toric variety X_Σ is smooth if and only if each cone $\sigma \in \Sigma$ is generated by a $\mathbb{Z}$-basis for the intersection of the linear span of σ with N.

PROOF. Consider a top-dimensional cone $\sigma \in \Sigma$, and form X_{Σ_σ} with the above property. Then the group G for X_{Σ_σ} is trivial, and $X_{\Sigma_\sigma} \simeq X_\sigma \simeq \mathbb{C}^r$. So X_Σ is locally smooth, hence smooth.

The converse is readily explained using $X_\sigma = \operatorname{Spec}\mathbb{C}[\check{\sigma}\cap M]$ to explicitly identify generators of the maximal ideal of X_σ at its origin with generators of the dual cone $\check{\sigma}$. Details are left to the reader. □

It is easy to check that all the fans given above satisfy this criterion, hence all the toric varieties are smooth.

7.2.2. Constructing Fans from Toric Varieties. In Sec. 7.2.1 we saw how much information can be read off from the fan. In this section, we explain how to construct the fan from a given normal toric variety.

The key idea is a slight modification of the description of the order-reversing correspondence given in Sec. 7.2.1. The new element is a description of the T-invariant subvarieties as closures of T-orbits.

Let us start with a toric variety X containing the torus $T \simeq (\mathbb{C}^*)^r$. We consider the lattice $N = \operatorname{Hom}(\mathbb{C}^*, T) \cong \mathbb{Z}^r$ and construct a fan $N_\mathbb{R}$.

Elements of N are homomorphisms $\psi : \mathbb{C}^* \to T$, which are called one-parameter subgroups. If we identify T with $(\mathbb{C}^*)^r$, we fix the identification

$$\mathbb{Z}^r \simeq N, \qquad (a_1, \ldots, a_r) \mapsto (t \mapsto (t^{a_1}, \ldots, t^{a_r}))\,.$$

Now let ψ be a one-parameter subgroup, and consider the induced map $f : \mathbb{C}^* \to X$ defined as $f(t) = \psi(t) \cdot 1$, where 1 denotes the identity element of T. The image of f is entirely contained in T. Suppose that $\lim_{t\to 0} f(t)$ exists in X. Then the orbit closure $Z_\psi = \overline{T \cdot \lim_{t\to 0} f(t)}$ is a non-empty T-invariant subvariety of X. From Sec. 7.2.1, we expect that there is a corresponding cone hiding somewhere in this description.

The extraction of the cone is simple. Consider the set of all ψ for which Z_ψ exists. On this set, we define the equivalence relation $\psi \simeq \psi'$ if $Z_\psi = Z_{\psi'}$.

Fixing an equivalence class, we take the closure of the convex hull in $N_{\mathbb{R}}$ of all one-parameter subgroups in the fixed equivalence class. This gives a cone. The collection of all cones obtained in this manner forms a fan Σ, and $X \simeq X_\Sigma$. We illustrate this below when we revisit Examples 7.1.2 and 7.2.3.

Here is a clarifying consequence of this construction. Suppose that $\psi \in N$ is contained in some cone of a fan Σ. Then there is precisely one cone $\sigma \in \Sigma$ such that ψ is contained in the relative interior of σ. For this σ, we have $Z_\psi = Z_\sigma$.

Note how this explains the compactness criterion for toric varieties as follows. Suppose that the union of the cones in Σ is a proper subset of $N_{\mathbb{R}}$, and let ψ be a one-parameter subgroup not contained in this set. Then $\psi(t)$ does not have a limit in X_Σ as $t \to 0$, so X_Σ cannot be compact.

We illustrate by continuing with Examples 7.1.2 and 7.2.3.

EXAMPLES 7.1.2 AND 7.2.3 REVISITED. *We will start from scratch with the description given in Example 7.1.2 of $\mathbb{CP}^2$ as a toric variety. We can use the rescaling of coordinates in $\mathbb{CP}^2$ to set the first coordinate of an element of T to 1. This identifies*

$$T = \{(1, t_1, t_2) \mid t_i \in \mathbb{C}^*\} \simeq (\mathbb{C}^*)^2.$$

We will use this isomorphism $T \simeq (\mathbb{C}^*)^2$ and the above construction of the fan to derive the fan for $\mathbb{CP}^2$ given in Example 7.2.3.

With this identification, the torus action is given by $(t_1, t_2)\cdot(x_1, x_2, x_3) = (x_1, t_1x_2, t_2x_3)$ for $(t_1, t_2) \in T = (\mathbb{C}^*)^2$ and $(x_1, x_2, x_3) \in \mathbb{CP}^2$.

The 1-parameter subgroups of T are indexed as above by $(a, b) \in \mathbb{Z}^2$, which represents the one-parameter subgroup $\psi_{a,b}(t) = (t^a, t^b) \in N = \mathrm{Hom}(\mathbb{C}^*, T)$.

Using the embedding of T in $\mathbb{CP}^2$, we can study $\lim_{t\to 0} \psi(t) \in \mathbb{CP}^2$. There are seven possibilities for these limit points and their orbit closures.

(7.6)

	$\lim_{t\to 0}\psi(t)$	closure of orbit of $\lim_{t\to 0}\psi(t)$
$a > 0,\ b > 0$	$(1,0,0)$	$\{(1,0,0)\}$
$a < 0,\ b > a$	$(0,1,0)$	$\{(0,1,0)\}$
$b < 0,\ b < a$	$(0,0,1)$	$\{(0,0,1)\}$
$a = b < 0$	$(0,1,1)$	$\{x_1 = 0\}$
$a > 0,\ b = 0$	$(1,0,1)$	$\{x_2 = 0\}$
$a = 0,\ b > 0$	$(1,1,0)$	$\{x_3 = 0\}$
$a = b = 0$	$(1,1,1)$	$\mathbb{CP}^2$

A pictorial description is given in Fig. 3, with limit points indicated. The closures of the regions defined in the first column of Eq. (7.6) define the seven cones in the fan for $\mathbb{CP}^2$ given in Fig. 1. Note that we have also recovered the correspondence between cones and non-empty T-invariant subvarieties given in Eq. (7.4).

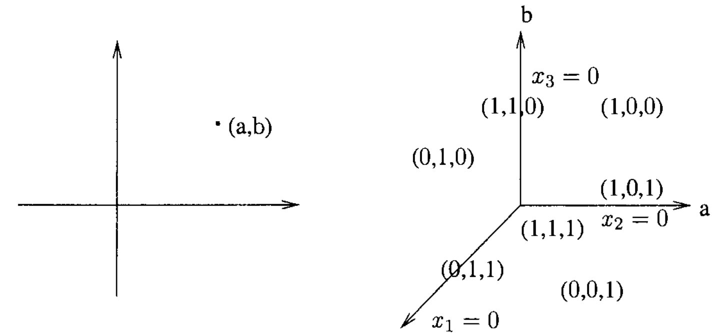


FIGURE 3. One-parameter subgroups and limit points for $\mathbb{CP}^2$

We close this section by giving an example explaining the need to restrict to normal toric varieties.

EXAMPLE 7.2.6. *Let $X \subset \mathbb{CP}^2$ be the plane curve defined by the equation $x_1x_2^2 = x_3^3$. This has a non-normal singularity at $(1, 0, 0)$, but it is a toric variety: The torus $T = \mathbb{C}^*$ is embedded in X via $t \mapsto (1, t^3, t^2)$. If we attempt to apply the above construction of a fan, we get the one-dimensional fan with two edges generated by $\{1\}$ and by $\{-1\}$. But this is the fan for $\mathbb{CP}^1$, not X. The intrinsic reason for the occurrence of $\mathbb{CP}^1$ is that $\mathbb{CP}^1$ is the normalization of X via the map $(x_1, x_2) \mapsto (x_1^3, x_2^3, x_1x_2^2)$.*

7.3. GLSM

The *gauged linear sigma model* (GLSM) is a two-dimensional gauge theory. We will explore gauge theories in more detail in Sec. 15.2. For present purposes, we restrict our attention to theories without a superpotential.

We consider a two-dimensional $U(1)^s$ gauge theory with vector superfields $V_1, \ldots, V_s$, and n chiral superfields $\Phi_1, \ldots, \Phi_n$. The charge of Φ_i under the a^{th} $U(1)$ will be denoted by $Q_{i,a}$, and the scalar component of Φ_i will be denoted by ϕ_i. The Lagrangian has the form

$$L = L_{\text{kin}} + L_{\text{gauge}} + L_{D,\theta}, \tag{7.7}$$

where the three terms are respectively the kinetic energy of the chiral superfields, the kinetic energy of the gauge fields, and a Fayet–Iliopoulos (FI) term and theta angle. Rather than describe these terms, we content ourselves with writing down the potential energy deduced from Eq. (7.7):

$$U(\phi_i) = \sum_{a=1}^{s} \frac{e_a^2}{2} \left(\sum_{i=1}^{n} Q_{i,a} |\phi_i|^2 - r_a \right)^2. \tag{7.8}$$

Here, the e_a are the gauge couplings and the r_i are real parameters ("FI parameters").

To find the supersymmetric ground states of this theory, we set the gauge fields to zero and find the zeros of the potential energy. This gives the system of equations

$$\sum_{i=1}^{n} Q_{i,a} |\phi_i|^2 = r_a, \qquad a = 1, \ldots, s. \tag{7.9}$$

The supersymmetric ground states are parametrized by the solutions of Eq. (7.9) modulo gauge equivalence.

Main Point. For general charge assignments and appropriate choice of FI parameters, the space of supersymmetric ground states is an $(n-s)$-dimensional normal toric variety whose fan has n edges.

We defer the geometric characterization of the fan to Sec. 7.4, where we will identify the charges $Q_{i,a}$ with certain intersection numbers.

We prepare to construct a fan Σ. First define the subgroup $G = (\mathbb{C}^*)^s \subset (\mathbb{C}^*)^n$ by the embedding

$$(t_1, \ldots, t_s) \mapsto \left(\prod_{a=1}^{s} t_a^{Q_{1,a}}, \ldots, \prod_{a=1}^{s} t_a^{Q_{n,a}} \right). \tag{7.10}$$

The torus is given by $T = (\mathbb{C}^*)^n / G$.

It is easy to see (essentially linear algebra) that we can choose a collection $S = (v_1, \ldots, v_n)$ of elements of N such that replacing $\Sigma(1)$ by S in the definition Eq. (7.2) of G as a subgroup of $(\mathbb{C}^*)^n$ yields Eq. (7.10). Note that the v_i have not been assumed distinct, although they will be distinct for most charge assignments. For ease of exposition, we assume that the v_i are distinct, and we let $\Sigma(1)$ be the set consisting of all elements of the collection S.

We now describe a fan Σ, assuming that the FI parameters have been appropriately chosen. For now, we take "appropriately chosen" to mean that there are sufficiently many solutions to Eq. (7.9), so that the set of solutions of Eq. (7.9) with all $\phi_i \neq 0$ modulo gauge equivalence projects surjectively onto $T = (\mathbb{C}^*)^n / G$. We will explain this condition geometrically in Sec. 7.4.

We consider all subsets $P = \{v_{i_1}, \ldots, v_{i_k}\} \subset S$ such that there are no solutions of Eq. (7.9) with $\phi_{i_1} = \cdots = \phi_{i_k} = 0$. If there are any such P consisting of a single element v, let $\Sigma(1) \subset S$ be the set obtained from S by removing all of these v. Then it can be shown that there is a unique fan Σ with edges equal to $\Sigma(1)$ with the following property: The subsets of $\Sigma(1)$ that do not span a cone of Σ are precisely those subsets P considered above that are subsets of $\Sigma(1)$.

We assert that for this Σ, the toric variety X_Σ is precisely the space of supersymmetric ground states. We do not explain the details here, but remark that the assertion is essentially a reformulation of the construction of toric varieties by symplectic reduction.

Note that the fan can depend on the choice of FI parameters. In such a case, the toric varieties can be related by birational transformations such as blow-ups or flops.

There may also be values of the FI parameters for which the space of supersymmetric ground states is not a toric variety. It can even be empty, as in the case of a $U(1)$ gauge theory with charges $(1,1)$. In that case, Eq. (7.9) reads

$$|\phi_1|^2 + |\phi_2|^2 = r,$$

which clearly has no solutions if $r < 0$.

The dependence of the theory on the FI parameters can be understood in terms of the *GKZ decomposition*.

Note that for general toric varieties, the group G need not be $(\mathbb{C}^*)^s$, as it may contain finite groups as factors. We would need an orbifold to produce such toric varieties as a space of supersymmetric ground states. We will return to this point in Sec. 7.5.

EXAMPLE 7.3.1. *We consider a $U(1)$ gauge theory with four chiral superfields with respective charges $(1,1,1,-3)$. We have already found a fan that produces the required group G: the fan in Example 7.2.5, yielding the group G given in Eq. (7.5).*

We have in this case a single FI parameter r. Then Eq. (7.9) in this case becomes the single equation

$$|\phi_1^2| + |\phi_2^2| + |\phi_3|^2 - 3|\phi_4|^2 = r.$$

If $r > 0$, then we cannot have $\phi_1 = \phi_2 = \phi_3 = 0$. This determines the fan Σ to be the fan of Example 7.2.5 (see especially the determination of $Z(\Sigma)$). So the space of supersymmetric ground states is the total space of the bundle $\mathcal{O}(-3)$ on $\mathbb{CP}^2$.

If $r < 0$, then we cannot have $\phi_4 = 0$. Here we do not get a fan with four edges; instead we remove the fourth edge generated by $(0,0,1)$ and get a cone over a triangle. As we will see in Example 7.6.3, this is a $\mathbb{Z}_3$ orbifold of $\mathbb{C}^3$, and the bundle $\mathcal{O}(-3)$ is obtained by blowing up this singularity.

7.4. Intersection Numbers and Charges

We begin this section by explaining how the charges in the GLSM are related to the toric variety of supersymmetric ground states. Later in this section we will relate the charges to intersection numbers in the toric variety. For ease of exposition, we assume that the toric variety is smooth.

Suppose we start with a GLSM with gauge group $U(1)^s$ and n chiral superfields $\Phi_1, \ldots, \Phi_n$. We use the construction of Sec. 7.3 to obtain a set $\Sigma(1) = \{v_1, \ldots, v_n\}$ of edges, where for ease of exposition we have assumed that $\Sigma(1) = S$ in the terminology of Sec. 7.3. By construction, the $Q_{i,a}$ are the relations among the v_i, i.e.,

$$\sum_{i=1}^{n} Q_{i,a} v_i = 0, \qquad a = 1, \ldots, s.$$

Conversely, if we start with the set $\Sigma(1)$, we can form the rank s lattice Λ of all $\mathbb{Z}$-linear relations among the $\{v_i\}$. A basis for Λ is a collection of relations

$$\sum_{i=1}^{n} Q'_{i,a} v_i = 0, \qquad a = 1, \ldots, s.$$

It is clear from linear algebra that the $Q'_{i,a}$ are precisely the charges of the original superfields Φ_i, with the understanding that the gauge group G may need to be written as a product of s copies of $U(1)$ in a different way, depending on the choice of basis for Λ.

EXAMPLE 7.4.1. *We look at $\mathbb{CP}^2$ again with the fan given in Fig. 1. The generators of the edges satisfy the linear relation*

$$1(1,0) + 1(0,1) + 1(-1,-1) = 0,$$

which generates the lattice of relations in this case. It is easy to see that $\mathbb{CP}^2$ arises as the space of supersymmetric ground states of a $U(1)$ GLSM with three chiral superfields with charge vector $(1,1,1)$.

Here and in what follows, it is convenient to organize the data in two matrices $P|Q$:

$$\left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{array}\right)$$

In general, row vectors of P are generators of the edges, and column vectors of Q are generators of the lattice Λ of relations. Each row corresponds to a field in the GLSM, and each column in Q corresponds to a $U(1)$ charge.

EXAMPLE 7.4.2. *We next turn to F_n given by the fan in Fig. 2. The lattice of relations is given by the matrices*

$$P|Q = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ -1 & -n & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & -n \end{array}\right)$$

This toric variety is therefore the space of supersymmetric ground states of a $U(1)^2$ gauge theory with four chiral superfields, having respective charges $(0,1)$, $(0,1)$, $(1,0)$, and $(1,-n)$, as can be checked directly.

We now describe the relationship between charges and intersection numbers. Let X_Σ be a toric variety. For each $\rho \in \Sigma(1)$, we let D_ρ be the T-invariant divisor Z_ρ (we have changed the symbol Z to D to emphasize that these are divisors). Note that

$$\mathbb{Z}^{\Sigma(1)} \simeq \oplus_{\rho \in \Sigma(1)} \mathbb{Z} \cdot D_\rho.$$

Each character $m \in M$ may be viewed as a holomorphic function on T. Its extension to X_Σ need not be holomorphic but is certainly at least a rational function. The zeros and poles of this rational function define the principal divisor

$$(m) = \sum_{\rho \in \Sigma(1)} \langle m, v_\rho \rangle D_\rho$$

which is naturally viewed as an element of $Z^{\Sigma(1)}$. We thus have a map $M \to \mathbb{Z}^{\Sigma(1)}$, which is an inclusion if the one-dimensional cones span $N_{\mathbb{R}}$. This inclusion map is given by a matrix P whose row vectors are the v_ρ with $\rho \in \Sigma(1)$. The examples in Sec. 7.4 give examples of such matrices P.

Here is the main result we need about divisors and divisor classes:

THEOREM 7.4.3. *$\Sigma a_\rho D_\rho$ and $\Sigma a'_\rho D_\rho$ are linearly equivalent*
$\Leftrightarrow$ They are homologically equivalent
$\Leftrightarrow$ They define the same line bundle
$\Leftrightarrow$ They differ by (m) for some $m \in M$

PROOF(SKETCH). If $\Sigma a_\rho D_\rho$ and $\Sigma a'_\rho D_\rho$ differ by (m), then they are linearly equivalent by definition. Linear equivalence of divisors D and D' is the same condition as $\mathcal{O}(D) \simeq \mathcal{O}(D')$ for any variety. Since the homology class $[D]$ of a divisor is the topological first Chern class $c_1(\mathcal{O}(D))$, it follows that linearly equivalent divisors are homologically equivalent. Proofs of the other equivalences will be omitted. □

Part of the assertion of Theorem 7.4.3 can be strengthened: It is a fact that any divisor is linearly equivalent to a T-invariant divisor. We therefore have an exact sequence

$$0 \to M \to \mathbb{Z}^{\Sigma(1)} \to A_{r-1}(X_\Sigma) \to 0, \tag{7.11}$$

where $A_{r-1}(X_\Sigma)$ is the Chow group of *all* divisors modulo linear equivalence. We see that $A_{r-1}(X_\Sigma)$ is a finitely generated abelian group of the form $\mathbb{Z}^s \oplus H$, where $s = n - r$ and H is a finite sum (possibly empty) of finite groups $\mathbb{Z}_{n_j}$. In particular, $A_{r-1}(X_\Sigma)/\text{torsion} \simeq \mathbb{Z}^s$.

The Chow group $A_k(X_\Sigma)$ of k-dimensional cycles modulo rational equivalence is also easy to describe from the toric data for any k, but we do not need this here.

Let us now apply $\text{Hom}(-, \mathbb{C}^*)$ to the exact sequence Eq. (7.11). We get an exact sequence

$$0 \to \text{Hom}(A_{r-1}(X_\Sigma), \mathbb{C}^*) \to \text{Hom}(\mathbb{Z}^{\Sigma(1)}, \mathbb{C}^*) \to \text{Hom}(M, \mathbb{C}^*) \to 0. \tag{7.12}$$

Note that the surjection in Eq. (7.12) is naturally identified with the map ϕ from Eq. (7.1). Comparing Eq. (7.12) with the definition Eq. (7.2) of G, we see that $G \simeq \text{Hom}(A_{r-1}(X_\Sigma), \mathbb{C}^*)$.

Recall from Sec. 7.3 that for toric varieties arising from the GLSM, we will get $G \simeq (\mathbb{C}^*)^s$. This will require that $A_{r-1}(X)$ has no torsion, $H = 0$. If H is nonzero, G acquires a finite abelian factor. For the rest of this section, we will assume that $H = 0$ and consequently $G \simeq (\mathbb{C}^*)^s$. The general situation can be dealt with as an orbifold of the special case considered here. Orbifolds will be considered in Sec. 7.5.

The key observation is that the exponents of the inclusion

$$G \hookrightarrow \text{Hom}(\mathbb{Z}^{\Sigma(1)}, \mathbb{C}^*) \simeq (\mathbb{C}^*)^{\Sigma(1)} \tag{7.13}$$

are given by the matrix Q whose column vectors are generators of the lattice Λ of relations. More precisely, identifying G with $(\mathbb{C}^*)^s$, Eq. (7.13) is given by the embedding Eq. (7.10). As discussed earlier in this section, Q can be identified with the charge matrix of the corresponding GLSM.

An element $(Q'_{1,a}, \ldots, Q'_{n,a}) \in \Lambda$ can be viewed as a linear functional on $Z^{\Sigma(1)}$ which takes the basis element ϕ_i to $Q'_{i,a}$. This functional annihilates the image of M in $\mathbb{Z}^{\Sigma(1)}$. By Eq. (7.11), it can therefore be viewed as an element in $\text{Hom}(A_{r-1}(X), \mathbb{Z})$, which is isomorphic to $H_2(X, \mathbb{Z})$.

This gives a practical guide to computations. The columns of Q correspond to a basis for Λ, i.e., to a basis for $H_2(X, \mathbb{Z})$. The rows of Q correspond to the T-invariant divisors $D_1, \ldots, D_n$. Since we are free to choose a convenient basis for Λ, we usually choose a basis of homology classes of irreducible curves $C_1, \ldots, C_s$. Unwinding the definitions, we conclude that

$$Q_{i,a} = D_i \cdot C_a. \tag{7.14}$$

For applications to mirror symmetry, it is best to choose the C_a to form a generating set for the Mori cone of classes of effective curves when this is possible. There is a systematic way to find generators.

THEOREM 7.4.4. *The Mori cone (the cone of effective one-cycles) is spanned by curves corresponding to $(r-1)$-dimensional cones.*

PROOF. See [**219**, Prop. 1.6]. □

A convenient interpretation of the intersection numbers in Eq. (7.14) is to use intersections with the C_j to put coordinates on the Chow group

$A_{r-1}(X_\Sigma)$. Then the intersection numbers in the i^{th} row of Q are coordinates of the divisor D_i in the Chow group.

We now relate this discussion to the GLSM, as promised earlier. Suppose we start with a charge matrix Q and choose a set of edges $S = \{v_1, \ldots, v_n\}$ as in Sec. 7.3. Note that $A_{n-1}(X_\Sigma)$ only depends on S, not on the actual fan Σ with $\Sigma(1) \subset S$, so we will denote this common Chow group by $A_{n-1}(S)$.

It is straightforward to see that the FI parameters naturally live in the Chow group $A_{n-1}(S)$: The assignment of an FI parameter to a charge vector is naturally an element of Λ^*, and we have already seen that Λ is dual to $A_{n-1}(X_\Sigma)$.

The divisor classes of the T-invariant divisors D_i span a cone $A^+_{n-1}(S) \otimes \mathbb{R} \subset A_{n-1}(S) \otimes \mathbb{R}$. If the FI parameters are chosen to lie in the interior of $A^+_{n-1}(S) \otimes \mathbb{R}$ using the identification described in the preceding paragraph, then the space of supersymmetric ground states forms a toric variety. This is the precise version of what we meant in Sec. 7.3 when we said that the FI parameters need to be "appropriately chosen." The GKZ decomposition alluded to earlier is a decomposition of $A^+_{n-1}(S) \otimes \mathbb{R}$ into subcones. We get different toric varieties of supersymmetric ground states when the FI parameters are picked in the interiors of different cones in the GKZ decomposition.

If, in addition, the FI parameters are chosen to lie in the Kähler cone of X_Σ, then the toric variety of supersymmetric ground states is precisely X_Σ. If we choose a basis for Λ that generates the Mori cone, then the condition that the FI parameters lie in the Kähler cone is simply the condition that all r_i are positive.

We now return to our examples.

EXAMPLE 7.4.1, REVISITED. *We rewrite the matrices $P|Q$ as*

$$\left(\begin{array}{ccc|c} \rho_1 & 1 & 0 & 1 \\ \rho_2 & 0 & 1 & 1 \\ \rho_3 & -1 & -1 & 1 \end{array}\right)$$

labeling the rows by the three edges ρ_i.

We see that $A_1(\mathbb{CP}^2) \simeq \mathbb{Z}$ and that the three coordinate lines associated to edges ρ_1, ρ_2, ρ_3 are in the same class in $A_1(\mathbb{CP}^2)$, the class of a line in $\mathbb{CP}^2$. The column of Q corresponds to the class L of a line as well; the fact that each entry of Q is 1 follows from the equality $D_i \cdot L = 1,\ i = 1, 2, 3$.

EXAMPLE 7.4.2, REVISITED. *We return to F_n. As we have seen in Example 7.2.4, the divisors D_1 and D_2 are fibers of F_n viewed as a $\mathbb{CP}^1$ bundle over $\mathbb{CP}^1$. We denote their common cohomology class by f. We let D_3 and D_4 have cohomology classes H and E respectively, as in Fig. 4. The configuration of the four divisors is also shown in Fig. 4. The divisors are also curves since F_n is two-dimensional.*

From the choice of coordinates, we see that $H = E + nf$. Thus the Mori cone is generated by f and E. We use these for the columns of Q. The intersection numbers in the first column are immediate from the geometry shown in Fig. 4: Clearly $f^2 = 0$, while $f \cdot H = D_1 \cdot D_3 = 1$ and $f \cdot E = D_1 \cdot D_4 = 1$. For the second column, all intersection numbers with E are clear, except E^2 but this can be calculated since $E \cdot E = E \cdot (H - nf) = 0 - n = -n$.

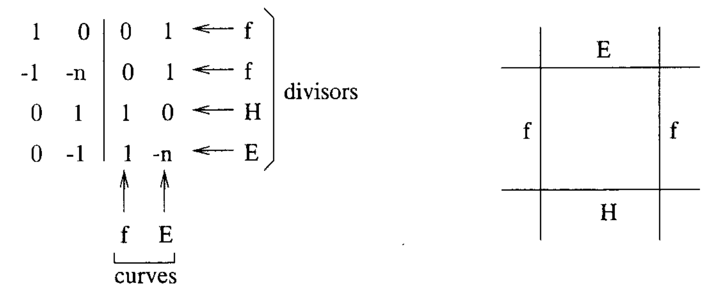
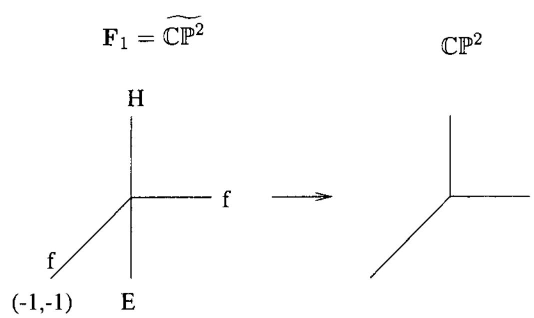


FIGURE 4. Divisors and intersections on F_n

If $n > 0$, the existence of a curve E with self-intersection number $-n$ shows that $F_n \neq \mathbb{CP}^1 \times \mathbb{CP}^1$.

In particular, if $n = 1$, then $E^2 = -1$, so that E is an exceptional divisor and can be blown down to a point on a smooth surface. Using toric geometry we will see in Sec. 7.6 that F_1 is $\mathbb{CP}^2$ blown up at a point. As a sneak preview, note that the fan for F_1 can be obtained from the fan for $\mathbb{CP}^2$ by inserting the edge corresponding to E and then subdividing the fan. More generally, we will see that subdividing a fan corresponds to blowing up.



EXAMPLE 7.4.5. *The same considerations hold even if X_Σ is not compact. We again take up the bundle $\mathcal{O}(-3)$ over $\mathbb{CP}^2$ from Example 7.2.5. For the matrices $P|Q$, we get*

$$\begin{array}{cl} \left(\begin{array}{ccc|c} -1 & -1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -3 \end{array}\right) & \begin{array}{l} \leftarrow x_1 = 0 \\ \leftarrow x_2 = 0 \\ \leftarrow x_3 = 0 \\ \leftarrow \textit{zero section} \end{array} \\ \qquad\qquad\qquad\quad \uparrow & \\ \qquad\qquad\qquad\quad C & \end{array}$$

Note that the charges $(1, 1, 1, -3)$ coincide with those of the GLSM considered in Example 7.3.1, as they must. Here the curve C associated with the column of Q is the zero section over a line in $\mathbb{CP}^2$. It clearly intersects each of the first three divisors at one point.

Recall from Example 7.2.5 that X_Σ is a line bundle over $\mathbb{CP}^2$. We now show conclusively that this bundle is $\mathcal{O}(-3)$, as claimed. Note that since C is contained in the zero section D_4, the intersection $C \cdot D_4$ is given by the degree of the normal bundle of D_4, restricted to C. Since $C \cdot D_4 = -3$, we conclude that the bundle is indeed $\mathcal{O}(-3)$, as claimed.

Finally, we note that for each of the examples in this section we have chosen our basis of Λ to correspond to generators of the Mori cone. Therefore, each of these toric varieties arises as the space of supersymmetric ground states of the GLSM with indicated charge vectors if we choose positive FI parameters.

7.5. Orbifolds

In this section we show how to analyze orbifolds.

DEFINITION 7.5.1. *A rational polyhedral cone is* simplicial *if it can be generated by a set of vectors $v_1, \ldots, v_k$, which form a basis for the vector space that they span. A fan Σ is* simplicial *if each cone in Σ is simplicial.*

We can now state the extension of the smoothness criterion to a criterion for orbifolds.

Orbifold criterion. A toric variety is an orbifold if and only if its fan is simplicial.

PROOF. Consider an r-dimensional cone $\sigma \in \Sigma$ generated by $v_1, \ldots, v_r$. Then we compute that G for X_{Σ_σ} is a finite group, so that $X_\sigma \simeq \mathbb{C}^r/G$ is an orbifold. Hence X_Σ is an orbifold.

The converse is non-trivial, but follows from the following statement in the literature: If X is a rationally smooth algebraic variety of dimension r admitting an action of a torus T with an isolated fixed point x and only finitely many T-invariant (closed irreducible) curves, then the number of such curves containing x equals r. If X_Σ is a toric orbifold and $\sigma \in \Sigma$ is an r-dimensional cone, then the point $x = Z_\sigma$ satisfies the stated hypothesis. Identifying the T-invariant curves containing x with the codimension 1 faces of σ, we conclude that σ has r codimension 1 faces, hence is simplicial. □

REMARK 7.5.2. *Intrinsically, G is the quotient of N by the sublattice generated by the v_i.*

We now consider certain global orbifolds. Suppose we have a simplicial fan, and in addition suppose that there is a sublattice $N' \subset N$ such that all top-dimensional cones in Σ are generated by a $\mathbb{Z}$ basis for N'.

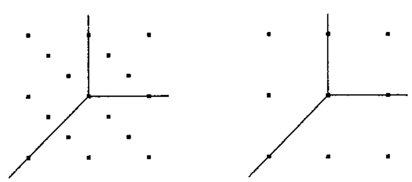
Since $N'_R = N_\mathbb{R}$, we can view Σ as a fan in $N'_\mathbb{R}$, obtaining an auxiliary toric variety $X_{\Sigma,N'}$ which is smooth. Note, however, that the torus has changed: We must take $T' = N' \otimes \mathbb{C}^*$. The natural map $T' \to T = N \otimes \mathbb{C}^*$ induced by the inclusion of N' in N is easily seen to be a finite quotient mapping (this is clear in coordinates). It is therefore not surprising that the original toric variety $X_\Sigma = X_{\Sigma,N}$ defined using the lattice N is a global orbifold of the smooth $X_{\Sigma,N'}$ by the finite group N/N'.

We give an example, which also illustrates how to describe certain orbifolds by toric geometry.

EXAMPLE 7.5.3. *We consider a particular $\mathbb{Z}_3$ orbifold of $\mathbb{CP}^2$ (chosen because it will be used to construct the mirror of plane cubic curves in Sec. 7.10).*

Recall that the embedding of $T = (\mathbb{C}^*)^2$ in $\mathbb{CP}^2$ is given by $(t_1, t_2) \in (\mathbb{C}^*)^2 \mapsto (1, t_1, t_2) \in \mathbb{CP}^2$. Consider the $\mathbb{Z}_3$ subgroup of $(\mathbb{C}^*)^2$ generated by (ω, ω^2), where $\omega = e^{2\pi i/3}$. This generator extends to act on $\mathbb{CP}^2$ as coordinatewise multiplication by $(1, \omega, \omega^2)$.

To construct a fan for $\mathbb{CP}^2/\mathbb{Z}_3$, the quotient of $\mathbb{CP}^2$ by this subgroup, we must first understand the torus $T' = T/\mathbb{Z}_3$ (note that T and T' are interchanged when comparing to the above general discussion). Observe that $t \mapsto (1, t^{1/3}, t^{2/3})$ is a well-defined one-parameter subgroup of T' which cannot be lifted to T, so that the lattice N' of one-parameter subgroups of T' is strictly larger than N. It is easy to see that $N' = N + \mathbb{Z}(1/3, 2/3)$. We simply take the same fan Σ, drawn relative to the lattice N' rather than N. These two fans are pictured in Fig. 5. The toric variety $X_{\Sigma,N'}$ is the orbifold $\mathbb{CP}^2/\mathbb{Z}_3$.



N' N

FIGURE 5. The fans for $\mathbb{CP}^2/\mathbb{Z}_3$ and $\mathbb{CP}^2$

The generators of the one-dimensional cones are $(-1,-1), (1,0), (0,1)$. If we change to coordinates in $N'_\mathbb{R} = N_\mathbb{R}$ adapted to the choice of generators $\{(2/3, 1/3), (1/3, 2/3)\}$ of N', then the generators of the edges have coordinates $(2,-1), (-1,2), (-1,-1)$.

To see the orbifold from this vantage point, consider the cone σ generated by $(2,-1), (-1,2)$. Since $\det\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3$, the vectors $(2,-1)$ and $(-1,2)$ generate a sublattice of N' of index 3, hence X_σ is the affine toric variety $\mathbb{C}^2/\mathbb{Z}_3$.

7.6. Blow-Up

In the F_1 case of Example 7.4.2, we mentioned that blow-ups of a toric variety can be obtained by subdividing the fan. We now explain this in a little more detail.

DEFINITION 7.6.1. *A fan Σ'* subdivides *the fan Σ if*

(1) *$\Sigma(1) \subset \Sigma'(1)$, and*
(2) *each cone of Σ' is contained in some cone of Σ.*

Note that $\Sigma'(1)$ is allowed to equal $\Sigma(1)$. See Example 7.6.4.

Suppose that Σ' subdivides Σ. Let $\Sigma'(1) = \{\rho_1, \ldots, \rho_m\}$, where the edges are ordered so that $\Sigma(1) = \{\rho_1, \ldots, \rho_n\}$. Then we assert that there is a well-defined map $X_{\Sigma'} \to X_\Sigma$ defined in terms of the homogeneous coordinates by projection onto the first n factors.

We need to check that (i) $\mathbb{C}^m - Z(\Sigma')$ projects into $\mathbb{C}^n - Z(\Sigma)$, and (ii) this projection is compatible with the group actions.

Requirement (i) follows immediately from the assumption that Σ' subdivides Σ, and requirement (ii) is easy to check. The map $X_{\Sigma'} \to X_\Sigma$ is clearly birational, since it is an isomorphism on a dense open set (the torus T).

To blow up a T-invariant smooth point $p \in X_\Sigma$, we find the r-dimensional cone $\sigma \in \Sigma$ corresponding to p. If the primitive generators of σ are $v_1, \ldots, v_r$, we add a new edge generated by

$$v_{r+1} = v_1 + \cdots + v_r,$$

and then we subdivide σ. Combining these new cones with the cones of Σ (except σ but including all proper faces of σ) we get a new fan Σ', yielding the blow-up.

In the GLSM, we would add a new field, and an extra $U(1)$ with charges $(1, \ldots, 1, -1, 0, \ldots)$ corresponding to the relation $v_1 + \cdots + v_r - v_{r+1} = 0$.

For general subdivisions, we wind up blowing up more general T-invariant ideals. This ideal is supported on the union of all the T-invariant subvarieties of X_Σ corresponding to cones in Σ that are not cones of Σ'. In the above example of the blow-up of a point, the only cone of Σ that is not a cone of Σ' is σ, so we conclude that the only thing that was blown up is the point Z_σ.

We can now do some interesting examples.

EXAMPLE 7.6.2. *We blow up the orbifold $\mathbb{C}^2/\{\pm 1\}$ at the origin, resolving this A_1 singularity. Using the technique of Example 7.5.3, we can choose coordinates for N so that the fan for $\mathbb{C}^2/\{\pm 1\}$ consists of the cone spanned by $v_1 = (1,0)$ and $v_2 = (1,2)$ as well as its faces. Inserting the edge spanned by $v_3 = (1,1)$ and subdividing, we obtain the fan Σ depicted in Fig. 6. The toric variety X_Σ is smooth and is equal to the blow-up of $\mathbb{C}^2/\{\pm 1\}$ at the singular point. Since v_3 has been added, the divisor D_3 is the exceptional divisor of the blow-up. The relation $v_1 + v_2 - 2v_3 = 0$ gives the charge vector $(1,1,-2)$, leading to $D_3^2 = -2$, the well-known result for the resolution of an A_1 singularity.*

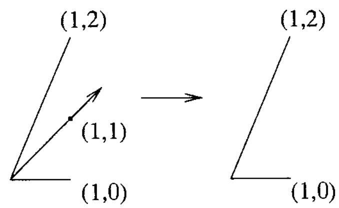


FIGURE 6. The fans for $\mathbb{C}^2/\{\pm 1\}$ and its blow-up

This example can be generalized to give the resolution of an A_n singularity. The A_n singularity can be written as $\mathbb{C}^2/\mathbb{Z}_{n+1}$, where the generator of the $\mathbb{Z}_{n+1}$ acts as multiplication by (ω, ω^n), with $\omega = \exp(2\pi i/(n+1))$. Its fan Σ can be taken to be the one generated by $(1,0)$ and $(1, n+1)$.

We subdivide Σ by inserting the edges spanned by $v_i = (1,i)$, for $i = 1, \ldots, n$. The resulting fan Σ' defines a smooth toric variety $X_{\Sigma'}$. The relations $v_{i-1} + v_{i+1} - 2v_i = 0$ lead as before to $D_i^2 = -2$. The D_i in fact form a chain of $\mathbb{CP}^1$s. This is the well-known resolution of an A_n singularity.

EXAMPLE 7.6.3. *Consider the simplicial fan Σ consisting of the cone spanned by $(-1,-1,1), (1,0,1), (0,1,1)$, as well as its faces. This defines an affine toric variety, which is in fact the cone over the anti-canonical embedding of $\mathbb{CP}^2$. It can be seen directly to be isomorphic to the orbifold $\mathbb{C}^3/\mathbb{Z}_3$, the $\mathbb{Z}_3$ generator acting as multiplication by (ω,ω,ω), with $\omega = \exp(2\pi i/3)$. It can be blown up by inserting the edge generated by $(0,0,1)$ and subdividing Σ to get a new fan Σ', which we recognize as the fan of $\mathcal{O}_{\mathbb{CP}^2}(-3)$ considered in Example 7.2.5.*

We thus see that $\mathcal{O}_{\mathbb{CP}^2}(-3)$ is the blow-up of $\mathbb{C}^3/\mathbb{Z}_3$ at its singular point. The map $X_{\Sigma'} \to X_\Sigma$ is the blow-down map. These fans are depicted in Fig. 7. We already saw this example from a different point of view in Example 7.3.1.

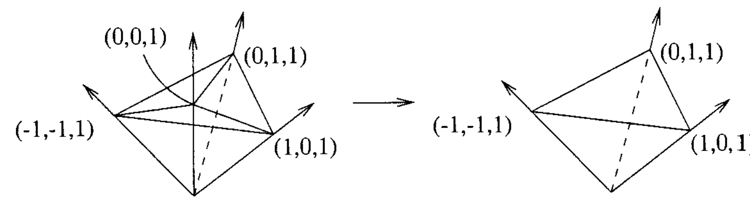


FIGURE 7. The fans for $\mathcal{O}_{\mathbb{CP}^2}(-3)$ and $\mathbb{C}^3/\mathbb{Z}_3$

EXAMPLE 7.6.4. *Consider the fan Σ consisting of the cone σ generated by $(1,0,0), (0,1,0), (0,0,1)$, and $(1,1,-1)$, as well as its faces. The toric variety X_Σ is singular, and the singularity is not an orbifold singularity since σ is not simplicial.*

This is the singularity called a node in the mathematics literature and a conifold singularity in the physics literature.

This singularity can be blown up in two distinct ways to yield smooth toric varieties, as depicted in Fig. 8.

There are no new edges added to the fan in either case, hence there is no exceptional divisor. In either case, there is a new two-dimensional cone σ (spanned by $(1,0,0),(0,1,0)$ and by $(0,0,1),(1,1,-1)$ in the respective cases), so there is an exceptional curve Z_σ, which can be seen to be a $\mathbb{CP}^1$ (in fact, any one-dimensional compact toric variety is necessarily $\mathbb{CP}^1$). The birational map between the two blow-ups is called a flop.

It is an essentially combinatorial result that any toric variety can be desingularized.

THEOREM 7.6.5. *There exists a refinement $\widetilde{\Sigma}$ of any fan Σ such that $X_{\widetilde{\Sigma}} \to X_\Sigma$ is a resolution of singularities.*

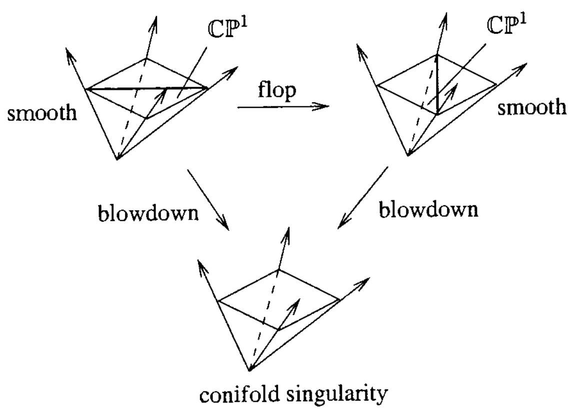


FIGURE 8. The fans for the conifold singularity and its blow-ups

7.7. Morphisms

We have seen several examples of morphisms of toric varieties in previous sections: $F_n \to \mathbb{CP}^1$, $\mathcal{O}(-3) \to \mathbb{CP}^2$, orbifolds, and blow-downs. A systematic understanding is helpful in applications. For instance, we can construct fairly general line bundles or projective bundles. The last construction is very useful for constructing Weierstrass fibrations used to build F-theory compactifications.

DEFINITION 7.7.1. *Let Σ be a fan in $N_{\mathbb{R}}$ and let Σ' be a fan in $N'_{\mathbb{R}}$. A* morphism *from Σ to Σ' consists of a homomorphism $\psi : N \to N'$ such that for each $\sigma \in \Sigma$, the image of σ under $\psi \otimes \mathbb{R}$ is contained in some cone of Σ'.*

The mapping $\psi : N \to N'$ induces a natural mapping of tori

$$T = N \otimes \mathbb{C} \to T' = N' \otimes \mathbb{C}.$$

We leave it to the reader to check that this extends to a mapping $X_\Sigma \to X_{\Sigma'}$.

The global orbifold considered in Sec. 7.5 gives a class of simple examples. In that case, we have N' is a sublattice of N, $\psi : N' \to N$ is the inclusion mapping, $\psi \otimes \mathbb{R}$ is the identity map, and $\Sigma = \Sigma'$.

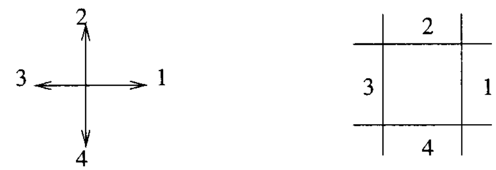


FIGURE 9. The fan of $\mathbb{CP}^1 \times \mathbb{CP}^1$ and T-invariant curves

EXAMPLE 7.7.2. *The fan Σ for $\mathbb{CP}^1 \times \mathbb{CP}^1$, depicted in Fig. 9, has edges spanned by $(1,0),(0,1),(-1,0),(0,-1)$. These in turn correspond to four T-invariant curves, whose configuration is also shown in Fig. 9.*

Note that projection onto either coordinate defines a morphism of fans from Σ to the standard fan for $\mathbb{CP}^1$ (whose edges are the positive and negative rays in $\mathbb{R}$). The corresponding morphisms of toric varieties are just the two projections onto the respective $\mathbb{CP}^1$ factors.

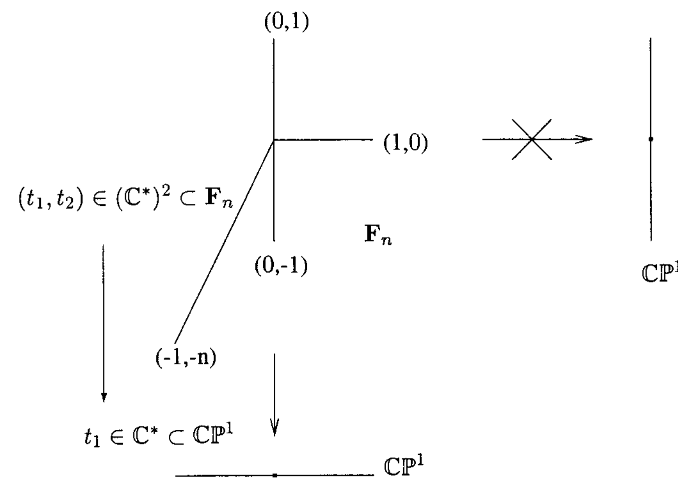


FIGURE 10. The $\mathbb{CP}^1$ bundle structure of F_n

EXAMPLE 7.7.3. *Now let $n > 0$ and consider instead the fan for F_n given in Example 7.2.4. In this case, projection onto the first factor, depicted in Fig. 10, maps this fan to the fan for $\mathbb{CP}^1$ as in the previous example, but projection onto the second factor is not a map of fans, since the image of the cone spanned by $(-1,-n)$ and $(0,1)$ under the second projection is all of $\mathbb{R}$, which is not contained in a cone of the fan for $\mathbb{CP}^1$. This reflects the fact, observed before, that F_n is a non-trivial $\mathbb{CP}^1$ bundle over $\mathbb{CP}^1$.*

We can see that this is a locally trivial $\mathbb{CP}^1$ bundle from the toric geometry using the ideas in this section. We restrict the bundle over the affine open subset $\mathbb{C} \subset \mathbb{CP}^1$ obtained by removing the edge of the fan for $\mathbb{CP}^1$ spanned by (-1). Correspondingly, we must remove the edge spanned by $(-1,-n)$ in the fan for F_n. We obtain the fan in Fig. 11, which is clearly a product $\mathbb{C} \times \mathbb{CP}^1$. We can similarly see the product structure over the other affine piece of $\mathbb{CP}^1$.

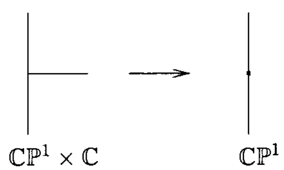


FIGURE 11. Local triviality of F_n as a $\mathbb{CP}^1$ bundle

EXAMPLE 7.7.4. *We return to the fan Σ of $\mathcal{O}_{\mathbb{CP}^2}(-3)$ and now note that it can be constructed directly from the fan Σ' of $\mathbb{CP}^2$. Each three-dimensional cone in Σ is spanned by $(v_1, 1), (v_2, 1), (0,0,1)$, where*

$$\{v_1, v_2\} \subset \{(-1,-1), (1,0), (0,1)\}$$

spans a two-dimensional cone of Σ'.

Projection onto the first two coordinates defines a map from Σ to Σ', which gives rise to the projection $\pi : \mathcal{O}_{\mathbb{CP}^2}(-3) \to \mathbb{CP}^2$. Reasoning as in Example 7.7.3, we see that this is a locally trivial line bundle. The global structure of this bundle (i.e., that it is $\mathcal{O}(-3)$) can be deduced directly from the toric data. The general rule is that if new edges are formed by lifting the

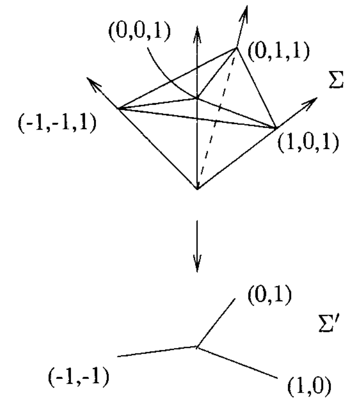


FIGURE 12. The fan description of $\mathcal{O}(-3)$ and its projection to $\mathbb{CP}^2$

edges spanned by $v_i \in N$ to $(v_i, k_i) \in N \oplus \mathbb{Z}$ (and adding the edge $(0,\ldots,0,1)$ over the origin), then the resulting bundle is $\mathcal{O}(-\sum k_i D_i)$, as can be checked.

Note that the edge spanned by $(0,0,1)$ projects to the 0 cone. Since $Z_{\{0\}} = \mathbb{CP}^2$, we conclude that $D_{(0,0,1)}$ maps surjectively to $\mathbb{CP}^2$. As a check, we have already seen that this divisor is the zero section. The other edges project to edges, so the other T-invariant divisors map to divisors in $\mathbb{CP}^1$:

$$\begin{aligned} D_{(-1,-1,1)} &= \pi^{-1}(\{x_1 = 0\}) \\ D_{(1,0,1)} &= \pi^{-1}(\{x_2 = 0\}) \\ D_{(0,1,1)} &= \pi^{-1}(\{x_3 = 0\}) \\ D_{(0,0,1)} &= \text{the zero section } \mathbb{CP}^2. \end{aligned}$$

EXAMPLE 7.7.5. *We can modify the discussion about line bundles to construct projective bundles or even weighted projective bundles, generalizing Example 7.7.3. Here is an example that arose in string theory. Consider* *the matrix P giving the edges of the fan Σ.*

$$\begin{pmatrix} -1 & -2 & -2 & -3 \\ 1 & 0 & -2 & -3 \\ 0 & 1 & -2 & -3 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -3 \end{pmatrix}.$$

We leave it as an exercise to the reader to determine what the correct cones are. Projection onto the first two coordinates maps the fan to the fan for F_2. The fibers are the weighted projective spaces $\mathbb{CP}(1,2,3)$ (note that $\mathbb{CP}(1,2,3)$ is the toric variety associated to the fan with edges generated by $(-2,-3),(1,0),(0,1)$).

This toric variety contains Calabi–Yau hypersurfaces whose fibers over F_2 are elliptic curves in $\mathbb{CP}(1,2,3)$. This is a typical way to construct elliptic fibrations for F-theory compactifications.

7.8. Geometric Engineering

The idea of *geometric engineering* is to construct geometric models with desired properties so that the resulting string theory, M-theory, or F-theory compactification has the desired physics. Toric geometry provides a useful way to engineer these geometries, as it is easy to do direct physical computations in that case. Geometric engineering will be revisited briefly in a broader geometric context in Sec. 36.1.

For example, one way to produce an $\mathcal{N}=2$ $SU(2)$ gauge theory in four dimensions is to produce a Calabi–Yau threefold X containing a surface F_n, which can be blown down to the base $\mathbb{CP}^1$. We consider type IIA string theory compactified on X. There are two massive states corresponding to D2-branes wrapping the fibers of F_n (with either orientation). In the limit where the fiber shrinks and the base $\mathbb{CP}^1$ gets large in such a way as to decouple gravity, we get a field theory in which these massive states become massless and join up with an existing $U(1)$ associated to the volume of the fiber to form an $SU(2)$ vector.

A local model for this geometry is the canonical bundle of F_n, with F_n embedded as the zero section. This can be constructed by toric geometry.

To do this, recall Example 7.7.4, where we showed how to construct the fan of $\mathcal{O}(K_{\mathbb{CP}^2}) = \mathcal{O}_{\mathbb{CP}^2}(-3)$ from the fan of $\mathbb{CP}^2$ and how to generalize this to more general bundles over toric varieties. In particular, let us use this method to construct the canonical bundle over F_2.

The result can be described in terms of the matrices $P|Q$ for F_2 and K_{F_2}:

$$\mathbf{F}_2 = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ -1 & -2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & -2 \end{array}\right)$$

$$K_{F_2} = \left(\begin{array}{ccc|cc} 1 & 0 & 1 & 0 & 1 \\ -1 & -2 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 & -2 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right)$$

Once we have the toric data, we can directly derive field theory results using mirror symmetry. The prepotential can be understood from an appropriate Picard–Fuchs system of equations, which can be deduced from the matrix Q of charges.

We can see more about this local geometry directly from the toric data. Projection onto the first coordinate defines a map $K_{F_2} \to \mathbb{CP}^1$ for which the divisors D_1 and D_2 are fibers. We can find the fan for either of these divisors using the description in Sec. 7.2.1. Using D_1 to illustrate, we need to project onto $\mathbb{Z}^3/\mathbb{Z}\cdot(1,0,1)$. We use the isomorphism

$$\mathbb{Z}^3/\mathbb{Z}\cdot(1,0,1) \simeq \mathbb{Z}^2, \qquad [(a,b,c)] \mapsto (c-a, b+c-a)$$

to identify D_1 as the toric variety associated with a two-dimensional fan with edges generated by $(1,0)$, $(1,1)$, and $(1,2)$. This is the resolution of an A_1 singularity from Example 7.6.2. So the local Calabi–Yau threefold looks like a resolution of an A_1 singularity fibered over a $\mathbb{CP}^1$. This geometry can be generalized to A_n singularities and their resolutions fibered over $\mathbb{CP}^1$.

The Picard–Fuchs equations for the mirror correspond to Picard–Fuchs equations for quantum cohomology. These equations can be proven to hold directly in many situations.

7.9. Polytopes

We now switch gears and discuss projective toric varieties and their relationship with polytopes. Our polytopes will be in $M_{\mathbb{R}}$, the dual space of $N_{\mathbb{R}}$.

DEFINITION 7.9.1. *An* integral polytope *in $M_{\mathbb{R}}$ is the convex hull of a finite set of points in M.*

In the sequel, we will drop the adjective "integral" and refer to these simply as polytopes.

The r-dimensional polytopes are the data needed to describe projective toric varieties.

7.9.1. Toric Varieties from Polytopes.

Consider an r-dimensional polytope $\Delta \subset M_{\mathbb{R}}$. We choose an ordering $m_0, \ldots, m_k$ of $\Delta \cap M$. Since $M = \mathrm{Hom}(T, \mathbb{C}^*)$, we interpret the m_i as nowhere vanishing holomorphic functions on T. These functions give rise to a map

$$f : T \to \mathbb{CP}^k, \qquad f(t) = (m_0(t), \ldots, m_k(t)). \tag{7.15}$$

It is easy to see that f is an embedding. We define $\mathbb{CP}_\Delta$ to be the closure of $f(T)$ in $\mathbb{CP}^k$. There is an action of T on $\mathbb{CP}^k$: The element $t \in T$ acts on $\mathbb{CP}^k$ as coordinatewise multiplication by $(m_0(t), \ldots, m_k(t))$. This gives $\mathbb{CP}_\Delta$ the structure of a toric variety. Note that this abstract variety structure does not depend on the chosen ordering of $\Delta \cap M$.

We can rewrite Eq. (7.15) as $y_i = m_i(t)$, where $(y_0, \ldots, y_k)$ are homogeneous coordinates on $\mathbb{CP}^k$. Now suppose that we have an additive relation $\sum a_i m_i = 0$ in M with $\sum a_i = 0$. Then $\mathbb{CP}_\Delta \subset \mathbb{CP}^k$ satisfies the homogeneous polynomial equation

$$\prod_{a_i > 0} y_i^{a_i} = \prod_{a_i < 0} y_i^{-a_i}. \tag{7.16}$$

It is frequently easy to use Eq. (7.16) to define $\mathbb{CP}_\Delta \subset \mathbb{CP}^k$ directly.

A simple modification of this construction gives non-normal toric varieties: Instead of using all of $\Delta \cap M$ to define Eq. (7.15), use a subset whose convex hull is still Δ. For example, if $\Delta = [0, k]$, then $\mathbb{CP}_\Delta$ is the rational normal curve of degree k in $\mathbb{CP}^k$, but if we use a proper subset $\{0, a_1, \ldots, a_l, k\}$ of $\Delta \cap M$, the closure of the image of $t \mapsto (1, t^{a_1}, \ldots, t^{a_l}, t^k)$ defines a non-normal curve of degree k in $\mathbb{CP}^{l+1}$. For instance, Example 7.2.6 arises using the subset $\{0, 2, 3\}$ of $[0, 3]$ to define the embedding Eq. (7.15).

If we stick to normal varieties, then we can construct the fan directly from the polytope Δ. First, for each face F of Δ, define the cone

$$\sigma_F = \{v \in N_{\mathbb{R}} \mid \langle m, v\rangle \le \langle m', v\rangle \text{ for all } m \in F \text{ and } m' \in \Delta\}.$$

Then the set of all of these cones, as F varies over all faces of Δ, forms a fan, the *normal fan* Σ_Δ.

THEOREM 7.9.2. $X_{\Sigma_\Delta} \simeq \mathbb{P}_\Delta$.

We will not prove this here, but will merely observe that the isomorphism is defined by

$$(x_1, \ldots, x_n) \mapsto \left(\prod_{i=1}^n x_i^{\langle m_0, v_i\rangle}, \ldots, \prod_{i=1}^n x_i^{\langle m_k, v_i\rangle}\right).$$

As usual, $(x_1, \ldots, x_n)$ are homogeneous coordinates in X_{Σ_Δ} and $\Sigma_\Delta(1) = \{v_1, \ldots, v_n\}$.

EXAMPLE 7.9.3. *Let $\Delta \subset \mathbb{R}^2$ be triangle with vertices $\{(0,0), (1,0), (0,1)\}$. Then Eq. (7.15) becomes $f(t_1, t_2) = (1, t_1, t_2)$. The image is dense, and $\mathbb{P}_\Delta \simeq \mathbb{CP}^2$.*

The normal fan can be computed from the definition to be precisely the fan for $\mathbb{CP}^2$ given in Example 7.2.3. Note that the edges of this fan are the inward-pointing normals to corresponding faces of Δ. This is how the normal fan gets its name.

In the next section, we will "derive" this directly from the geometry of $\mathbb{CP}^2$.

EXAMPLE 7.9.4. *Let $\Delta \subset \mathbb{R}^2$ be triangle with vertices $\{(0,0), (2,0), (0,2)\}$. Note that the normal fan is unchanged from Example 7.9.3, since the shape of the polytope is unchanged. The toric variety is still $\mathbb{CP}^2$ but the embedding has changed. There are six points of*

$$\Delta \cap M = \{(0,0), (1,0), (2,0), (0,1), (1,1), (0,2)\}.$$

The torus is therefore embedded in $\mathbb{CP}^5$ as

$$(t_1, t_2) \mapsto (1, t_1, t_1^2, t_2, t_1 t_2, t_2^2). \tag{7.17}$$

This extends to $\mathbb{CP}^2$ as the well-known Veronese embedding of $\mathbb{CP}^2$:

$$(x_1, x_2, x_3) \mapsto (x_1^2, x_1 x_2, x_2^2, x_1 x_3, x_2 x_3, x_3^2).$$

EXAMPLE 7.9.5. *Let $\Delta \subset \mathbb{R}^2$ be the quadrilateral with vertices*

$$\{(0,0),(2,0),(1,1),(0,1)\}.$$

This also contains the lattice point $(1,0)$. Note that we have obtained this from the polytope of Example 7.9.4 by cutting off the corner $(0,2)$.

The normal fan changes to the fan of F_1 (Example 7.2.4), which is the blow-up of $\mathbb{CP}^2$ at a point (Example 7.4.2). So $\mathbb{CP}_\Delta \simeq F_1$.

The embedding in Eq. (7.17) is modified to

$$(t_1,t_2) \mapsto (1,t_1,t_1^2,t_2,t_1t_2).$$

If we tried to extend this to a map from $\mathbb{CP}^2$, we would get

$$(x_1,x_2,x_3) \mapsto (x_1^2,x_1x_2,x_2^2,x_1x_3,x_2x_3). \tag{7.18}$$

But this is not defined at $(0,0,1)$! So we must blow up $(0,0,1)$ to make Eq. (7.18) well-defined, and it is not difficult to see that it is an embedding. So we see directly that $\mathbb{P}_\Delta \simeq F_1$ as well.

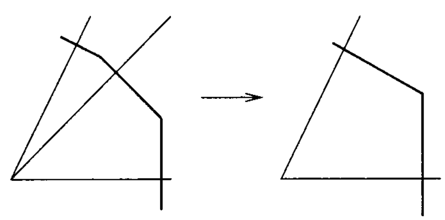


FIGURE 13. Blowing up the toric variety associated to a polytope

More generally, blowing up corresponds to cutting out an edge of Δ. Fig. 13 gives an illustration of part of a polytope and the normal fan, before and after blowing up. It demonstrates how cutting off an edge corresponds to subdividing a fan.

We close this section with an easy way to picture the topology of $\mathbb{CP}_\Delta$ directly from Δ. We content ourselves with examples here.

For the first example, consider the map $\mu : \mathbb{CP}^2 \to \mathbb{R}^2$ given by

$$(x_1,x_2,x_3) \mapsto \left(\frac{|x_2|^2}{|x_1|^2+|x_2|^2+|x_3|^2}, \frac{|x_3|^2}{|x_1|^2+|x_2|^2+|x_3|^2}\right).$$

The image of μ is the set of all $(a,b) \in \mathbb{R}^2$ satisfying $a \geq 0, b \geq 0, a+b \leq 1$, forming the triangle $\Delta \subset \mathbb{R}^2$ from Example 7.9.3. The fiber of a point in the

interior of the triangle is a compact torus $S^1 \times S^1$, the fiber over an interior point of an edge is an S^1, and the fiber over a vertex is a point.

An even simpler example is $\mathbb{CP}^1$ with bundle $\mathcal{O}_{\mathbb{CP}^1}(1)$, in which case Δ is the interval $[0,1]$. We have the map $\mu : \mathbb{CP}^1 \to \mathbb{R}$ given by

$$(x_1,x_2) \mapsto \frac{|x_2|^2}{|x_1|^2+|x_2|^2}.$$

The image of μ is $[0,1]$. The fiber of μ over an interior point of $[0,1]$ is S^1, and the fiber of μ over each of the endpoints is a point. This description leads immediately to the homeomorphism $\mathbb{P}_\Delta \simeq S^2$. Both of these examples are pictured in Fig. 14. The example of $\mathbb{CP}^1$ is actually embedded in the example of $\mathbb{CP}^2$ as the line $x_3 = 0$. Note that the bottom edge of the triangle in the left half of the figure can be identified with $[0,1]$, compatibly with the lattice.

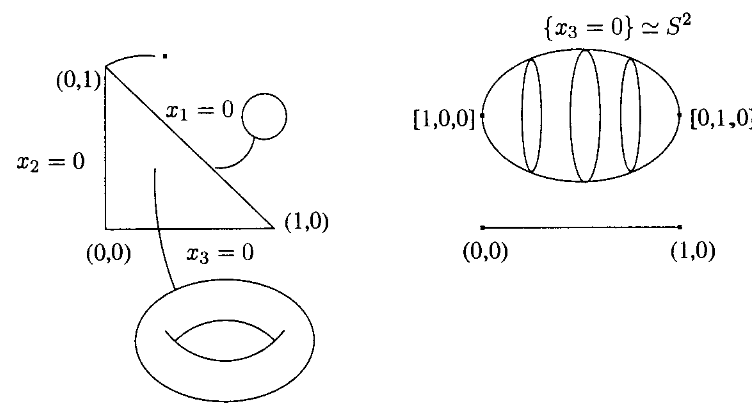


FIGURE 14. The topology of $\mathbb{CP}^2$ and a coordinate line as described by its polytope

In general, there is a continuous map $\mu : \mathbb{CP}_\Delta \to \Delta$ such that the fiber of μ over an interior point of a k-dimensional face of Δ is homeomorphic to $(S^1)^k$.

7.9.2. Polytopes from Toric Varieties. In this section, we construct polytopes from projective toric varieties. The idea is simple. Suppose we have a toric variety $T \subset X$ embedded in a projective space $\mathbb{CP}^k$. This defines a hyperplane class $\mathcal{O}_X(1)$ on X. We will need to assume that the action of T extends to an action on $\mathbb{CP}^k$, acting by coordinatewise multiplications.

To construct a polytope Δ, we need to choose an isomorphism between $\mathcal{O}_X(1)$ and $\mathcal{O}(D)$, where D is some fixed T-invariant divisor. Making a different choice for D will result in a translation of Δ, so the choice of D is essentially irrelevant.

In the usual way, we identify sections of $\mathcal{O}(D)$ with meromorphic functions f on X such that $(f)+D \geq 0$, where (f) is the divisor of f. Thus each coordinate function x_i on $\mathbb{CP}^k$ is identified with a meromorphic function f_i on X. The condition that the T-action extends to $\mathbb{CP}^k$ implies that the restriction of f_i to T is a character of T. We let m_i denote this restriction and identify it with an element of M. Then Δ is the convex hull of the $\{m_i\}$.

EXAMPLE 7.9.3, REVISITED. *We consider $\mathbb{CP}^2$ with hyperplane class identified with $\mathcal{O}(D_1)$. The isomorphism between $\mathcal{O}_X(1)$ and $\mathcal{O}(D_1)$ is defined by division by x_1.*

Thus, the coordinates $\{x_1, x_2, x_3\}$ correspond to the meromorphic functions $\{(1, x_2/x_1, x_3/x_1)\}$ respectively. The coordinates on the torus are given by

$$(t_1, t_2) = (x_2/x_1, x_3/x_1),$$

so the characters are $(0,0), (1,0)$, and $(0,1)$, and we arrive at the polytope Δ that led us to $\mathbb{CP}^2$. We illustrate this polytope in Fig. 15, together with the corresponding monomials on $\mathbb{CP}^2$.

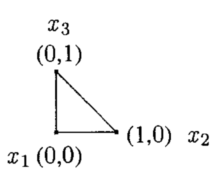


FIGURE 15. The polytope for $\mathbb{CP}^2$ with bundle $\mathcal{O}(1)$

EXAMPLE 7.9.6. *For a toric variety X defined by a fan Σ, we have $\mathcal{O}(-K_X) \simeq \mathcal{O}(\sum_{\rho\in\Sigma(1)} D_\rho)$. In particular, $\mathcal{O}_{\mathbb{CP}^2}(3) \simeq \mathcal{O}(-K_{\mathbb{CP}^2}) \simeq \mathcal{O}(D_1+D_2+D_3)$, where D_i is defined by the section x_i of $\mathcal{O}_{\mathbb{CP}^2}(1)$, $i = 0,1,2$.*

A basis of $\Gamma(\mathcal{O}_{\mathbb{CP}^2}(3))$ is given by the ten homogeneous monomials of degree 3 in x_1, x_2, x_3.

Then with our choice $\mathcal{O}_{\mathbb{CP}^2}(3) \simeq \mathcal{O}(D_1+D_2+D_3)$, the degree 3 polynomial s is identified with the meromorphic function $s/(x_1x_2x_3)$ on $\mathbb{CP}^2$. The T action on the vector space

$$V = \left\{ \frac{s}{x_1x_2x_3} : s \in \Gamma(\mathcal{O}_{\mathbb{CP}^2}(3)) \right\}$$

has weights spanning the polytope $\Delta \subset M_{\mathbb{R}}$ depicted in Fig. 16. The cones over the proper faces of Δ form a fan in $M_{\mathbb{R}}$ also depicted in Fig. 16, which we recognize as the fan of $\mathbb{CP}^2/\mathbb{Z}_3$ in Example 7.9.4. We will use this to illustrate mirror symmetry in Sec. 7.10.

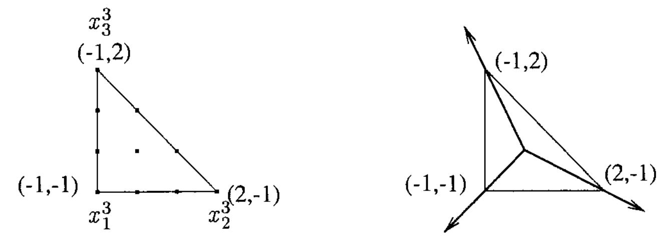


FIGURE 16. The polytope for $\mathbb{CP}^2$ with bundle $\mathcal{O}_{\mathbb{CP}^2}(3)$

7.10. Mirror Symmetry

In this final section, we relate toric geometry to mirror symmetry. First we explain Batyrev's construction of mirror symmetry. Then we relate this to the physical description of mirror symmetry in Ch. 20.

7.10.1. Batyrev's Construction.

Batyrev has introduced a beautiful construction of mirror symmetry for Calabi–Yau hypersurfaces in toric varieties, based on the notion of duality for reflexive polytopes.

DEFINITION 7.10.1. *An integral polytope is* reflexive *if*

(1) *for each codimension 1 face $F \subset \Delta$, there is an $n_F \in N$ with $F = \{m \in \Delta \mid \langle m, n_F\rangle = -1\}$, and*

(2) $0 \in \mathrm{int}(\Delta)$.

The polar polytope Δ° *of* Δ *is the convex hull of the* n_F *in* $N_{\mathbb{R}}$.

THEOREM 7.10.2. *A polytope* Δ *is reflexive if and only if* $\mathbb{CP}_\Delta$ *is Gorenstein and Fano. A polytope* Δ *is reflexive if and only if* Δ° *is reflexive.*

PROOF. See [**16**]. □

The Gorenstein condition on a variety is a condition on its singularities. This means that even though there is no notion of top degree holomorphic forms at the singularities, the canonical bundle extends to a bundle at the singularities. Once there is a canonical bundle, then the Fano condition means as usual that the anti-canonical bundle is positive.

Batyrev's construction can be described as follows. Start with a reflexive polytope Δ. Then its normal fan Σ_Δ coincides with the fan formed by taking the cones over the faces of Δ°. Anti-canonical hypersurfaces are given by sections of the anti-canonical bundle $\mathcal{O}_{X_{\Sigma_\Delta}}(\sum_{v_i \in \Sigma_\Delta(1)} \mathcal{O}(D_i))$. These define Calabi–Yau hypersurfaces $X \subset \mathbb{CP}_\Delta$.

Theorem 7.10.2 says that Δ° is reflexive, so we can apply the same construction starting with Δ° in place of Δ. The result is a family of Calabi–Yau hypersurfaces $X^\circ \subset \mathbb{CP}_{\Delta^\circ}$.

The assertion is that the family X is mirror to the family X°.

Note that if we use the usual embedding $\mathbb{CP}_\Delta \hookrightarrow \mathbb{CP}^k$ with $k = |\Delta \cap M| - 1$, then anti-canonical hypersurfaces in $\mathbb{CP}_\Delta$ are defined by linear equations in the coordinates of $\mathbb{CP}^k$.

REMARK 7.10.3. *We actually need to blow up* X_{Δ_Σ} *by subdividing as in Theorem 7.6.5. The required result is actually a bit stronger: there is a subdivision for which the blow-up is projective.*

Here is an example.

EXAMPLE 7.10.4. *We consider an example of one-dimensional mirror symmetry. A one-dimensional Calabi–Yau is an elliptic curve. Perhaps the simplest algebraic examples are the plane cubic curves in* $\mathbb{CP}^2$. *Let us find the mirror family.*

First we need the polytope Δ *for* $\mathbb{CP}^2$ *with bundle* $\mathcal{O}(3)$ *described in Example 7.9.6.*

The one-dimensional faces of Δ *are defined by the linear inequalities*

$$-a - b \geq -1,\ a \geq -1,\ b \geq -1$$

respectively. So Δ° *is the convex hull of the points* $(-1,-1)$, $(1,0)$, *and* $(0,1)$ *respectively. The cones over the faces of* Δ° *form the fan for* $\mathbb{CP}^2$, *as it must. The polytopes are shown in Fig. 17.*

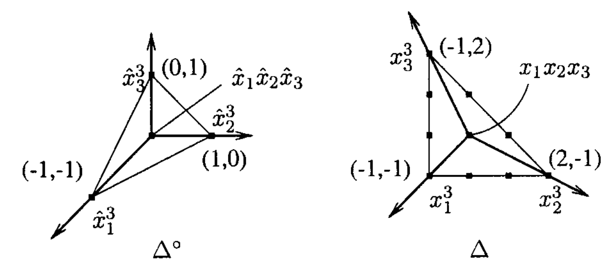


FIGURE 17. The polytopes for $\mathbb{CP}^2$ and its mirror

As in Example 7.9.6, let $f : \mathbb{CP}^2 \to \mathbb{CP}^9, (x_1, x_2, x_3) \mapsto (x_1^3, \ldots, x_3^3)$ *be the 3-fold Veronese embedding, which is also the anti-canonical embedding of* $\mathbb{CP}^2$. *Each section of* $\mathcal{O}(3)$ *defines a cubic curve in* $\mathbb{CP}^2$. *Each monomial corresponds to a character in* M. *Multiplicative relations among sections correspond to additive relations among characters. For example,* $\rho_8 + \rho_{10} = 2\rho_9$ *tells us that the image of* $\mathbb{CP}^2$ *under* f *is contained in the hypersurface* $\{y_8 y_{10} = y_9^2\} \subset \mathbb{CP}^9$, *where* $(y_1, \ldots, y_{10})$ *are homogeneous coordinates on* $\mathbb{CP}^9$. *See Fig. 18. The equations defining the Veronese image can all be found similarly.*

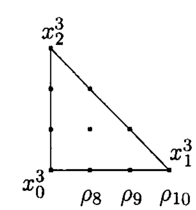


FIGURE 18. The polytope for $\mathbb{CP}^2$ and bundle $\mathcal{O}(3)$ with a dependency

In Example 7.5.3, we saw that cones over the proper faces of Δ *form a fan* Σ° *in* $M_{\mathbb{R}}$ *defining* $\mathbb{CP}^2/\mathbb{Z}_3$. *The anti-canonical class of* $\mathbb{CP}^2/\mathbb{Z}_3$ *consists of* $\mathbb{Z}_3$*-invariant cubics. We can take* $\hat{x}_1^3, \hat{x}_2^3, \hat{x}_3^3, \hat{x}_1\hat{x}_2\hat{x}_3$ *as a basis for the* $\mathbb{Z}_3$*-invariant cubics in* $\mathbb{CP}^2$, *where the* $\hat{x}_i$ *are coordinates in the* $\mathbb{CP}^2$ *with the* $\mathbb{Z}_3$ *action.*

These monomials correspond to lattice points of the polytope Δ° *in* $N_{\mathbb{R}}$, *and cones over proper faces of* Δ° *form a fan* Σ *that defines* $\mathbb{CP}^2$.

The polytope description gives an embedding $\mathbb{CP}^2/\mathbb{Z}_3 \hookrightarrow \mathbb{CP}^3$ *defined by* $X_i = \hat{x}_i^3$, $i = 1,2,3$ *and* $X_0^3 = \hat{x}_1\hat{x}_2\hat{x}_3$, *where the* X_i *are coordinates on* $\mathbb{CP}^2/\mathbb{Z}_3$. *This equation can be deduced from the relation*

$$(1,0) + (0,1) + (-1,-1) = 3 \cdot (0,0).$$

EXAMPLE 7.10.5. *We now consider the famous example of quintic hypersurfaces in* $\mathbb{CP}^4$. *The construction of the mirror family by Greene and Plesser consists of invariant quintic hypersurfaces in* $\mathbb{CP}^4/\mathbb{Z}_5^3$, *where* $\mathbb{Z}_5^3$ *is the group of all automorphisms of the form*

$$(\alpha_1, \ldots, \alpha_5) \text{ with } \alpha_i^5 = 1,\ \prod \alpha_i = 1. \tag{7.19}$$

This can be seen by Batyrev's construction. One way to see this is to start with the fan for $\mathbb{CP}^4$ *with edges given by*

$$\begin{pmatrix} -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.20}$$

Quintic hypersurfaces in $\mathbb{CP}^4$ *are anti-canonical, so the construction of Batyrev applies.*

The polytope Δ° *for the mirror is the convex hull of the rows of Eq. (7.20). Note that* $\Delta^\circ \cap N$ *consists of six points, namely the points represented by the rows of Eq. (7.20) together with the origin, which we denote by* v_0. *These give an embedding* $\mathbb{CP}_{\Delta^\circ} \subset \mathbb{CP}^5$. *We let the coordinates on* $\mathbb{CP}^5$ *be* $(y_0, \ldots, y_5)$, *with* y_0 *corresponding to the origin in* Δ°. *From the relation* $\sum v_i = 5v_0$, *we deduce the equation*

$$y_1 \cdots y_5 = y_0^5 \tag{7.21}$$

for $\mathbb{CP}_{\Delta^\circ}$.

The orbifold description of Greene and Plesser follows immediately: The transformation

$$y_i = \hat{x}_i{}^5, \quad i = 1, \ldots, 5, \qquad y_0 = \hat{x}_1 \cdots \hat{x}_5 \tag{7.22}$$

is invariant under the $\mathbb{Z}_5^3$ *automorphism group Eq. (7.19) and defines the isomorphism* $\mathbb{CP}_{\Delta^\circ} \simeq \mathbb{CP}^4/\mathbb{Z}_5^3$. *The anti-canonical hypersurfaces in* $\mathbb{CP}_{\Delta^\circ}$ *are given by linear expressions in the* y_i. *Under the isomorphism induced by Eq. (7.22), these correspond to* $\mathbb{Z}_5^3$*-invariant quintics, as claimed.*

Alternatively, the identification $\mathbb{CP}_{\Delta^\circ} \simeq \mathbb{CP}^4/\mathbb{Z}_5^3$ *could have been deduced by identifying the normal fan of* Δ° *with the fan consisting of the cones over the proper faces of the polytope* Δ *corresponding to the sections of* $\mathcal{O}(5)$ *on* $\mathbb{CP}^4$, *then applying the methods of Sec. 7.5.*

7.10.2. Relation to the Physical Description of Mirror Symmetry. This final section is not self-contained, as it refers to material to be presented in Ch. 20. We include it here while the ideas of toric geometry are fresh in the reader's mind. We describe part of the relation between toric geometry and the field theoretic description which will be given in Ch. 20.

We introduce n twisted chiral fields $Y_1, \ldots, Y_n$; r twisted chiral fields $\Sigma_1, \ldots, \Sigma_r$; and parameters $t_1, \ldots, t_s$ (mirror to the Kähler parameters of X_Σ). The gauge group is $U(1)^s$. As usual, r is the dimension of X_Σ, and s is the number of independent charges; if Σ has n edges, then $s = n - r$. The charge matrix will again be denoted as $Q = Q_{i,a}$, where $1 \le i \le n$ and $1 \le a \le s$.

Then the required superpotential is

$$W = \left(\sum_{a=1}^{s} \Sigma_a \left(\sum_{i=1}^{n} Q_{i,a} Y_i - t_a\right)\right) + \sum_{i=1}^{n} e^{-Y_i}. \tag{7.23}$$

EXAMPLE 7.10.6. *We return to the quintic. The quintic is related to the non-compact theory of* $\mathbb{CP}^4$ *with bundle* $\mathcal{O}(-5)$. *This can be described by a* $U(1)$ *gauge theory with charges* $(1,1,1,1,1,-5)$. *Labeling the charged twisted chiral fields as* $Y_1, \ldots, Y_5, Y_P$, *the superpotential Eq. (7.23) becomes*

$$W = \Sigma\left(Y_1 + \cdots + Y_5 - 5Y_P - t\right) + \sum_{i=1}^{5} e^{-Y_i} + e^{-Y_P}.$$

The Σ*-constraint gives* $Y_1 + \cdots + Y_5 = 5Y_P + t$. *Exponentiating gives*

$$\prod_{i=1}^{5} e^{Y_i} = q e^{5Y_P}, \tag{7.24}$$

where $q = e^t$. For $q = 1$, this is precisely the same as the equation of the toric variety $\mathbb{CP}_{\Delta^\circ}$ given in Eq. (7.21) after the change of variables $y_i = e^{Y_i}$ and $y_0 = e^{Y_P}$. The case of general q requires a rescaling.

Part 2

Physics Preliminaries

CHAPTER 8

What Is a QFT?

One of the central developments of the past century in theoretical physics was the development of a subject called quantum field theory. This subject is still being developed by physicists. This was at first motivated by an attempt to understand quantum electrodynamics. However it is now believed that all of physics should be based on some quantum field theory. This is mainly because all the known forces and matter in nature can be described by some quantum field theory.

This is also precisely the main obstacle in rigorously connecting modern physics with mathematics. Many of the constructions in quantum field theories, though based on sound physical arguments, are mathematically conjectural and very few quantum field theories have rigorously been proven to exist.

The aim of Part 2 is to develop QFT in as much detail as is essential in understanding mirror symmetry. However, mathematical rigor will not be our main focus, for the reason mentioned above. Instead, we will aim at familiarizing the reader as to how to think about QFT. So our aim is not to define what a quantum field theory is, but to introduce it through a number of examples. We start with easy examples and build toward more difficult and interesting ones. In a sense this section can be viewed as a "practical guide" to quantum field theories.

8.1. Choice of a Manifold M

The starting point for defining a quantum field theory is the choice of a manifold M of dimension d. For most, but not all, QFTs the manifold is viewed as a Riemannian manifold with a smooth metric on it. If the metric is positive definite we sometimes refer to it as a Euclidean QFT. For many physical applications we will also consider manifolds with $d-1$ positive directions and one negative direction of the metric, as in the d-dimensional Minkowski space. The manifold M may or may not have boundaries. In case it does have boundaries some additional information is needed at the boundaries to define the quantum field theory.

8.2. Choice of Objects on M and the Action S

The next ingredient is the choice of objects to consider over M. Roughly speaking, in the QFT one aims to integrate over the space parametrizing these objects. The objects are also called fields. The operation of integration over the fields is also called the *path-integral.* For example, we may consider a principal bundle over M with a connection. In physics terminology the choice of the connection is called "picking a gauge field." We may also be considering sections of a vector bundle over M. These fields are sometimes called matter fields. Quantum field theories associated with connections and sections of associated vector bundles are called "quantum gauge theories."

As another example of QFTs we may consider the space of maps

$$X: \quad M \to N \tag{8.1}$$

for some target manifold N. The field theories associated with integrating over the space of such maps are called sigma models. Sometimes we may be interested in considering various choices of metrics on M. Integrating over such choices is called "quantum gravity."

In integrating over the field space we have to choose a measure on it. In most cases there is a natural choice of a measure on these spaces. The measure is also usually weighted (in the case of Euclidean signature) by $\exp(-S)$, where S is a functional on the space of fields in question and is called the action. In the Minkowski signature the measure is modified by the weight $\exp(iS)$.

8.3. Operator Formalism and Manifolds with Boundaries

One can also consider the case where M has some boundary components:

$$\partial M = \cup_i B_i.$$

This can only occur when the dimension of M is greater than or equal to 1. In such a case, in defining the integration over the field space we have to specify boundary conditions for fields on B_i. The space of field configurations on each B_i gives rise to a Hilbert space $\mathcal{H}_i$, and the path-integral, as we shall see, can be viewed as a multilinear map

$$\otimes_i \mathcal{H}_i \to \mathbb{C}.$$

If we glue two manifolds along their boundaries, the path-integral can be performed by pairing the states corresponding to the boundaries that were glued. This is compatible with the definition of the path-integral as corresponding to the sum over all field configurations (i.e., we fix the field configuration on the boundary we are gluing and then sum over all possible field configurations on the glued boundary).

In the case $M = N \times I$, where N is a manifold without boundaries and I is an interval of length T, the path-integral gives rise to a linear map (by dualizing the Hilbert space corresponding to one of the boundaries):

$$U(T): \mathcal{H} \to \mathcal{H}.$$

Using the sewing property of QFTs we learn that $U(T_1)U(T_2) = U(T_1 + T_2)$. This in turn defines an operator H as the generator of U,

$$U(T) = \exp(-TH)$$

in the Euclidean case, or

$$U(T) = \exp(-iTH)$$

in the case where I corresponds to the negative direction in the signature (the "time"). H is called the Hamiltonian and in most theories is a Hermitian operator.

8.4. Importance of Dimensionality

As is clear from these examples, in quantum field theories we are typically interested in integrating over infinite-dimensional spaces. It turns out that the greater the dimension d of M, the more complicated the integrations over these spaces. In fact (ignoring gravitational theories), the only non-trivial quantum field theories that are believed to exist (i.e., for which some kind of integration over the infinite-dimensional space exists) have $d \leq 6$ and most of the standard ones have $d \leq 4$.

Quantum field theories in different dimensions can be related to each other by an operation known as "Kaluza–Klein reduction." Roughly speaking this means considering the situation where

$$M = N \times K$$

and where K is much less than N. The action S may be very large for field configurations that are not constant over K, so the path-integral, which is weighted by e^{-S}, localizes to field configurations that are constant along K. This gives rise to an "effective" path-integral over field configurations that have only constant modes along K.

Certain path-integrals do not depend on the metric on the manifold. In such cases taking the volume of K to be small reduces the path-integral to a simpler one on N, which is a lower-dimensional manifold (and can possibly be 0-dimensional) and is easier to compute.

Luckily for us, the study of mirror symmetry entails studying quantum field theories with $d = 2$, so our aim is to study mainly low-dimensional quantum field theories. We start with quantum field theories with dimension $d = 0$ and work our way up gradually to $d = 2$.

One nice feature of this way of proceeding is that in cases of $d = 0, 1$ we can make many things (if not everything) mathematically rigorous. Moreover, many of the ideas relevant for the more complicated case of $d = 2$ already show up in these cases.

The case of $d = 0$, corresponding to when M is a point, is already very interesting. In this case QFT is equivalent to carrying out some finite-dimensional integrals, which of course can be rigorously studied. We use this simple case to set up the basic ingredients of quantum field theories and also introduce fermionic fields and supersymmetry, which are quite important in the study of mirror symmetry. Already in this context we can discuss rigorously the important notions of localization and deformation invariance that often arise in supersymmetric quantum field theories.

The case of $d = 1$ is also known as quantum mechanics, as the quantum aspects of particles are captured by it (where M corresponds to the world-line of the particle). In this case we introduce the notion of supersymmetric sigma models as well as supersymmetric Landau–Ginzburg models (sigma models with extra potential functions on the target manifold). For $d = 1$ we can introduce the notion of the operator formulation of quantum theories. The operator formulation on manifolds M arises when it has some boundaries (which occurs only for $d \geq 1$). This is related to the fact that such quantum field theories need extra data at the boundary to make sense of them.

We then move on to the case of $d = 2$ QFTs. We start with some relatively simple examples, involving essentially free theories (sigma models with target manifolds being flat tori). These are already complicated enough to provide us with the basic example of mirror symmetry known as T-duality. We then move on to more complicated cases involving sigma models on Kähler manifolds, their reformulation in terms of gauge theories, and their connection to Landau–Ginzburg theories. The notion of superspace is introduced and used effectively. It turns out that properties of superspace play a crucial role in the formulation and physical proof of mirror symmetry, and we devote a large portion of this part of the book to developing these ideas.

CHAPTER 9

QFT in $d = 0$

In this section we will consider zero-dimensional quantum field theories, i.e., when M is a point. The simplest case is taking the field X to correspond to maps from $M \to \mathbb{R}$, which in this case can be identified with a variable X. The action $S[X]$ in this case is just a function of the variable X. The partition function is an integral given by

$$Z := \int dX \; e^{-S[X]}. \tag{9.1}$$

The correlation functions in this zero-dimensional QFT are just weighted integrals given by

$$\langle f(X) \rangle := \int dX \; f(X)\, e^{-S[X]}. \tag{9.2}$$

Sometimes it is useful to consider normalized correlation functions given by

$$\frac{\int dX\, f(X)\, e^{-S[X]}}{\int dX\; e^{-S[X]}}. \tag{9.3}$$

Another way of determining the correlation functions is to deform the action

$$S \mapsto S' = S + \sum_i a_i f_i(X). \tag{9.4}$$

Then the correlation functions are given by the derivatives of the partition function with respect to the parameter a_i,

$$\langle f_i(X) \rangle = \left. \frac{\partial Z(\alpha, a_i)}{\partial a_i} \right|_{a_i = 0}, \tag{9.5}$$

where α is a parameter of S and

$$Z(\alpha, a_i) = \int dX e^{-S'(X)}. \tag{9.6}$$

As an example, consider the toy model with action

$$S[X] = \frac{\alpha}{2} X^2 + i\epsilon X^3. \tag{9.7}$$

We typically want the action to have certain reality properties but here we will not worry about that. The partition function here depends on two parameters; we write $Z(\alpha, \epsilon)$. Notice that for $\epsilon = 0$ the action is just quadratic and we can write down the exact partition function,

$$Z(\alpha, 0) = \sqrt{\frac{2\pi}{\alpha}}. \tag{9.8}$$

We often define the normalization of the measure of integration such that we get rid of the factor $\sqrt{\frac{2\pi}{\alpha}}$, i.e., we consider the normalized partition function $\frac{Z(\alpha,\epsilon)}{Z(\alpha,0)}$.

If $\epsilon \ll 1$, then we can expand the partition function in powers of ϵ to obtain a perturbative expansion,

$$Z = \int dX\; e^{-\frac{\alpha}{2}X^2 - i\epsilon X^3} = \int dX \sum_{n=0}^{\infty} e^{-\frac{\alpha}{2}X^2} \frac{(-i\epsilon X^3)^n}{n!}. \tag{9.9}$$

We assume that the perturbative expansion exists and do not worry about issues of convergence.

Now we will introduce the machinery of Feynman diagrams, which are very useful methods for perturbative computations in QFTs. Even though the introduction of this machinery is not necessary in this rather simple example, setting it up in a simple situation will help in understanding Feynman diagrams in the more complicated case of higher-dimensional quantum field theories. Consider the function

$$f(\alpha, J) = \int e^{-\frac{\alpha}{2}X^2 + JX}. \tag{9.10}$$

J is known as "the source" in physics. We can perform the integration by completing the square,

$$f(\alpha, J) = \int e^{-\frac{\alpha}{2}(X - \frac{J}{\alpha})^2 + \frac{J^2}{2\alpha}} = \sqrt{\frac{2\pi}{\alpha}}\; e^{\frac{J^2}{2\alpha}}. \tag{9.11}$$

Using the function $f(\alpha, J)$ we can write down some other useful integrals as the derivatives of this function. In particular, we have

$$\int X^r e^{-\frac{\alpha}{2}X^2}\, dX = \left.\frac{\partial^r f(\alpha, J)}{\partial J^r}\right|_{J=0}. \tag{9.12}$$

Pairs of $\frac{\partial}{\partial J}$ act together for a non-vanishing contribution to the above quantity. This can be seen from the form $f \propto \exp(J^2/2\alpha)$. First $\frac{\partial}{\partial J}$ brings down a term $\frac{J}{\alpha}$ from the exponent, then another $\frac{\partial}{\partial J}$ absorbs it. That there must be a second one to absorb it can be seen from the fact that if it were

not absorbed, setting $J = 0$ at the end would yield zero. Since each $\frac{\partial}{\partial J}$ corresponds to an X, we see that in computing the integral of X^r with the Gaussian measure, we have to consider all ways of choosing pairs of them. This operation when used for computing such integrals is called "choosing pairs" and "contracting them." This contraction is also called Wick contraction.

Each pair of $\frac{\partial}{\partial J}$ gives a factor of $\frac{1}{\alpha}$ and therefore $\frac{d^r f(\alpha,J)}{dJ^r}$ gives

$$(\frac{1}{\alpha})^{r/2} \times (\#\text{ of ways of contracting}). \tag{9.13}$$

Sometimes we draw lines to show possible contractions. Such a line is called a propagator. Therefore each propagator is weighted with a factor of $\frac{1}{\alpha}$.

Let us go back to computing the partition function Z for our toy model, Eq. (9.7). Consider the first non-trivial correction to $Z(\alpha, 0)$,

$$\mathcal{O}(\epsilon)^2 \;:\; \frac{(-i\epsilon)^2}{2!} \int dX\; X^3 \times X^3 \times e^{-\frac{\alpha}{2}X^2}. \tag{9.14}$$

The graphical representation of this integral is shown in Fig. 1.

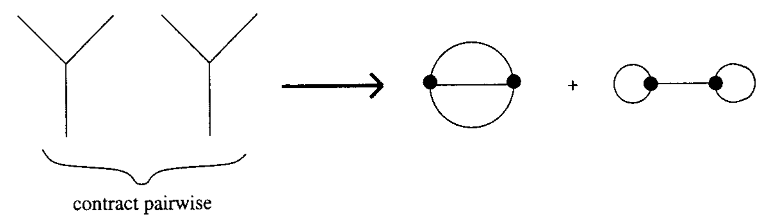


FIGURE 1. There is one vertex for each X^3 and the three edges emanating from the vertex are in one-to-one correspondence with the three X's

The vertices of the graph come from terms in the action with higher powers of X. In general, a term of the form X^k leads to a vertex with k edges emanating from it. The above example involves the case $k = 3$. The first graph gives a factor of $\frac{1}{2}(-i\epsilon)^2(\frac{1}{\alpha})^3 \times 3!$, and the second graph gives a factor of $\frac{1}{2}(-i\epsilon)^2(\frac{1}{\alpha})^3 \times 3^2$. The numbers $3!$ and 3^2 reflect the number of ways the contraction can be done to yield the same diagram. Note that altogether we have $3! + 3^2 = 15$ possible pairs of contractions, this is as expected because the total number of X's is six and choosing a pair of them can be done in $6 \times 5/2 = 15$ ways. The total value of the integral in Eq.

(9.14) is the sum of factors from the two diagrams. These diagrams are called Feynman diagrams. In general we obtain both connected and disconnected diagrams.

EXERCISE 9.0.1. *Show that*

$$Z(\alpha, \epsilon) = e^{\sum \text{connected graphs}}. \tag{9.15}$$

Moreover, show that the combinatorial factor associated to each connected graph is given by $(-3!i\epsilon)^V \alpha^{-E}/|Aut(G)|$, *where* V *is the number of vertices of the graph,* E *is the number of edges, and* $|Aut(G)|$ *denotes the order of the automorphism group of the graph.*

$F := -\ln Z$ is usually called the free energy and is given by minus the sum of the connected graphs.

9.1. Multivariable Case

Consider the case of multiple variables X_i with $(i = 1, \ldots, N)$ and the action given by

$$S(X_i, M, C) = \tfrac{1}{2} X^i M_{ij} X^j + C_{ijk} X^i X^j X^k. \tag{9.16}$$

We assume that the matrix M is positive definite and invertible. Since for $C = 0$ the action is quadratic, we can evaluate the partition function to obtain

$$Z(M, C = 0) = \int \prod_i dX^i \, e^{-\frac{1}{2} X^i M_{ij} X^j} = \frac{(2\pi)^{N/2}}{\sqrt{\det(M)}}. \tag{9.17}$$

The term $C_{ijk} X^i X^j X^k$ in the action leads to a vertex as shown in Fig. 2, with three lines meeting at a point and a factor of $-C_{ijk}$.

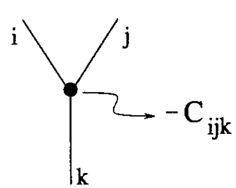


FIGURE 2. The Feynman diagrams with more fields will have edges labeled by the fields. To each vertex we associate a factor $-C_{ijk}$

To determine the partition function for small C we have to expand the exponential of the cubic (and higher-order terms) in powers of C (and other higher-order couplings) and use Feynman rules to determine the coefficients in the perturbative expansion as shown before in the case of a single variable. In this case a propagator connecting X^i and X^j carries a factor of $(M^{-1})_{ij}$.

9.2. Fermions and Supersymmetry

We are interested mainly in supersymmetric quantum field theories. These theories, apart from having ordinary (also called "bosonic") variables such as X^i, also have Grassmann variables ψ^a, which are called "fermionic" or "odd" fields. These form an associative and up to sign, commutative algebra. There is a $\mathbb{Z}_2$ gradation that assigns to all the bosonic variables a $+1$ and to all the fermionic variables a -1, and is compatible with the multiplication in the algebra. The fermionic variables have commutation properties given by

$$X^i \psi^a = \psi^a X^i, \quad \psi^a \psi^b = -\psi^b \psi^a. \tag{9.18}$$

The second property in the above equation implies that $(\psi^a)^2 = 0$. Note that pairs of ψ^i behave like bosonic variables since

$$\psi^a(\psi^b\psi^c) = (\psi^b\psi^c)\psi^a. \tag{9.19}$$

The rules of integration over Grassmann variables are different from bosonic variables and are defined by

$$\int d\psi = 0, \quad \int \psi \, d\psi = 1. \tag{9.20}$$

In the case of many Grassmann variables we have

$$\int \psi^1 \cdots \psi^n \, d\psi^1 \cdots d\psi^n = 1. \tag{9.21}$$

The integrals involving permutations of the n fields are given by ± 1 depending on the parity of the permutation. Any other integral over the fermionic fields (i.e., with less than n fermionic fields) is zero. The action $S(X^i, \psi^a)$ is Grassmann even, which means that we need to have an even number of ψ^a's in each term. In order to evaluate the partition function

$$Z = \int \prod_i dX^i \prod_a d\psi^a \; e^{-S(X,\psi)}, \tag{9.22}$$

we have to expand it in powers of ψ^a and keep only the terms having each ψ^a exactly once. As an example, consider the case when the action only has fermionic variables,

$$S(\psi) = \frac{1}{2}\psi^i M_{ij}\psi^j. \tag{9.23}$$

The partition function in this case is given by

$$Z = \int \prod_k d\psi^k e^{-\frac{1}{2}\psi^i M_{ij}\psi^j} = \mathrm{Pf}(M). \tag{9.24}$$

$\mathrm{Pf}(M)$ is the Pfaffian of M and is such that $\mathrm{Pf}(M)^2 = \det(M)$.

The smallest number of fermionic variables that can have a non-trivial action is two (as the action has to have an even number of them). Consider the most general action of one bosonic variable and two fermionic variables given by

$$S(X, \psi^1, \psi^2) = S_0(X) - \psi^1\psi^2 S_1(X). \tag{9.25}$$

The partition function is given by

$$\begin{aligned} Z &= \int dX d\psi^1 d\psi^2 e^{-S_0 + \psi^1\psi^2 S_1(X)} \\ &= \int dX d\psi^1 d\psi^2 e^{-S_0}(1 + \psi^1\psi^2 S_1(X)) \\ &= \int dX d\psi^1 d\psi^2 e^{-S_0} + \int dX d\psi^1 d\psi^2 e^{-S_0}\psi^1\psi^2 S_1(X). \end{aligned} \tag{9.26}$$

The first term vanishes due to ψ^i integration, and we get

$$Z = \int dX e^{-S_0} S_1(X). \tag{9.27}$$

We thus see that we can integrate out the odd variables and end up with an integral purely in terms of bosonic variables.

For a special choice of $S_0(X)$ and $S_1(X)$ the above theory has a symmetry, known as supersymmetry. Let

$$S_0(X) = \frac{1}{2}(\partial h)^2 \quad \text{and} \quad S_1(X) = \partial^2 h, \tag{9.28}$$

where h is a real function of X and $\partial h := h'$. In other words, consider the zero-dimensional QFT defined by the action

$$S(X, \psi_1, \psi_2) := \tfrac{1}{2}(\partial h)^2 - \partial^2 h\, \psi_1\psi_2. \tag{9.29}$$

There are symmetries of this action generated by odd parameters, which are symmetries that exchange bosonic fields with fermionic fields and are known as supersymmetries. Consider the following transformation of the fields:

$$\begin{aligned} \delta_\epsilon X &= \epsilon^1\psi_1 + \epsilon^2\psi_2, \\ \delta\psi_1 &= \epsilon^2\partial h, \\ \delta\psi_2 &= -\epsilon^1\partial h. \end{aligned} \tag{9.30}$$

Here ϵ^i and ψ_i are Grassmann odd variables, therefore they anti-commute with each other. They denote the infinitesimal parameters generating the supersymmetry. It is easy to check that the action is invariant under this transformation.

EXERCISE 9.2.1. *Show that the integration measure $dX d\psi_1 d\psi_2$ is also invariant under this transformation. (In showing this you will develop a concept known as superdeterminant and its infinitesimal version, the supertrace, which one encounters when dealing with both even and odd variables).*

9.3. Localization and Supersymmetry

In the context of this very simple supersymmetric quantum field theory we will illustrate an important principle that occurs in supersymmetric theories in general. This phenomenon, known as localization, allows one to compute partition functions (and certain correlation functions) of supersymmetric theories by showing that the relevant path-integrals defining the quantum field theory reduce to a much smaller-dimensional integral, and in ideal situations reduce to counting contributions of certain points in the field space.

Suppose ∂h is nowhere zero. Then we will show that

$$Z := \int e^{-S}\, dX\, d\psi_1\, d\psi_2 = 0\,. \tag{9.31}$$

The basic idea is to trade one of the fermionic fields with the supersymmetry transformation variable. Put differently, we choose the supersymmetry transformation to set one of the fermions in the action to be zero, and then use the rules of Grassmann integration to get zero. For example, if we consider $\epsilon^1 = \epsilon^2 = -\psi_1/\partial h$, which is allowed if $\partial h \neq 0$, then the ψ_1 field will be eliminated from the action. This motivates us to consider the change of variables

$$\begin{aligned} \widehat{X} &:= X - \frac{\psi_1\psi_2}{\partial h(X)}, \\ \widehat{\psi}_1 &:= \alpha(X)\psi_1, \\ \widehat{\psi}_2 &:= \psi_1 + \psi_2, \end{aligned} \tag{9.32}$$

where α is an arbitrary function of X. Since the action is invariant under the supersymmetry transformation $(X, \psi_1, \psi_2) \to (\widehat{X}, 0, \widehat{\psi}_2)$, we have

$$S(X, \psi_1, \psi_2) = S(\widehat{X}, 0, \widehat{\psi}_2). \tag{9.33}$$

The integration measure is written in the new variables as

$$dX\, d\psi_1\, d\psi_2 = \left(\alpha(\widehat{X}) - \frac{\partial^2 h(\widehat{X})}{(\partial h(\widehat{X}))^2}\widehat{\psi}_1\widehat{\psi}_2\right) d\widehat{X}\, d\widehat{\psi}_1\, d\widehat{\psi}_2\,. \tag{9.34}$$

Thus the partition function is given by

$$\begin{aligned} Z \;=\; & \int d\widehat{\psi}_1 \int e^{-S(\widehat{X},0,\widehat{\psi}_2)}\alpha(\widehat{X}) d\widehat{X}\, d\widehat{\psi}_2 \\ & - \int e^{-S(\widehat{X},0,\widehat{\psi}_2)} \frac{\partial^2 h(\widehat{X})}{(\partial h(\widehat{X}))^2}\widehat{\psi}_1\widehat{\psi}_2 d\widehat{X}\, d\widehat{\psi}_1\, d\widehat{\psi}_2. \end{aligned} \tag{9.35}$$

The first term vanishes since $\widehat{\psi}_1$ does not appear in the integrand and the integral over $\widehat{\psi}_1$ gives zero by the following rule of Grassmann integration:

$$\int d\widehat{\psi}_1\, 1 = 0. \tag{9.36}$$

The second term survives the Grassmann integration, but it also vanishes since it is a total derivative in $\widehat{X}$.

Now let us consider a more general situation where ∂h may be zero for some X's. In this case the change of variable above is singular at such X's. Let us integrate over the fermionic fields and the X, with an infinitesimal neighborhood of points where $\partial h = 0$ is deleted. Then the above argument still applies and for this part of the contribution we get zero. On the other hand, if $\partial h = 0$ then $\delta\psi_i = 0$. That is, in the vicinity of the points where $\partial h = 0$, we cannot trade the supersymmetry transformation variable with one of the fermionic fields, i.e., the points where $\partial h = 0$ are the fixed points of odd symmetry shown in Eq. (9.30). Thus we see that the computation of the partition function localizes to the vicinity of the fixed point set. This is the localization principle: *The path-integral is localized at loci where the R.H.S. of the fermionic transformation under supersymmetry is zero.* This principle holds for any QFT with supersymmetry. We will now use this result to compute the above partition function in a simple way.

We know from the localization principle that the partition function gets contributions only from the critical points of h. Let us consider the case in which h is a generic polynomial of order n with isolated critical points. Then it has at most $n-1$ critical points.

Near the critical point X_c, h can be written as

$$h(X) = h(X_c) + \frac{\alpha_c}{2}(X - X_c)^2 + \cdots\,. \tag{9.37}$$

Since the partition function localizes at the critical points we can consider the infinitesimal neighborhood of such points and keep only the leading terms in the action suitable for this infinitesimal neighborhood. In other words, we can forget about the higher-order terms. Near each critical point X_c the partition function becomes (including the suitable normalization of the measure discussed before)

$$\sum_{X_c} \int \frac{dX\, d\psi^1 d\psi^2}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha_c^2(X-X_c)^2 + \alpha_c\psi^1\psi^2} = \sum_{X_c} \frac{\alpha_c}{|\alpha_c|} = \sum_{X_c} \frac{h''(X_c)}{|h''(X_c)|}. \tag{9.38}$$

Thus we see that the partition function is an integer given by

$$Z = \sum_{x_0 : \partial h|_{x_0} = 0} \frac{\partial^2 h(x_0)}{|\partial^2 h(x_0)|}. \tag{9.39}$$

This result implies that if n, the order of h, is odd, then $Z = 0$, because there are as many critical points with positive $\partial^2 h$ as with negative, and if the order of h is even, then $Z = \pm 1$, the sign depending on whether the leading term in h is positive or negative (because the number of positive and negative $\partial^2 h$ differ by one).

The fact that the partition function turns out to be an integer is at first surprising. It seems as if it is counting something. This turns out to be explainable when we discuss a related one-dimensional QFT, in which case the same computation arises and is related to counting the dimension of a subspace (the ground states) of a Hilbert space.

From the above result we see not only a localization principle, but also a hint of a deformation invariance of the result. In other words, the partition function seems to be sensitive (up to sign) only to the order of the polynomial in h. We will now explain this deformation invariance, which is another general property shared by supersymmetric quantum field theories.

9.4. Deformation Invariance

If we have a quantum field theory with a symmetry, meaning that the action and the measure are invariant, then the correlation function of quantities that are variations of other fields under the symmetry vanish. In other words, if $f = \delta g$, where δg denotes the variation of g under some symmetry, then

$$\langle f \rangle = \int f e^{-S} = \int \delta g e^{-S} = \int \delta(g e^{-S}) = 0 \tag{9.40}$$

This follows from a change of variables of the integral and is valid as long as the "integration by parts" that could potentially lead to boundary terms is absent. In other words, as long as g is not too big at infinity in field space this should be valid. This general idea applies to both bosonic and fermionic symmetries. Here we wish to apply it to fermionic symmetries.

For the supersymmetric quantum field theory at hand we take $g = \partial\rho(X)\psi_1$ and consider the variation of g under the supersymmetry transformation shown in Eq. (9.30) with $\epsilon^1 = \epsilon^2 = \epsilon$ and $f = \delta_\epsilon g$, which is given by

$$\begin{aligned} f = \delta_\epsilon g &= \partial^2\rho\delta X\psi_1 + \partial\rho(X)\delta\psi_1 \\ &= \epsilon(\partial\rho\partial h - \partial^2\rho\psi_1\psi_2). \end{aligned} \tag{9.41}$$

Thus since $\langle \delta_\epsilon g \rangle = 0$ we see that

$$\langle \partial\rho\partial h - \partial^2\rho\,\psi_1\psi_2 \rangle = 0. \tag{9.42}$$

Since

$$S = \frac{1}{2}(\partial h)^2 - \partial^2 h\,\psi_1\psi_2, \tag{9.43}$$

we see that under the change $h \mapsto h + \rho$ in the action

$$\delta_\rho S = \partial h \partial\rho - \partial^2\rho\psi_1\psi_2. \tag{9.44}$$

Thus it follows from Eq. (9.42) that

$$\langle \delta_\rho S \rangle = 0. \tag{9.45}$$

This implies that the partition function is invariant under the change in the superpotential. This is true as long as ρ is small at infinity in field space compared to h (otherwise the boundary terms in the vanishing argument discussed above will be present). If h is a polynomial of order n, then ρ could be a lower-order polynomial with the vanishing argument still applicable

(ρ can even be of degree n as long as the X^n term is smaller than that in h). In particular transformations of the form $h \mapsto \lambda h$ with $\lambda > 0$ do not change the partition function if the leading term in h is not changed. Thus we see that the partition function is invariant under a large class of deformations of the action.

This idea can also be used to evaluate the partition function. For example, consider rescaling $h \longrightarrow \lambda h$ with $\lambda \gg 1$. In this case the action is very large and $\exp(-S)$ very small, except in the vicinity of the critical points of h. This effectively reduces the problem to the local computations we have already encountered in the context of the localization principle.

In fact without any computations we can also gain insight into the result for the partition function by considering the deformations of h. Since the partition function is invariant under deformation of h, it is easy to see from Fig. 3 that, if h is a polynomial of order n, then we can deform h such that it has no critical points if n is odd and only one critical point if n is even. Using the invariance under the rescaling of h we can now see that if n is odd the partition function vanishes as it has no critical points and if n is even the answer comes from a single point and the answer is ± 1 with the sign determined by the sign of $\partial^2 h$ at the critical point.

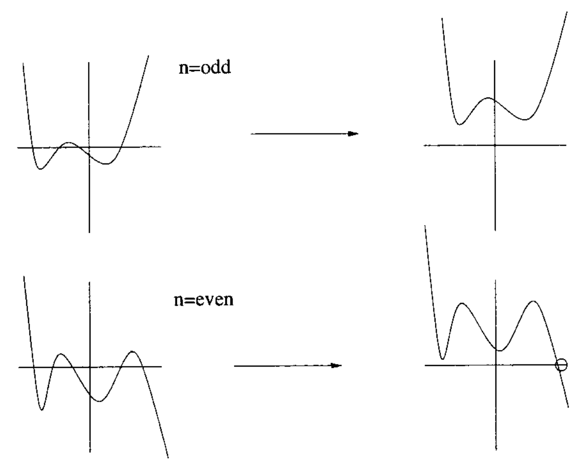


FIGURE 3. Deformation invariance is a powerful tool in computation of partition function in supersymmetric theories

9.5. Explicit Evaluation of the Partition Function

One of the advantages of considering such a simple example is that we can actually do the integral directly and check the results we obtained based on localization and deformation invariance principle. We integrate out the fermionic fields to obtain

$$Z = \frac{1}{\sqrt{2\pi}} \int dX\, d\psi_1\, d\psi_2 e^{-\frac{1}{2}(\partial h)^2 + \partial^2 h \psi_1 \psi_2} = \frac{1}{\sqrt{2\pi}} \int dX \partial^2 h\, e^{-\frac{1}{2}(\partial h)^2}. \tag{9.46}$$

We define a new coordinate $y = \partial h$. Then the above partition function is

$$Z = D \frac{1}{\sqrt{2\pi}} \int dy\, e^{-\frac{1}{2}y^2} = D, \tag{9.47}$$

where D denotes the degree of the map $X \mapsto y = \partial h(X)$. Here D enters the equation because the change of variable from X to $y = \partial h$ is not one-to-one. From the property of the degree of the map (which counts the number of preimages of a given point taking into account the relative orientation of each preimage with respect to its image), we know that D is zero when n is odd and ± 1 when n is even. In other words, we find

$$Z = 0, \text{if} \quad n = \text{odd} \quad \text{and} \quad \pm 1 \quad \text{if} \quad n = \text{even}. \tag{9.48}$$

This result is in agreement with what we obtained using localization and deformation invariance arguments.

9.6. Zero-Dimensional Landau–Ginzburg Theory

Now we consider the complex analogue of the theory considered before. The variables are doubled: $(X, \psi_1, \psi_2) \mapsto (z, \overline{z}, \psi_1, \psi_2, \overline{\psi}_1, \overline{\psi}_2)$, where z is a complex bosonic variable and ψ_i are complex fermionic variables, with $\overline{\psi}_i$ denoting the complex conjugate variable. The action is given by

$$S(z, \bar{z}, \psi_1, \psi_2, \bar{\psi}_1, \bar{\psi}_2) = |\partial W|^2 - (\partial^2 W)\psi_1\psi_2 - (\overline{\partial^2 W})\overline{\psi}_1\overline{\psi}_2, \tag{9.49}$$

where $W(z)$ is a holomorphic function of z.[1] The action is invariant under the transformations

$$\begin{aligned} \delta z &:= \epsilon^1 \psi_1 + \epsilon^2 \psi_2, \quad \delta \overline{z} := 0, \\ \delta \psi_1 &:= \epsilon^2 \overline{\partial W}, \quad \delta \overline{\psi}_1 := 0, \\ \delta \psi_2 &:= -\epsilon^1 \overline{\partial W}, \quad \delta \overline{\psi}_2 := 0 \end{aligned} \tag{9.50}$$

and

$$\begin{aligned} \overline{\delta} \overline{z} &:= \overline{\epsilon}^1 \overline{\psi}_1 + \overline{\epsilon}^2 \overline{\psi}_2, \quad \overline{\delta} z := 0, \\ \overline{\delta \psi}_1 &:= \overline{\epsilon}^2 \partial W, \quad \overline{\delta} \psi_1 := 0, \\ \overline{\delta \psi}_2 &:= -\overline{\epsilon}^1 \partial W, \quad \overline{\delta} \psi_2 := 0. \end{aligned} \tag{9.51}$$

So now we have four real (or two complex) supersymmetry transformations. Note that if we restrict to the transformations with $\epsilon^1 = \epsilon^2$ and $\overline{\epsilon^1} = \overline{\epsilon^2}$, then the above SUSY transformations are such that $\delta^2 = 0, \overline{\delta}^2 = 0$.

The localization principle discussed before, applied to this case, implies localization near the critical points of W. If the critical points of W are isolated and non-degenerate, then near the critical point z_c

$$W(z) = W(z_c) + \frac{\alpha}{2}(z - z_c)^2 + \dots, \tag{9.52}$$

$$e^{-S} = e^{-|\alpha(z - z_c)|^2 + \alpha\psi_1\psi_2 + \overline{\alpha}\overline{\psi}_1\overline{\psi}_2}, \tag{9.53}$$

$$\begin{aligned} Z &:= \frac{1}{2\pi} \int e^{-S} dz d\overline{z} d\psi_1 d\psi_2 d\overline{\psi}_1 d\overline{\psi}_2 \\ &= \sum_{z_c : \partial W(z_c) = 0} \frac{1}{2\pi} |\alpha|^2 \int e^{-|\alpha(z - z_c)|^2} dz d\overline{z} \\ &= \sum_{z_c : \partial W(z_c) = 0} 1 = \#\text{ of critical points of } W. \end{aligned} \tag{9.54}$$

Thus the partition function of this theory counts the number of critical points of the holomorphic function $W(z)$.

In general the computation of correlation functions in supersymmetric theories (other than the function 1, which is the partition function) is not easy. However, if we have enough supersymmetry, we can compute correlation functions of certain fields that are invariant under some of the supersymmetries. The fact that we have so many supersymmetries in this example suggests that we should be able to compute some correlation functions in

[1] We sometimes write $f(z)$ for a holomorphic function of z, and $f(z, \bar{z})$ for a non-holomorphic function.

this theory. In fact, as we will now see, there is an interesting relation between supersymmetry and holomorphicity for this QFT. If we consider the correlation function $\langle f \rangle$, where $f = z^i \overline{z}^l$ with nonzero i and l, this would in general lead to a rather complicated integral that is not possible to evaluate using any localization principle. This is in accord with the fact that this f is not invariant under any of the supersymmetries. However, we can restrict to either functions of z or functions of $\overline{z}$. These f's do preserve half the supersymmetry since $\delta \overline{z} = 0$ and $\overline{\delta} z = 0$. Thus correlation functions of holomorphic or anti-holomorphic quantities can be calculated using the localization principle. In particular, for holomorphic f we apply the localization principle to the $\overline{\delta}$ supersymmetry variation. This implies that again the correlation function localizes to the points where $\partial W = 0$:

$$\begin{aligned} \langle f(z) \rangle &= \int \frac{dz d\overline{z} d\psi^1 d\psi^2 d\overline{\psi}^1 d\overline{\psi}^2}{2\pi} f(z) e^{-S} \\ &= \int \frac{dz d\overline{z}}{2\pi} f(z) |\partial^2 W|^2 e^{-\frac{1}{2}|\partial W|^2}. \end{aligned} \tag{9.55}$$

Due to localization we only need to determine the partition function near the critical points of W,

$$\begin{aligned} \langle f(z) \rangle &= \sum_{z_c : \partial W(z_c) = 0} f(z_c) \int \frac{dz d\overline{z}}{2\pi} |\partial^2 W|^2 e^{-\frac{1}{2}|\partial W|^2} \\ &= \sum_{z_c : \partial W(z_c) = 0} f(z_c). \end{aligned} \tag{9.56}$$

Similarly, if $g(\overline{z})$ is an anti-holomorphic function, by considering the δ supersymmetry variation we have

$$\langle g(\overline{z}) \rangle = \sum_{\overline{z}_c : \overline{\partial W}(\overline{z}_c) = 0} g(\overline{z}_c). \tag{9.57}$$

9.6.1. Chiral Ring. We saw above that we can calculate the correlation functions of fields that are invariant under the $\overline{\delta}$ transformations. Such fields are called *chiral fields*. Note that the product of two chiral fields is again a chiral field, because

$$\overline{\delta}(fg) = (\overline{\delta} f) g + f (\overline{\delta} g). \tag{9.58}$$

Among fields made up only of bosonic fields the chiral fields are holomorphic functions of z. We can also construct fields that are trivially chiral. Consider fields of the form given by $h = \overline{\delta} \Lambda$. Since $\overline{\delta}^2 = 0$ (recall we are taking $\overline{\epsilon}^1 = \overline{\epsilon}^2$)

it follows that $\overline{\delta} h = 0$. It is natural to consider the $\overline{\delta}$ cohomology, i.e., the equivalence classes of chiral fields modulo the addition of trivially chiral fields. As usual the cohomology elements can be viewed as

$$\{\overline{\delta} \Phi = 0\} / \{\Phi = \overline{\delta} \Lambda\}. \tag{9.59}$$

The study of this cohomology is also very natural to consider from the viewpoint of the QFT, because the addition of trivially chiral fields to the chiral fields does not affect the correlation functions:

$$\langle f + \overline{\delta} \Lambda \rangle = \langle f \rangle. \tag{9.60}$$

This follows from the $\overline{\delta}$ symmetry of the action. The QFT gives a natural evaluation on the cohomology elements (analogous to the integration of top forms on manifolds in the context of de Rham cohomology).

We can also study the corresponding cohomology ring. We consider the product of chiral fields and consider only the cohomology class of the product (as usual, one can check that the product does not depend on the choice of the representatives). In the present context this cohomology ring is called the *chiral ring*.

We will now evaluate the chiral ring for bosonic fields. Note that if $f(z)$ is a holomorphic function of z then

$$\overline{\delta}_{\overline{\epsilon}^1 = \overline{\epsilon}^2} (f(z) \overline{\psi}_1) = f(z) \partial W(z). \tag{9.61}$$

This implies that the bosonic chiral fields (which are holomorphic functions of z) are trivially chiral if they have a $\partial W(z)$ as a factor. In other words, we find that the chiral ring is given by

$$\mathcal{R} = \mathbb{C}[z] / \{\mathcal{I}\}, \tag{9.62}$$

where $\mathcal{I}$ is the ideal generated by the ∂W. As an example, consider

$$W(z) = \frac{1}{n+1} z^{n+1} - \lambda z, \tag{9.63}$$

(where λ is a constant). Since $\partial W = z^n - \lambda$, this implies that the chiral ring is generated by one element z with the relation $z^n = \lambda$. Thus the ring elements are given by $\mathcal{R} = \{1, z, z^2, \ldots, z^{n-1}\}$. Moreover, since the correlation functions make sense as evaluations on the cohomology elements, we learn that the correlation functions of z^{i+kn} and $z^i \lambda^k$ are equal. This can also be checked directly from the computation of the correlation function for chiral fields, as shown in Eq. (9.56). In fact in this case one easily sees that

$\langle z^r \rangle$ is zero for all r except when $r \equiv 0 \bmod n$, in which case the correlation function is

$$\langle z^{kn} \rangle = n\lambda^k. \tag{9.64}$$

9.6.2. Multivariable Case. The supersymmetric quantum field theories we have studied can of course be naturally extended to many variables, both in the real case as well as in the Landau–Ginzburg case. Here we will write the LG case explicitly and leave the other case as an exercise for the reader.

For multi-variable LG theory we have variables $(z_i, \psi_1^i, \psi_2^i)$ and their complex conjugates, where $i = 1, \cdots N$. The action is a simple generalization of the action considered before and is given by

$$S(z_i, \psi_1^i, \psi_2^i) = \sum_{i=1}^{N} |\partial_i W(z_1, \ldots, z_N)|^2 - \partial_i \partial_j W \, \psi_1^i \psi_2^j - \overline{\partial_i \partial_j W} \, \overline{\psi}_1^i \overline{\psi}_2^j. \tag{9.65}$$

Localization implies that the partition function and correlation functions of holomorphic functions (or anti-holomorphic functions) localize at the critical points of W, $\partial_i W = 0 \, \forall i$. The chiral ring in this case is given by

$$\mathcal{R} = \mathbb{C}[z_1, \ldots, z_N]/\{\mathcal{I}\}, \tag{9.66}$$

where $\mathcal{I}$ is the ideal generated by $\partial_i W$.

An interesting set of examples we will encounter later involves LG theories with a quasi-homogeneous superpotential W. These are W's that are polynomials in the z_i with the property that

$$W(\lambda^{q_1} z_1, \ldots, \lambda^{q_N} z_N) = \lambda W(z_1, \ldots, z_N) \tag{9.67}$$

for some weights q_i. We can think about this as introducing a gradation on the fields, where z_i has grade q_i and the products of fields are compatible with the addition of the gradation. In physics terminology one calls this a $U(1)$ charge. In this case the chiral ring $\mathcal{R}$ will also inherit the gradation.

We will mainly encounter cases where W corresponds to an isolated singularity. This means that if we consider $\partial_i W = 0$ for all i, the only solution is at the origin, $z_i = 0$.

EXERCISE 9.6.1. *Show that for an isolated quasi-homogeneous singularity the* Poincaré polynomial *of the chiral ring (also known as the singularity*

ring) defined by $P(t) = \sum_{X_\alpha \in \mathcal{R}} t^{Q_\alpha}$, *where* Q_α *is the gradation of the chiral field* X_α, *is given by*

$$P(t) = \prod_i \frac{(1 - t^{1-q_i})}{(1 - t^{q_i})}, \tag{9.68}$$

Show that this implies that the dimension of $\mathcal{R}$ *is*

$$\dim \mathcal{R} = \prod_i \frac{(1 - q_i)}{q_i} \tag{9.69}$$

and that for every element of charge Q_α *there is an element of charge* $D - Q_\alpha$ *(the analogue of Poincaré duality for LG theories), where*

$$D = \sum_i (1 - 2q_i). \tag{9.70}$$

This is why we sometimes say that the corresponding LG theory has dimension D *given by the above formula.*

CHAPTER 10

QFT in Dimension 1: Quantum Mechanics

In this chapter we consider one-dimensional quantum field theories, also known as quantum mechanics. We give a brief introduction to quantum mechanics and discuss certain aspects of it in the context of supersymmetric quantum mechanics.

We introduce various examples. In particular we consider supersymmetric quantum mechanical systems corresponding to maps from one-dimensional space to target spaces that are Riemannian manifolds (we also specialize to the case of Kähler manifolds). These are known as sigma models. We discuss the operator formalism of supersymmetric quantum mechanics and relate the Hilbert space in this context with the space of differential forms on the manifold. The supersymmetry operator gets identified with the d operator and the Hamiltonian with the Laplacian acting on differential forms on the manifold. Above all, the supersymmetric ground states will be the main focus of the discussion. These turn out to correspond to cohomology elements of the manifold. We also consider introducing a "potential" on the manifold (i.e., a choice of function on the manifold) which deforms the theory, and relate certain aspects of this quantum-mechanical system to Morse theory.

These examples will serve as simple concrete models to appreciate the structure of the supersymmetry algebra. It is also a good preparation for the $(1 + 1)$-dimensional supersymmetric field theories to be discussed in upcoming chapters.

10.1. Quantum Mechanics

We start with a brief introduction to quantum mechanics without supersymmetry. In the path-integral formalism, which generalizes our discussion of zero-dimensional QFT, the partition function and the correlation functions are expressed as integrations over fields defined on a one-dimensional manifold. Also, we will have an alternative formulation — the operator formalism — based on states and operators, which only exists for QFTs with $d \geq 1$. As noted before, this arises when we consider manifolds with boundaries, which in this context corresponds to considering an interval as the manifold.

The one-dimensional space on which we formulate the QFT is either a finite interval I, the real line $\mathbb{R}$ or the circle S^1. It is parametrized by time t. We first consider the case of a single bosonic field X, a map into a target manifold that for the moment we take to be $\mathbb{R}$:

$$X : I, \mathbb{R} \text{ or } S^1 \to \mathbb{R}. \tag{10.1}$$

We consider the action

$$S = \int L\, dt = \int \left\{ \frac{1}{2}\left(\frac{dX}{dt}\right)^2 - V(X) \right\} dt. \tag{10.2}$$

Here L is known as the Lagrangian. This is the action of a particle (of mass 1) moving in the target space $\mathbb{R}$ under the influence of the potential $V(X)$. The equation of motion for the particle can be obtained by looking at configurations $X(t)$, which extremize the above action for a fixed boundary value. That is,

$$\delta S = \int \left\{ \frac{dX}{dt}\delta\left(\frac{dX}{dt}\right) - \frac{dV}{dX}\delta X \right\} dt = 0. \tag{10.3}$$

Using integration by parts, we obtain the equation of motion (the Euler–Lagrange equation),

$$\frac{d^2X}{dt^2} = -\frac{dV}{dX}. \tag{10.4}$$

In the zero-dimensional case considered in the previous chapter we had no time derivatives and the action had only a potential term. The action, as shown by Eq. (10.2), has no explicit time dependence and the system has time translation symmetry. Namely, the action is invariant under $X(t) \to X(t+\alpha)$ for a constant α. If we let α depend on t, $\alpha = \alpha(t)$, the action varies as

$$\delta S = \int dt\, \dot{\alpha}(t) \left(\frac{1}{2}\dot{X}^2 + V(X) \right), \tag{10.5}$$

where the dot over the field denotes d/dt. For a configuration that obeys the equation of motion, Eq. (10.5) must be zero for any $\alpha(t)$. Integration

by parts yields

$$\frac{d}{dt}\left(\frac{1}{2}\dot{X}^2 + V(X) \right) = 0. \tag{10.6}$$

The quantity

$$H = \frac{1}{2}\dot{X}^2 + V(X) \tag{10.7}$$

is a constant of motion. This is the energy of this system, or the Hamiltonian in the canonical formalism. In general, following the same procedure one can find a constant of motion, or a conserved charge, for each symmetry of the action. This is called Noether's procedure and the constant of motion is called the Noether charge.

Let us consider the integral

$$Z(X_2, t_2; X_1, t_1) = \int DX(t)\, e^{iS(X)}, \tag{10.8}$$

where integration is over all paths connecting the points X_1, X_2 such that $X(t_1) = X_1$ and $X(t_2) = X_2$ as shown in Fig. 1. This integral is called a

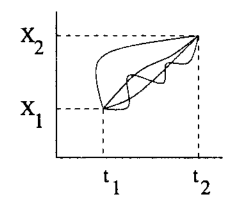


FIGURE 1.

path-integral for the obvious reason. Since $S(X)$ is real we are summing up phases associated with different paths and the convergence of the integral is a subtle problem. One can actually avoid this difficulty by considering the "Euclidean theory" (which will also be useful for other purposes). This is obtained by "Euclideanizing" the time coordinate t by the so-called Wick rotation:[1]

$$t \longrightarrow -i\tau. \tag{10.9}$$

[1]The reason that it is called "Euclidean theory" will become clear when we consider (1+1)- or higher-dimensional quantum field theory.

Then the action becomes $S(X) \longrightarrow iS_E(X)$, where $S_E(X)$ is the Euclidean action

$$S_E(X) = \int \left\{ \frac{1}{2}\left(\frac{dX}{d\tau}\right)^2 + V(X) \right\} d\tau. \tag{10.10}$$

The path-integral is now given by

$$Z_E(X_2, \tau_2; X_1, \tau_1) = \int_{X(\tau_1)=X_1}^{X(\tau_2)=X_2} DX(\tau)\, e^{-S_E(X)}. \tag{10.11}$$

Note that the kinetic term is positive semi-definite and the integral has a better convergence property (as long as the potential $V(X)$ grows at infinity in X). We can also consider the partition function as the Euclidean path-integral on the circle S^1_β of circumference β:

$$Z_E(\beta) = \int_{X(\tau+\beta)=X(\tau)} DX(\tau)\, e^{-S(X)}. \tag{10.12}$$

The most subtle part of the story is to define the measure of integration. One way of defining it is to divide the time coordinate into intervals and use a single variable in each interval. After the integration is done over all the intervals we can take the size of the interval to zero. There are technical issues here about how to make sense of this process. For the one-dimensional path-integrals there are ways of rigorously defining the path-integral using random walk techniques. In a "free field theory," by which we mean the action is quadratic, we can define it as a generalization of Eq. (9.17) where the matrix M is now of infinite size. As we will see, one can define the determinant of such an infinite matrix by so-called zeta function regularization. If the theory is not free but the interaction term is small, one can define the path-integral as the perturbation series in the small coupling constant, as was done in the zero-dimensional example. In particular, just as in the zero-dimensional QFT of Ch. 9, we can formulate a notion of Feynman diagrams, with propagators and vertices etc.

EXERCISE 10.1.1. *Formulate Feynman diagram perturbation theory for quantum mechanics by following steps similar to those for the zero-dimensional QFT.*

Starting from path-integrals, we can move to the operator formalism, which is how quantum mechanics was historically formulated. In general terms, the Hilbert space and operator formulation arises when we consider manifolds with boundaries. To each boundary we associate a Hilbert space that corresponds to fixing the field configurations at the boundary. In the case at hand, i.e., one-dimensional QFT, the boundary is just a point. Fixing the value of the field at the boundary corresponds to choosing delta function distributions in this case. More precisely, the Hilbert space $\mathcal{H}$ in this case is the space of complex-valued square-normalizable functions of the variable X, i.e., $\mathcal{H} = L^2(\mathbb{R}; \mathbb{C})$, with its standard inner product

$$\langle f, g \rangle = \int \overline{f(X)} g(X) dX. \tag{10.13}$$

This Hilbert space is considered to be the space of "states." Let us consider a mapping of a state at time t_1 to a state at time t_2,

$$Z_{t_2;t_1} : \mathcal{H} \longrightarrow \mathcal{H}, \tag{10.14}$$

given by

$$f(X_1) \mapsto (Z_{t_2;t_1} f)(X_2) = \int Z(X_2, t_2; X_1, t_1) f(X_1) dX_1. \tag{10.15}$$

This is the operator representing the time evolution of the states. If the action is invariant under the time translation, as in Eq. (10.2), then

$$Z(X_2, t_2; X_1, t_1) = Z(X_2, t_2 - t_1; X_1, 0) =: Z_{t_2-t_1}(X_2, X_1) \tag{10.16}$$

and $Z_{t_2;t_1} = Z_{t_2-t_1;0} =: Z_{t_2-t_1}$. By definition, we have

$$\int Z_{t_3-t_2}(X_3, X_2) Z_{t_2-t_1}(X_2, X_1) dX_2 = Z_{t_3-t_1}(X_3, X_1), \tag{10.17}$$

which expresses the obvious fact that the time evolution from t_1 to t_2 and then from t_2 to t_3 is the same as the evolution from t_1 to t_3. In short, $Z_t Z_{t'} = Z_{t+t'}$. Thus, the time evolution operator can be written as

$$Z_t = e^{-itH} \tag{10.18}$$

for some operator H. The Noether charge in the classical theory corresponds, in the quantum theory, to the generator of the associated symmetry transformation.[2] The generator H of the time translation is called the "Hamiltonian." It is a Hermitian operator and the time evolution operator $Z_t = e^{-itH}$ is a unitary operator.

It turns out that H can be described in a systematic fashion for quantum-mechanical systems. In particular, in the system with the classical action

[2]It is a good exercise to show this using the path-integral.

described by Eq. (10.2), the Hamiltonian, which is also known as the energy of the system, is given by Eq. (10.7) or

$$H = \frac{1}{2}p^2 + V(X), \tag{10.19}$$

where p is the conjugate momentum of X, $p = \delta S/\delta\dot{X} = \partial L/\partial\dot{X}$, with S and L as in Eq. (10.2). In the classical theory, X and p obey the relation

$$\{X, p\} = 1, \tag{10.20}$$

where $\{\,,\}$ is the Poisson bracket. It turns out that in quantum theory H corresponds to the operator given by the same expression, where the Poisson brackets are replaced by commutators and X and p satisfy the commutation relation

$$[X, p] = i. \tag{10.21}$$

From the above commutator it follows that when acting on the space of functions of X we can identify X with multiplication by X and p with the operator

$$p := -i\frac{d}{dX}. \tag{10.22}$$

Thus X and p become Hermitian operators (we ignore boundedness issues for the moment). In the Euclidean theory, e^{-itH} is replaced with $e^{-\tau H}$, which is not a unitary operator. We will not show why this dictionary between the path-integral and operator formulations of quantum mechanics works as indicated here, but just use it and check in examples how it works.

Now consider the partition function on the circle S^1_β of circumference β. This can be considered to be the Euclidean path-integral on the interval of length β with the values of X at the initial and final end points identified and integrated over. Thus, it is given by

$$Z_E(\beta) = \int dX_1 Z_{E,\beta}(X_1, X_1) = \mathrm{Tr}\, e^{-\beta H}. \tag{10.23}$$

10.1.1. Examples. Simple Harmonic Oscillator. Consider the Lagrangian

$$L = \frac{1}{2}\dot{X}^2 + \frac{1}{2}X^2. \tag{10.24}$$

The Hamiltonian is given by

$$H = \frac{p^2}{2} + \frac{X^2}{2} = \frac{1}{2}(p + iX)(p - iX) + \frac{1}{2}, \tag{10.25}$$

where the last term is due to the fact that $[X, p] = i$. We define new operators

$$a = \frac{1}{\sqrt{2}}(p - iX), \quad a^\dagger = \frac{1}{\sqrt{2}}(p + iX), \tag{10.26}$$

so that the Hamiltonian has a simple expression,

$$H = a^\dagger a + \frac{1}{2}. \tag{10.27}$$

The operators a and $a^\dagger$ obey the commutation relations

$$[a, a^\dagger] = 1, \tag{10.28}$$

$$[a, a] = [a^\dagger, a^\dagger] = 0, \tag{10.29}$$

from which it follows that

$$[H, a] = -a, \quad [H, a^\dagger] = a^\dagger. \tag{10.30}$$

Thus, if $|\psi\rangle$ is a state of energy E, i.e., if it satisfies

$$H|\psi\rangle = E|\psi\rangle, \tag{10.31}$$

then we have

$$Ha|\psi\rangle = (E-1)a|\psi\rangle, \text{ and } Ha^\dagger|\psi\rangle = (E+1)a^\dagger|\psi\rangle. \tag{10.32}$$

Namely, a and $a^\dagger$ lower and raise the energy by one unit and for this reason they are called the lowering and raising operators respectively. The ground state $|0\rangle$ is defined as the state annihilated by the lowering operator,

$$a|0\rangle = 0. \tag{10.33}$$

This state has energy $E_0 = 1/2$. The corresponding wave-function obeys the differential equation $(-i\frac{d}{dX} - iX)\Psi_0(X) = 0$ that corresponds to $(p - iX)|0\rangle = 0$. There is a unique solution (up to an overall constant) given by

$$\Psi_0(X) = e^{-\frac{1}{2}X^2}. \tag{10.34}$$

The Hilbert space is spanned by states $|n\rangle = (a^\dagger)^n|0\rangle$ of energy

$$E_n = n + \frac{1}{2}. \tag{10.35}$$

Since we have determined the spectrum we can evaluate the partition function:

$$Z(\beta) = \mathrm{Tr}\, e^{-\beta H} = \sum_{n=0}^{\infty} e^{-\beta(n+\frac{1}{2})} = \frac{1}{2\sinh(\beta/2)}. \tag{10.36}$$

We can also evaluate the partition function in the path-integral formalism

$$Z(\beta) = \int_{X(t+\beta)=X(t)} DX(t) \exp\left(-\int dt \left(\frac{1}{2}\dot{X}^2 + \frac{1}{2}X^2\right)\right). \tag{10.37}$$

The Euclidean action can be written as

$$\frac{1}{2}\int dt \left(\frac{1}{2}\dot{X}^2 + \frac{1}{2}X^2\right) = \frac{1}{2}\int dt X\Theta X, \tag{10.38}$$

where $\Theta = -\frac{d^2}{dt^2} + 1$. Let $f_n(t)$ be the orthonormal eigenfunctions of the operator Θ,

$$\Theta f_n(t) = \lambda_n f_n(t), \quad \int \bar{f}_n(t) f_m(t) dt = \delta_{n,m}. \tag{10.39}$$

Then we can expand $X(t)$ in terms of the eigenfunctions $f_n(t)$, $X(t) = \sum_n c_n f_n(t)$. We can use c_n as the new variables in the path-integral,

$$\mathrm{e}^{-S} = \mathrm{e}^{-\frac{1}{2}\sum_n \lambda_n c_n^2} \tag{10.40}$$

$$DX(t) = \prod_n \frac{dc_n}{\sqrt{2\pi}}. \tag{10.41}$$

The path-integral then becomes

$$Z(\beta) = \prod_n \lambda_n^{-1/2} = \frac{1}{\sqrt{\det(\Theta)}}. \tag{10.42}$$

The eigenvalues of the operator Θ are

$$\lambda_n = 1 + \left(\frac{2\pi n}{\beta}\right)^2, \tag{10.43}$$

where n runs over all non-negative integers and there is one mode (constant mode) for $n = 0$ and there are two modes ($\cos(2\pi nX/\beta)$ and $\sin(2\pi nX/\beta)$) for $n \geq 1$. Thus, we have

$$Z(\beta) = \prod_{n=1}^{\infty} \left(1 + \left(\frac{2\pi n}{\beta}\right)^2\right)^{-1}. \tag{10.44}$$

We can write the above product as

$$Z(\beta) = \prod_{n=1}^{\infty} \left(\frac{2\pi n}{\beta}\right)^{-2} \prod_{n=1}^{\infty} \left(1 + \left(\frac{2\pi n}{\beta}\right)^{-2}\right)^{-1}. \tag{10.45}$$

The second factor is a convergent product and is given by $\beta/(2\sinh(\beta/2))$. The first factor is divergent and requires a regularization. This is done by the zeta function regularization, as we now show. We consider a function

$$\zeta_1(s) = \sum_{n=1}^{\infty} \left(\frac{2\pi n}{\beta}\right)^{-2s}, \tag{10.46}$$

which is convergent for sufficiently large $\mathrm{Re}(s)$ and can be analytically continued to near $s = 0$. If we take the derivative at $s = 0$, we obtain $\zeta_1'(0) = \sum_{n=1}^{\infty} \log(2\pi n/\beta)^{-2}$ and the infinite product can be identified as $\prod_{n=1}^{\infty}(2\pi n/\beta)^{-2} = \exp \zeta_1'(0)$. We note that the function $\zeta_1(s)$ is related to Riemann's zeta function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ by $\zeta_1(s) = (\beta/2\pi)^{2s}\zeta(2s)$ and therefore $\zeta_1'(0) = 2\log(\beta/2\pi)\zeta(0) + 2\zeta'(0)$. Using the property $\zeta(0) = -1/2$ and $\zeta'(0) = -(1/2)\log(2\pi)$ of Riemann's zeta function, we obtain $\zeta_1'(0) = -\log(\beta/2\pi) - \log(2\pi) = -\log\beta$. Thus, the first factor of Eq. (10.45) is regularized as $\exp\zeta_1'(0) = 1/\beta$ and the partition function is given by

$$Z(\beta) = \frac{1}{\beta} \cdot \frac{\beta}{2\sinh(\beta/2)}. \tag{10.47}$$

This agrees with the result obtained in the operator formalism.

Sigma Model on a Circle. As another example we consider the case when the target space is the circle S^1_R of circumference R and the potential is trivial, $V(X) = 0$. The field X is now a periodic variable

$$X \sim X + R. \tag{10.48}$$

The action is given by

$$S(X) = \int \frac{1}{2}\dot{X}^2 dt, \tag{10.49}$$

and the Hamiltonian is

$$H = \frac{1}{2}p^2 = -\frac{1}{2}\frac{d^2}{dX^2}. \tag{10.50}$$

The eigenfunctions and the eigenvalues of the Hamiltonian are

$$\psi_n = \mathrm{e}^{2\pi i nX/R}, \quad E_n = \frac{2\pi^2 n^2}{R^2}, n \in \mathbb{Z}. \tag{10.51}$$

Using the operator formalism we find the partition function to be

$$Z(\beta) = \mathrm{Tr}\,\mathrm{e}^{-\beta H} = \sum_{n=-\infty}^{\infty} \mathrm{e}^{-\beta 2\pi^2 n^2/R^2}. \tag{10.52}$$

In the path-integral approach we have

$$Z(\beta) = \int DX\, e^{-S_E(X)} = \int DX\, e^{-\int_0^\beta \frac{1}{2}(\frac{dX}{d\tau})^2 d\tau}. \tag{10.53}$$

Here the integration is over all maps of S^1_β to S^1_R. The topological type of the map (i.e., the connected component in the space of all maps) is classified by the winding number, m, which is an integer. Thus, the path-integral is the sum over all possible winding sectors

$$Z(\beta) = \sum_{m=-\infty}^{\infty} \int DX_m\, e^{-S_E(X_m)}, \tag{10.54}$$

where X_m is a variable that represents a map of winding number m, $X_m(\beta) = X(0) + Rm$. It is convenient to express the variable X_m as

$$X_m(\tau) = \frac{m\tau R}{\beta} + X_0(\tau), \tag{10.55}$$

where $X_0(\tau)$ is a periodic function. The action for this X_m is given by

$$S_E(X_m) = \frac{m^2R^2}{2\beta} + \int_0^\beta X_0\left(-\frac{1}{2}\frac{d^2}{d\tau^2}\right)X_0 d\tau. \tag{10.56}$$

Then the path-integral becomes

$$Z(\beta) = \sum_{m=-\infty}^{\infty} e^{-\frac{m^2R^2}{2\beta}} \int DX_0\, e^{-\int_0^\beta X_0\left(-\frac{1}{2}\frac{d^2}{d\tau^2}\right)X_0 d\tau}. \tag{10.57}$$

The integrals over X_0 are common to all m:

$$\frac{R\sqrt{\beta}}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{\det'\left(-\frac{d^2}{d\tau^2}\right)}}. \tag{10.58}$$

The first factor is from integration over the zero mode (constant mode). The factor $1/\sqrt{2\pi}$ comes from the definition of the measure (as in Eq. (10.41)) and the factor $\sqrt{\beta}$ arises because the normalized zero mode is $1/\sqrt{\beta}$ and therefore the integration variable takes values in $[0, R\sqrt{\beta}]$ rather than $[0, R]$. On the other hand, $\det'\left(-\frac{d^2}{d\tau^2}\right)$ in the second factor is the determinant of the operator $-\frac{d^2}{d\tau^2}$ acting on the nonzero modes. For each $n \neq 0$ there is one mode with eigenvalue $(2\pi n/\beta)^2$ for $-\frac{d^2}{d\tau^2}$. Thus, the determinant is

$$\det{}'\left(-\frac{d^2}{d\tau^2}\right) = \prod_{n\neq 0}\left(\frac{2\pi n}{\beta}\right)^2 = \beta^2, \tag{10.59}$$

where the zeta function regularization is assumed (and the computation in the previous example is directly applied). Thus the path-integral gives

$$Z(\beta) = \frac{R}{\sqrt{2\pi\beta}} \sum_{m=-\infty}^{\infty} e^{-\frac{m^2R^2}{2\beta}}. \tag{10.60}$$

This looks different from the result obtained from the operator formalism, Eq. (10.52), but in fact it is exactly equal to that due to an identity known as the Poisson resummation formula.[3]

Sigma Model on the Real Line $\mathbb{R}$. Let us finally consider the theory of single bosonic field X without a potential, $V(X) = 0$. The action is simply

$$S = \int \frac{1}{2}\dot{X}^2\, dt. \tag{10.61}$$

This theory can be considered to be the sigma model on the real line $\mathbb{R}$. The Hamiltonian is given by

$$H = \frac{1}{2}p^2. \tag{10.62}$$

For any k, the plane-wave

$$\Psi_k(X) = e^{ikX} \tag{10.63}$$

is the momentum eigenstate of momentum $p = k$. This is of course the Hamiltonian eigenstate of energy

$$E_k = \frac{1}{2}k^2. \tag{10.64}$$

Unlike in the previous two examples, the wave-functions Ψ_k are not square-normalizable but satisfy the orthogonality relation

$$\int \Psi_k^*(X)\Psi_{k'}dX = 2\pi\delta(k - k'). \tag{10.65}$$

Also, the spectrum is continuous and the partition function $Z(\beta) = \mathrm{Tr}\, e^{-\beta H}$ is not well defined. If we consider this theory to be the $R \to \infty$ of the sigma

[3] The Poisson resummation formula can be obtained as follows. We first note the identity

$$\sum_{n=-\infty}^{\infty} \delta(x + 2\pi n) = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} e^{imx}.$$

Multiplying by $e^{-\frac{\alpha}{2}x^2}$ and integrating over x, this identity yields

$$\sum_{n=-\infty}^{\infty} e^{-\frac{\alpha}{2}(2\pi n)^2} = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty}\int e^{imx - \frac{\alpha}{2}x^2}dx = \frac{1}{\sqrt{2\pi\alpha}}\sum_{m=-\infty}^{\infty} e^{-\frac{1}{2\alpha}m^2}.$$

In the case at hand, $\alpha = \beta/R^2$.

model on S^1_R, then by using Eq. (10.60) the partition function can be written as

$$Z(\beta) = \lim_{R\to\infty} \frac{R}{\sqrt{2\pi\beta}} \sum_{m=-\infty}^{\infty} e^{-\frac{m^2R^2}{2\beta}} = \frac{\lim_{R\to\infty} R}{\sqrt{2\pi\beta}}. \tag{10.66}$$

10.1.2. More General Sigma Models. So far we have considered a rather simple target space, namely the flat space or a circle. We can also consider quantum-mechanical systems with the target being manifolds with non-trivial topology and metric. These more general cases are also known as non-linear sigma models.

Consider the case of a non-linear sigma model with target space a Riemannian manifold with metric $g_{ij}(X)$. The action in this case is

$$S = \tfrac{1}{2}\int dt\; g_{ij}(X)\tfrac{dX^i}{dt}\tfrac{dX^j}{dt}. \tag{10.67}$$

We can expand the metric in Riemann normal coordinates around any point,

$$g_{ij}(X) = \delta_{ij} + C_{ijkl}X^kX^l + \cdots. \tag{10.68}$$

Thus we see that we have a quadratic term in the action as well as quartic and higher-order terms (involving the curvature). This makes explicit computations in the path-integral more difficult. It is possible to obtain the path-integral as a perturbation series, starting from the quadratic term in the fields, but it will be very hard to obtain the exact result in this way. In this case, it turns out that the operator approach is more powerful.

Recall that in the quantum theory X is the position operator and the associated conjugate momentum is

$$P_i := \frac{\delta S}{\delta \dot{X}^i} = g_{ij}\partial_t X^j. \tag{10.69}$$

X and P satisfy the commutation relation

$$[X^i, P_j] = i\delta^i_j. \tag{10.70}$$

To define the Hamiltonian, we start from the classical expression of the energy for this system, which is given by

$$H = \frac{1}{2}g^{ij}(X)P_iP_j.$$

In the quantum theory the above expression for the Hamiltonian is ambiguous, because X and P do not commute. Requiring H to be Hermitian places some constraint but is not strong enough to fix H uniquely. It is clear from the above expression that H is a kind of Laplacian acting on functions over

the manifold. But one has many inequivalent quantum choices for H that reduce to the same classical object.

EXERCISE 10.1.2. *Show why the above Hamiltonian is related to the Laplacian acting on functions on the manifold.*

This ambiguity in the choice of quantization of this system is related to different ways of making sense of the measure in the path-integral. As we will see when we discuss the supersymmetric sigma model, maintaining supersymmetry fixes the ambiguity in operator-ordering for the Hamiltonian.

At any rate, once we fix a choice of Hamiltonian we can compute, for example, the partition function on a circle, which in the operator formulation is given by $\mathrm{Tr}\, e^{-\beta H}$, in terms of the spectrum of the Laplacian on the manifold.

10.1.3. Semi-Classical Approximation. If the action is not quadratic in the fields it is difficult to determine the spectrum exactly and to compute the partition function. In such cases an approximation scheme can be used to express the partition function in terms of an expansion parameter.

Let $S(X)$ be the action and X_{cl} be a solution of the classical equations of motion, i.e.,

$$\left.\frac{\delta S}{\delta X}\right|_{X=X_{cl}} = 0. \tag{10.71}$$

Then we can expand the action around the classical solution,

$$S(X) = S(X_{cl}) + \frac{(\delta X)^2}{2}\left.\frac{\delta^2 S}{\delta X^2}\right|_{X=X_{cl}} + \cdots. \tag{10.72}$$

Keeping only the terms in the action up to quadratic order in δX, we can evaluate the partition function as

$$\begin{aligned} Z &= \int \mathcal{D}X\; e^{\frac{i}{\hbar}S(X)}, \\ &= e^{iS(X_{cl})}\int \mathcal{D}\delta X\; e^{i\frac{(\delta X)^2}{2}\frac{\delta^2 S(X_{cl})}{\delta X^2}+\cdots} \\ &\approx e^{iS(X_{cl})}\frac{1}{\sqrt{\det(\frac{\delta^2 S(X_{cl})}{\delta X^2})}}. \end{aligned} \tag{10.73}$$

A good approximation to the path-integral is to take the above Z summed over *all* the classical solutions to the system, and include the determinant of the operator obtained by integrating over the quadratic terms near each classical solution. This is called the semi-classical approximation. In general this is only an approximation valid when the fields do not vary too much from the classical configurations. As we will see later in the context of supersymmetric theories, however, for certain computations the semi-classical computation is *exact*. In fact, we have already seen examples of this in the context of the zero-dimensional supersymmetric QFTs, where we saw that the sum of the contributions of the path-integral near the critical points of a superpotential, which are analogues of the classical solutions in this context, give the exact result. The analogue of the determinants in that context gave us the $+/-$ sign contributions.

10.2. The Structure of Supersymmetric Quantum Mechanics

We now embark on the study of quantum mechanics with supersymmetry, or supersymmetric quantum mechanics. In quantum mechanics, in general, it is very hard to find exact information such as the spectrum of the Hamiltonian and the correlation functions. This is also true for supersymmetric quantum mechanics. However, a particular class of data can be obtained exactly in supersymmetric theories, the most important of which are the supersymmetric ground states. This will be the focus of the present section. Also, one can exactly evaluate correlation functions of operators that preserve a part of the supersymmetry. We will see that these data can be obtained by employing the localization principle and deformation invariance, as discussed before in the context of zero-dimensional supersymmetric QFTs.

10.2.1. Single-Variable Potential Theory. We start our study with a specific example. The example is the supersymmetric generalization of our potential theory with a single variable x. The theory has a superpartner of x that is a complex fermion ψ. The Lagrangian is given by

$$L = \frac{1}{2}\dot{x}^2 - \frac{1}{2}\left(h'(x)\right)^2 + \frac{i}{2}\left(\overline{\psi}\dot{\psi} - \dot{\overline{\psi}}\psi\right) - h''(x)\overline{\psi}\psi, \tag{10.74}$$

where $\overline{\psi}$ is the complex conjugate of ψ, $\overline{\psi}=\psi^\dagger$. The second term, $-\frac{1}{2}(h'(x))^2$, is the potential term $-V(x)$. Needless to say, ψ and $\overline{\psi}$ are anti-commuting variables. The Lagrangian is real, as one can check by using the property $(\overline{\psi}\psi)^\dagger = \psi^\dagger\overline{\psi}^\dagger = \overline{\psi}\psi$.

Let us consider a transformation of the fields

$$\begin{aligned} \delta x &= \epsilon\overline{\psi} - \overline{\epsilon}\psi, \\ \delta\psi &= \epsilon(i\dot{x} + h'(x)), \\ \delta\overline{\psi} &= \overline{\epsilon}(-i\dot{x} + h'(x)), \end{aligned} \tag{10.75}$$

where $\epsilon = \epsilon_1 + i\epsilon_2$ is a complex fermionic parameter and $\overline{\epsilon}$ is its complex conjugate, $\overline{\epsilon} = \epsilon^*$. Under this variation of fields, the Lagrangian changes by a total derivative in time $\delta L = \frac{d}{dt}(\cdots)$ and therefore the action is invariant:

$$\delta S = \int \delta L\, dt = 0, \tag{10.76}$$

as long as the boundary variation vanishes. Thus, the system has a symmetry associated with the transformation shown in Eq. (10.75). Since the variation parameter is fermionic, such a symmetry is called a fermionic symmetry. We can also see that (up to the equations of motion)

$$[\delta_1, \delta_2]x = 2i(\epsilon_1\overline{\epsilon}_2 - \epsilon_2\overline{\epsilon}_1)\dot{x}, \quad [\delta_1, \delta_2]\psi = 2i(\epsilon_1\overline{\epsilon}_2 - \epsilon_2\overline{\epsilon}_1)\dot{\psi}, \tag{10.77}$$

where δ_i is the fermionic transformation Eq. (10.75) with the variation parameter $\epsilon = \epsilon_i$ $(i = 1, 2)$. Roughly speaking, the square of the fermionic transformation is proportional to the time derivative. Such a fermionic transformation is called a supersymmetry. We refer to this situation by saying that *the classical system with the Lagrangian shown in Eq. (10.74) has supersymmetry generated by Eq. (10.75).* This QFT is a one-dimensional generalization of the supersymmetric zero-dimensional QFT discussed before. In fact, if we take the one-dimensional space to be a circle S^1 of radius β, in the limit where $\beta \to 0$, the path-integral is dominated by configurations which are independent of the position on the S^1.

EXERCISE 10.2.1. *Show this in the Euclidean formulation of the path-integral.*

In other words, in this limit we can consider the fields x and ψ to be independent of t. It is then easy to see that the action as well as the supersymmetry transformations reduce to what we have given for the case of the zero-dimensional supersymmetric QFT.

To find the conserved charges corresponding to the supersymmetry, we follow the Noether procedure. Namely, we take the variational parameter

ϵ to be a function of time, $\epsilon = \epsilon(t)$, and see how the action varies. The variation takes the form

$$\delta \int L dt = \int (-i\dot{\epsilon}\, Q - i\dot{\bar{\epsilon}}\, \overline{Q})\, dt, \tag{10.78}$$

where

$$Q = \overline{\psi}\left(i\dot{x} + h'(x)\right), \tag{10.79}$$

$$\overline{Q} = \psi\left(-i\dot{x} + h'(x)\right). \tag{10.80}$$

These are the conserved charges associated with the supersymmetry. We call them supercharges. As one can see, Q and $\overline{Q}$ are complex conjugates of each other,

$$\overline{Q} = Q^\dagger, \tag{10.81}$$

and the number of supercharges is two in real units.

Let us quantize this system. Conjugate momenta for x and ψ are given by $p = \partial L/\partial \dot{x}$ and $\pi_\psi = \partial L/\partial \dot{\psi} = i\overline{\psi}$.[4] The idea behind $\pi_\psi = i\overline{\psi}$ is that by partial integration the fermionic part of the action is given by $\int (i\overline{\psi}\dot{\psi} - h''(x)\overline{\psi}\psi)dt$. We consider this as the first order formalism of the classical mechanics $S = \int\{pdq - H(p,q)dt\}$ (which will also yield that the fermionic part of the classical Hamiltonian is $h''(x)\overline{\psi}\psi$). By moving from the classical system to the quantum system, we have the canonical commutation relation given by

$$[x, p] = i, \tag{10.82}$$

and $\{\psi, \pi_\psi\} = i$ or

$$\{\psi, \overline{\psi}\} = 1, \tag{10.83}$$

with all the other (anti-)commutators vanishing. Here the only novel feature is that between pairs of fermionic operators we have anti-commutation relations rather than commutation relations.[5] The Hamiltonian is given by

$$H = \frac{1}{2}p^2 + \frac{1}{2}(h'(x))^2 + \frac{1}{2}h''(x)(\overline{\psi}\psi - \psi\overline{\psi}). \tag{10.84}$$

Here we have chosen a specific ordering in the last term. In the classical theory $h''(x)(c\overline{\psi}\psi - (1-c)\psi\overline{\psi})$ are equivalent for any c, but in the

[4]The ordering for Grassmann derivatives has been chosen such that $(\partial/\partial\psi_1)(\psi_1\psi_2) = -\psi_2$.

[5]$\{a, b\} := ab + ba$.

quantum theory the change in c alters the Hamiltonian because of the anti-commutation relation shown in Eq. (10.83). Later we will see the reason behind the choice $c = 1/2$. To complete the quantization we must determine the representation of these operators. In the case of a bosonic variable, the (bosonic) Hilbert space is the space of square-normalizable wave-functions and the action of the operators on such a function $\Psi(x)$ is given by

$$\widehat{x}\Psi(x) = x\,\Psi(x), \; p\Psi(x) = -i\frac{d}{dx}\Psi(x). \tag{10.85}$$

(The $\widehat{x}$ notation emphasizes that x is being thought of as an operator.) For the fermionic variables, we note that the anti-commutation relations $\{\psi, \overline{\psi}\} = 1$ and $\{\psi, \psi\} = \{\overline{\psi}, \overline{\psi}\} = 0$ look like the algebra of lowering and raising operators: $[a, a^\dagger] = 1$ and $[a, a] = [a^\dagger, a^\dagger] = 0$, which we found in the simple harmonic oscillator. Indeed, if we define the fermion number operator F such that

$$F = \overline{\psi}\psi, \tag{10.86}$$

it satisfies the commutation relation with ψ and $\overline{\psi}$:

$$[F, \psi] = -\psi, \; [F, \overline{\psi}] = \overline{\psi}. \tag{10.87}$$

As in the quantization of the harmonic oscillator, we define a state $|0\rangle$ annihilated by the "lowering operator"

$$\psi|0\rangle = 0. \tag{10.88}$$

Then one can build up a tower of states multiplying $|0\rangle$ by powers of the "raising operator" $\overline{\psi}$. However, by the fermionic statistics, $\overline{\psi}^2 = 0$ and the height of the tower is just 1. Namely, the space is the two-dimensional space spanned by [6]

$$|0\rangle, \; \overline{\psi}|0\rangle. \tag{10.89}$$

With respect to this basis the operators are represented by the matrices

$$\psi = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; \overline{\psi} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{10.90}$$

The total Hilbert space of states is thus given by

$$\mathcal{H} = L^2(\mathbb{R}, \mathbb{C})|0\rangle \oplus L^2(\mathbb{R}, \mathbb{C})\overline{\psi}|0\rangle. \tag{10.91}$$

[6]We note that the algebra of ψ and $\overline{\psi}$ is the same as the Clifford algebra on $\mathbb{R}^2$. The above representaion is its unique irreducible representation.

We denote the first and second components as

$$\mathcal{H}^B = L^2(\mathbb{R},\mathbb{C})|0\rangle, \tag{10.92}$$

$$\mathcal{H}^F = L^2(\mathbb{R},\mathbb{C})\overline{\psi}|0\rangle, \tag{10.93}$$

and call them the space of bosonic states and the space of fermionic states respectively. The operator $F=\overline{\psi}\psi$ is zero on $\mathcal{H}^B$ and $F=1$ on $\mathcal{H}^F$. Thus, there is a $\mathbb{Z}_2$ grading on $\mathcal{H}$ given by $(-1)^F$.

The charges Q and $Q^\dagger=\overline{Q}$ given by Eq. (10.79) and Eq. (10.80), or

$$Q=\overline{\psi}\left(i\,p+h'(x)\right), \tag{10.94}$$

$$\overline{Q}=\psi\left(-i\,p+h'(x)\right), \tag{10.95}$$

commute with the Hamiltonian

$$[H,Q]=[H,\overline{Q}]=0, \tag{10.96}$$

and are indeed conserved charges in the quantum theory.

EXERCISE 10.2.2. *Verify the above commutation relation using the commutation relations of x, p, ψ and $\overline{\psi}$. Also show that the supercharges generate the fermionic symmetry shown in Eq. (10.75). Namely, for any combination of $(x,\psi,\overline{\psi})$, $\mathcal{O}=\mathcal{O}(x,\psi,\overline{\psi})$, we have*

$$\delta\mathcal{O}=[\widehat{\delta},\mathcal{O}],\ \ \widehat{\delta}:=\epsilon Q+\overline{\epsilon}\overline{Q}. \tag{10.97}$$

Note that the Hermiticity, as in Eq. (10.81), means $\widehat{\delta}^\dagger=-\widehat{\delta}$ (e.g., $(\epsilon Q)^\dagger=Q^\dagger\epsilon^\dagger=\overline{Q}\overline{\epsilon}=-\overline{\epsilon}\overline{Q}$), which is consistent with $(\delta\mathcal{O})^\dagger=\delta\mathcal{O}^\dagger$ since $[\widehat{\delta},\mathcal{O}]^\dagger=[\mathcal{O}^\dagger,\widehat{\delta}^\dagger]$.

The supercharges act on the Hilbert space and map bosonic states to fermionic states and vice versa. This can be considered the consequence of the relation

$$Q(-1)^F=-(-1)^FQ,\quad \overline{Q}(-1)^F=-(-1)^F\overline{Q}, \tag{10.98}$$

which follows from

$$[F,Q]=Q,\ [F,\overline{Q}]=-\overline{Q}. \tag{10.99}$$

Because of the relations $\psi^2=\overline{\psi}^2=0$, the supercharges are nilpotent:

$$\{Q,Q\}=\{\overline{Q},\overline{Q}\}=0. \tag{10.100}$$

Now let us compute the anti-commutation relation between Q and $\overline{Q}$:

$$\begin{aligned}\{Q,\overline{Q}\}=&\{\overline{\psi}(ip+h'(x)),\psi(-ip+h'(x))\}\\ =&\{\overline{\psi}ip,\psi(-i)p\}+\{\overline{\psi}h'(x),\psi h'(x)\}\\ &+i\{\overline{\psi}p,\psi h'(x)\}-i\{\overline{\psi}h'(x),\psi p\}\\ =&p^2+(h'(x))^2+i\overline{\psi}p\psi h'(x)+i\psi h'(x)\overline{\psi}p\\ &-i\overline{\psi}h'(x)\psi p-i\psi p\overline{\psi}h'(x)\\ =&p^2+(h'(x))^2+i(\overline{\psi}\psi-\psi\overline{\psi})[p,h'(x)]\\ =&p^2+(h'(x))^2+h''(x)(\overline{\psi}\psi-\psi\overline{\psi}).\end{aligned} \tag{10.101}$$

We note that this is equal to $2H$. Specifically, the supercharges obey the anti-commutation relation

$$\{Q,\overline{Q}\}=2H. \tag{10.102}$$

We shall call a quantum mechanics with a $\mathbb{Z}_2$ grading $(-1)^F$ a *supersymmetric quantum mechanics* when there are operators Q and $\overline{Q}$ obeying the (anti-)commutation relation given above. Such a quantum mechanics has special properties which will be described below. Note that we have chosen the operator ordering in Eq. (10.84) so that the resulting theory is a supersymmetric quantum mechanical system.

10.2.2. The General Structure of Hilbert Space and the Supersymmetric Index.

We now derive some general properties of supersymmetric quantum mechanics.

By definition, supersymmetric quantum mechanics (with two supercharges) is a quantum mechanics with a positive definite $\mathbb{Z}_2$-graded Hilbert space of states $\mathcal{H}$ with an even operator H as the Hamiltonian and odd operators Q and $Q^\dagger$ as supercharges. These operators obey the following commutation relations:

$$Q^2=Q^{\dagger 2}=0, \tag{10.103}$$

$$\{Q,Q^\dagger\}=2H. \tag{10.104}$$

As a consequence, the supercharges are conserved:

$$[H,Q]=[H,Q^\dagger]=0. \tag{10.105}$$

The operator defining the $\mathbb{Z}_2$-grading is denoted by $(-1)^F$. Hereafter we use $Q^\dagger$ and $\overline{Q}$ interchangeably. Since the Hamiltonian is even and the supercharges are odd, $H(-1)^F = (-1)^F H$, $Q(-1)^F = -(-1)^F Q$, $\overline{Q}(-1)^F = -(-1)^F\overline{Q}$. We denote the even subspace of $\mathcal{H}$ (on which $(-1)^F = 1$) by $\mathcal{H}^B$ and the odd subspace (on which $(-1)^F = -1$) by $\mathcal{H}^F$. The Hamiltonian preserves the decomposition $\mathcal{H} = \mathcal{H}^B \oplus \mathcal{H}^F$ while the supercharges map one subspace to the other:

$$Q, Q^\dagger : \mathcal{H}^B \longrightarrow \mathcal{H}^F, \tag{10.106}$$

$$Q, Q^\dagger : \mathcal{H}^F \longrightarrow \mathcal{H}^B. \tag{10.107}$$

The first consequence of the algebra and the positive-definiteness of the Hilbert space is that the Hamiltonian is a non-negative operator

$$H = \tfrac{1}{2}\{Q, Q^\dagger\} \geq 0. \tag{10.108}$$

A state has zero energy if and only if it is annihilated by Q and $Q^\dagger$:

$$H|\alpha\rangle = 0 \iff Q|\alpha\rangle = \overline{Q}|\alpha\rangle = 0. \tag{10.109}$$

Due to the non-negativity of the Hamiltonian, a zero energy state is a ground state. States annihilated by Q or $\overline{Q}$ are states invariant under the supersymmetry and are called supersymmetric states. What we have seen above is that a zero energy ground state is a supersymmetric state and vice versa. Thus, in what follows we call such a state a *supersymmetric ground state.*

The Hilbert space can be decomposed in terms of eigenspaces of the Hamiltonian

$$\mathcal{H} = \bigoplus_{n=0,1,\dots} \mathcal{H}_{(n)}, \quad H|_{\mathcal{H}_{(n)}} = E_n. \tag{10.110}$$

We accept the convention that $E_0 = 0 < E_1 < E_2 < \cdots$ (if there is no zero energy state we set $\mathcal{H}_{(0)} = 0$). Since Q, $\overline{Q}$ and $(-1)^F$ commute with the Hamiltonian, these operators preserve the energy levels:

$$Q, \overline{Q}, (-1)^F : \mathcal{H}_{(n)} \longrightarrow \mathcal{H}_{(n)}. \tag{10.111}$$

In particular, each energy level $\mathcal{H}_{(n)}$ is decomposed into even and odd (or bosonic and fermionic) subspaces

$$\mathcal{H}_{(n)} = \mathcal{H}^B_{(n)} \oplus \mathcal{H}^F_{(n)}, \tag{10.112}$$

and the supercharges map one subspace to the other:

$$Q, \overline{Q} : \mathcal{H}^B_{(n)} \longrightarrow \mathcal{H}^F_{(n)}\,;\ \mathcal{H}^F_{(n)} \longrightarrow \mathcal{H}^B_{(n)}. \tag{10.113}$$

Let us consider the combination $Q_1 := Q + Q^\dagger$, which obeys

$$Q_1^2 = 2H. \tag{10.114}$$

This operator preserves each energy level, mapping $\mathcal{H}^B_{(n)}$ to $\mathcal{H}^F_{(n)}$ and vice versa. Since $Q_1^2 = 2E_n$ at the nth level, as long as $E_n > 0$, Q_1 is invertible and defines an isomorphism

$$\mathcal{H}^B_{(n)} \cong \mathcal{H}^F_{(n)}. \tag{10.115}$$

Thus, the bosonic and fermionic states are paired at each excited level. At the zero energy level $\mathcal{H}_{(0)}$, however, the operator Q_1 squares to zero and does not lead to an isomorphism. In particular the bosonic and fermionic supersymmetric ground states do not have to be paired.

Now, let us consider a continuous deformation of the theory (i.e., the spectrum of the Hamiltonian deforms continuously) while preserving supersymmetry. Then the excited states (the states with positive energy) move in bosonic/fermionic pairs due to the isomorphism discussed above. Some excited level may split to several levels but the number of bosonic and fermionic states must be the same at each of the new levels. Some of the zero energy states may acquire positive energy and some positive energy states may become zero energy states, but those states must again come in pairs of bosonic and fermionic states. This means that the number of bosonic ground states minus the number of fermionic ground states is invariant. This invariant can also be represented as

$$\dim \mathcal{H}^B_{(0)} - \dim \mathcal{H}^F_{(0)} = \mathrm{Tr}\,(-1)^F e^{-\beta H}. \tag{10.116}$$

This is because in computing the trace on the right-hand side the states with positive energy come in pairs that cancel out when weighted with $(-1)^F$, and only the ground states survive. This invariant is called the supersymmetric index or the Witten index and is sometimes also denoted by the shorthand notation $\mathrm{Tr}\,(-1)^F$.

Since $Q^2 = 0$ we have a $\mathbb{Z}_2$-graded complex of vector spaces

$$\mathcal{H}^F \xrightarrow{Q} \mathcal{H}^B \xrightarrow{Q} \mathcal{H}^F \xrightarrow{Q} \mathcal{H}^B, \tag{10.117}$$

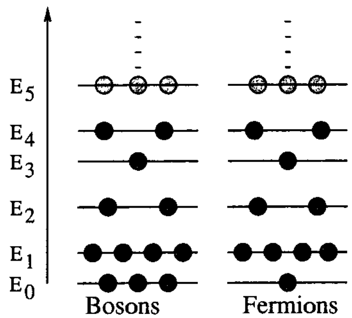


FIGURE 2.

and thus we can consider the cohomology of this complex,

$$H^B(Q) := \frac{\operatorname{Ker} Q : \mathcal{H}^B \to \mathcal{H}^F}{\operatorname{Im} Q : \mathcal{H}^F \to \mathcal{H}^B}, \qquad H^F(Q) := \frac{\operatorname{Ker} Q : \mathcal{H}^F \to \mathcal{H}^B}{\operatorname{Im} Q : \mathcal{H}^B \to \mathcal{H}^F}. \tag{10.118}$$

The complex shown in Eq. (10.117) decomposes into energy levels. At each of the excited levels, it is an exact sequence, and the cohomology vanishes. This is seen by noting that if the vector $|\alpha\rangle$ at the nth level is Q-closed, $Q|\alpha\rangle = 0$, then by the relation $1 = (QQ^\dagger + Q^\dagger Q)/(2E_n)$ that holds on $\mathcal{H}_{(n)}$ we have $|\alpha\rangle = QQ^\dagger|\alpha\rangle/(2E_n)$; namely $|\alpha\rangle$ is Q-exact. At the zero energy level $\mathcal{H}_{(0)}$, the coboundary operator is trivial, $Q = 0$, and the cohomology is nothing but $\mathcal{H}^B_{(0)}$ and $\mathcal{H}^F_{(0)}$ themselves. Thus, we have seen that the cohomology groups come purely from the supersymmetric ground states

$$H^B(Q) = \mathcal{H}^B_{(0)}, \quad H^F(Q) = \mathcal{H}^F_{(0)}. \tag{10.119}$$

In other words, the space of supersymmetric ground states is characterized as the cohomology of the Q-operator.

So far, we have assumed only the $\mathbb{Z}_2$-grading denoted by $(-1)^F$. However, in some cases there can be a finer grading such as a $\mathbb{Z}$-grading that reduces modulo 2 to the $\mathbb{Z}_2$-grading under consideration. Such is the case if there is a Hermitian operator F with integral eigenvalues such that $e^{\pi i F} = (-1)^F$. In fact, the example we discussed earlier has a fermion number F that gives a $\mathbb{Z}$ grading (although in the Hilbert space only two values of F were realized). The Hilbert space $\mathcal{H}$ can be decomposed with respect to the eigenspaces of F as $\mathcal{H} = \oplus_{p\in\mathbb{Z}}\mathcal{H}^p$ and the bosonic and fermionic subspaces

are simply $\mathcal{H}^B = \oplus_{p \text{ even}}\mathcal{H}^p$ and $\mathcal{H}^F = \oplus_{p \text{ odd}}\mathcal{H}^p$. Furthermore, if Q has charge 1,

$$[F, Q] = Q, \tag{10.120}$$

the $\mathbb{Z}_2$-graded complex shown in Eq. (10.117) splits into a $\mathbb{Z}$-graded complex

$$\cdots \xrightarrow{Q} \mathcal{H}^{p-1} \xrightarrow{Q} \mathcal{H}^p \xrightarrow{Q} \mathcal{H}^{p+1} \xrightarrow{Q} \cdots, \tag{10.121}$$

and there is a cohomology group for each $p \in \mathbb{Z}$:

$$H^p(Q) = \frac{\operatorname{Ker} Q : \mathcal{H}^p \to \mathcal{H}^{p+1}}{\operatorname{Im} Q : \mathcal{H}^{p-1} \to \mathcal{H}^p}. \tag{10.122}$$

Of course, the space of supersymmetric ground states is the sum of these cohomology groups and the bosonic/fermionic decomposition corresponds to

$$\mathcal{H}^B_{(0)} = \bigoplus_{p \text{ even}} H^p(Q), \ \mathcal{H}^F_{(0)} = \bigoplus_{p \text{ odd}} H^p(Q). \tag{10.123}$$

The Witten index is then the Euler characteristic of the complex

$$\operatorname{Tr}(-1)^F = \sum_{p\in\mathbb{Z}} (-1)^p \dim H^p(Q). \tag{10.124}$$

It is possible to generalize this consideration to the case with a $\mathbb{Z}_{2k}$-grading. This is left as an exercise for the reader.

Finally, we provide a path-integral expression for the Witten index $\operatorname{Tr}(-1)^F e^{-\beta H}$ together with that for the partition function $Z(\beta) = \operatorname{Tr} e^{-\beta H}$ on a circle of circumference β. These are given as

$$Z(\beta) = \operatorname{Tr} e^{-\beta H} = \int \mathcal{D}X\mathcal{D}\psi\mathcal{D}\overline{\psi}|_{\mathrm{AP}}\ e^{-S(X,\psi,\overline{\psi})}, \tag{10.125}$$

$$\operatorname{Tr}(-1)^F = \operatorname{Tr}(-1)^F e^{-\beta H} = \int \mathcal{D}X\mathcal{D}\psi\mathcal{D}\overline{\psi}|_{\mathrm{P}}\ e^{-S(X,\psi,\overline{\psi})}, \tag{10.126}$$

where the subscript AP and P on the measure means that we impose anti-periodic and periodic boundary conditions on the fermionic fields:

$$\begin{aligned} \mathrm{AP}: \psi(0) = -\psi(\beta)\,,\ \overline{\psi}(0) = -\overline{\psi}(\beta)\,, \\ \mathrm{P}: \psi(0) = +\psi(\beta)\,,\ \overline{\psi}(0) = +\overline{\psi}(\beta)\,. \end{aligned} \tag{10.127}$$

The fact that inserting $(-1)^F$ operator corresponds to changing the boundary conditions on fermions is clear from and follows from the fact that fermions anti-commute with $(-1)^F$. So before the trace is taken, the fermions are multiplied by an extra minus sign. What is not completely obvious is that without the insertion of $(-1)^F$ the fermions have anti-periodic boundary condition along the circle. To understand this, let us consider the correlation functions on the circle with insertions of fermions. Due to the fermion number symmetry, the number of ψ insertions must be the same as the number of $\overline{\psi}$ insertions for the correlators to be non-vanishing. We consider the simplest case with the insertion of $\overline{\psi}(t_1)$ and $\psi(t_2)$. Let us start with $t_2 = 0 < t_1 < \beta$, and increase t_2 so that it passes through t_1 and "comes back" to β. Due to the anti-commutativity of the fermionic operators, when t_2 passes through t_1, the correlation function receives an extra minus sign. Thus, the ordinary correlation function $\langle\overline{\psi}(t_1)\psi(t_2)\rangle_{S^1_\beta}$, which corresponds to the trace without $(-1)^F$, is antiperiodic under the shift $t_2 \to t_2 + \beta$. The rule (10.125)-(10.126) will also be confirmed when we explicitly compute the partition functions in simple models, both in the path-integral and operator formalisms.

We saw in the operator representation that $Tr(-1)^F\,\mathrm{e}^{-\beta H}$ is independent of β. What this means in this context is that in the path-integral representation on a circle of radius β with periodic boundary conditions, the path-integral is independent of the radius of the circle. One can directly see this in the path-integral language as well. Namely, the change of the circumference is equivalent to insertion of H in the path-integral. This can in turn be viewed as the Q variation of the field $\overline{Q}$ (in view of the commutation relation $\{Q, \overline{Q}\} = 2H$). For periodic boundary conditions on the circle, Q is a symmetry of the path-integral (this only exists for periodic boundary conditions for fermions because there is no constant non-trivial ϵ that is anti-periodic along S^1). And as in our discussion in the context of zero-dimensional QFT, the correlators that are variations of fields under symmetry operations are zero. Thus the insertion of H in the path-integral gives zero, which is equivalent to β independence of the Witten index in the path-integral representation.

10.2.3. Determination of Supersymmetric Ground States. Let us find the supersymmetric ground states of the supersymmetric potential theory. The supercharges are represented in the $(|0\rangle, \overline{\psi}|0\rangle)$ basis as

$$Q = \overline{\psi}(ip + h'(x)) = \begin{pmatrix} 0 & 0 \\ d/dx + h'(x) & 0 \end{pmatrix}, \tag{10.128}$$

$$\overline{Q} = \psi(-ip + h'(x)) = \begin{pmatrix} 0 & -d/dx + h'(x) \\ 0 & 0 \end{pmatrix}. \tag{10.129}$$

We are looking for a state $\Psi = f_1(x)|0\rangle + f_2(x)\overline{\psi}|0\rangle$ annihilated by the supercharges, $Q\Psi = \overline{Q}\Psi = 0$. The conditions on the functions $f_1(x)$ and $f_2(x)$ are the differential equations

$$\left(\frac{d}{dx} + h'(x)\right) f_1(x) = 0, \tag{10.130}$$

$$\left(-\frac{d}{dx} + h'(x)\right) f_2(x) = 0. \tag{10.131}$$

The equation itself is solved by

$$f_1(x) = c_1\,\mathrm{e}^{-h(x)}, \quad f_2(x) = c_2\,\mathrm{e}^{h(x)}. \tag{10.132}$$

It appears there are two solutions, but we are actually looking for square-normalizable functions. Whether $\mathrm{e}^{-h(x)}$ or $\mathrm{e}^{h(x)}$ is normalizable or not depends on the behaviour of the function $h(x)$ at infinity, $x \to \pm\infty$. We consider three different asymptotic behaviors of $h(x)$. (We assume polynomial growth of $|h(x)|$ at large x.)

- Case I: $h(x) \to -\infty$ as $x \to -\infty$ and $h(x) \to +\infty$ as $x \to +\infty$ (Fig. 3 (I)), or the opposite case where the sign of $h(x)$ is flipped. In this case the functions $\mathrm{e}^{-h(x)}$ and $\mathrm{e}^{h(x)}$ are diverging in either one of the infinities $x \to \pm\infty$ and are both non-normalizable. Thus, there is no supersymmetric ground state. The supersymmetric index is of course zero:

$$\mathrm{Tr}\,(-1)^F = 0. \tag{10.133}$$

- Case II: $h(x) \to \infty$ at both infinities $x \to \pm\infty$ (Fig. 3 (II)). In this case $\mathrm{e}^{-h(x)}$ decays rapidly at infinity and is normalizable, but $\mathrm{e}^{h(x)}$ is not. Thus, there is one supersymmetric ground state given by

$$\Psi = \mathrm{e}^{-h(x)}|0\rangle. \tag{10.134}$$

Since this state belongs to $\mathcal{H}^B$, the supersymmetric index is

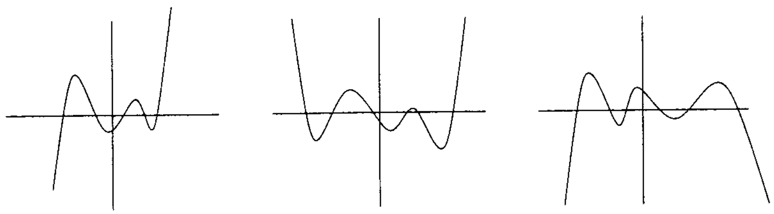
$$\mathrm{Tr}\,(-1)^F = 1. \tag{10.135}$$

- Case III: $h(x) \to -\infty$ at both infinities $x \to \pm\infty$ (Fig. 3 (III)). In this case $\mathrm{e}^{-h(x)}$ is not normalizable but $\mathrm{e}^{h(x)}$ is. Thus, there is again one supersymmetric ground state given by

$$\Psi = \mathrm{e}^{h(x)}\overline{\psi}|0\rangle. \tag{10.136}$$

This time this state belongs to $\mathcal{H}^F$ and the index is

$$\mathrm{Tr}\,(-1)^F = -1. \tag{10.137}$$



I II III

FIGURE 3

10.2.4. Example: Harmonic Oscillator. We now consider the example of a supersymmetric harmonic oscillator. Namely, the case where the function $h(x)$ is given by

$$h(x) = \frac{\omega}{2}x^2, \tag{10.138}$$

so that the potential $V(x) = \frac{1}{2}(h'(x))^2$ is that of the harmonic oscillator

$$V(x) = \frac{\omega^2}{2}x^2. \tag{10.139}$$

Note that we have a parameter ω which was set equal to ± 1 in the treatment of bosonic harmonic oscillator, see Fig. 4. As we will see later this is an important example that provides the basis of the semi-classical treatment of the more general models. (This semi-classical method will be one of the main tools in our discussion of supersymmetric QFTs in subsequent sections).

Following the previous analysis, which is valid for any polynomial $h(x)$, we find that there is one supersymmetric ground state in both the $\omega > 0$ and $\omega < 0$ cases. For $\omega > 0$, since $h(x)$ grows to $+\infty$ at infinity, $|x| \to \infty$, the supersymmetric ground state is given by

$$\Psi_{\omega>0} = e^{-\frac{1}{2}\omega x^2}|0\rangle. \tag{10.140}$$

For $\omega < 0$, $h(x)$ descends to $-\infty$ at infinity, and the state is given by

$$\Psi_{\omega<0} = e^{-\frac{1}{2}|\omega|x^2}\overline{\psi}|0\rangle. \tag{10.141}$$

Note that in both cases, the x dependence of the wave-function Ψ is of the form

$$\exp\left(-\frac{1}{2}|\omega|\,x^2\right). \tag{10.142}$$

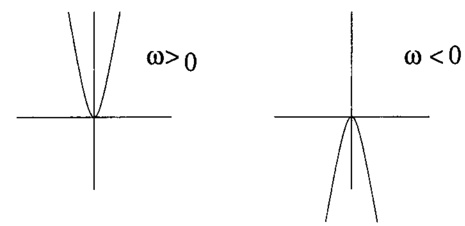


FIGURE 4.

In this model, not only the supersymmetric ground states but also the exact spectrum of the Hamiltonian can be obtained. The Hamiltonian is given by

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2x^2 + \frac{1}{2}\omega[\overline{\psi}, \psi]. \tag{10.143}$$

The part $(1/2)p^2 + (\omega^2/2)x^2 =: H_{\mathrm{osc}}$ is the same as the Hamiltonian for the simple harmonic oscillator and has the spectrum

$$\frac{|\omega|}{2},\ \frac{|\omega|}{2} + |\omega|,\ \frac{|\omega|}{2} + 2|\omega|, \ldots \tag{10.144}$$

each with multiplicity 1, as was analyzed before in the case $|\omega| = 1$. (The two pieces of H commute, so we analyze the spectra independently.) Note that the first eigenvalue $|\omega|/2$ is positive; it is called the zero point oscillation energy. Now the "fermionic part" of the Hamiltonian $(\omega/2)[\overline{\psi}, \psi] =: H_f$ is represented as the matrix

$$H_f = \frac{\omega}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{10.145}$$

in the $(|0\rangle, \overline{\psi}|0\rangle)$ basis. Note that one of the eigenvalues, $-|\omega|/2$, is negative and we call it the fermionic zero point energy. Thus the spectrum of the total Hamiltonian H is given by

$$\begin{aligned}(10.146)\qquad &\left\{\begin{array}{l} 0,\ |\omega|, 2|\omega|, \cdots \\ |\omega|, 2|\omega|, 3|\omega|, \cdots \end{array}\right. \quad \omega > 0, \\ &\left\{\begin{array}{l} |\omega|, 2|\omega|, 3|\omega| \cdots \\ 0,\ |\omega|, 2|\omega|, \cdots \end{array}\right. \quad \omega < 0.\end{aligned}$$

In both the $\omega > 0$ and $\omega < 0$ cases, the zero energy is attained as a consequence of the cancellation of the zero point oscillation energy $|\omega|/2$ and the fermionic zero point energy $-|\omega|/2$. Note the boson–fermion pairing for positive energy, as was anticipated by our general discussion of supersymmetric theories.

We now calculate the partition function and the Witten index. The Hilbert space factorizes as

$$(10.147)\qquad \mathcal{H} = \left(L^2(\mathbb{R},\mathbb{C}) \otimes |0\rangle\right) \oplus \left(L^2(\mathbb{R},\mathbb{C}) \otimes \overline{\psi}|0\rangle\right), = L^2 \otimes \mathbb{C}^2$$

where $L^2 := L^2(\mathbb{R},\mathbb{C})$ is the Hilbert space of the bosonic harmonic oscillator, on which H_{osc} acts non-trivially, and $\mathbb{C}^2 := \mathbb{C}|0\rangle \oplus \mathbb{C}\overline{\psi}|0\rangle$ is the space on which H_f acts non-trivially. Given this factorization, the partition function and the Witten index are given by

(10.148)
$$\begin{aligned} Z(\beta) &:= \mathrm{Tr}_{\mathcal{H}}\, \mathrm{e}^{-\beta H} = \mathrm{Tr}_{L^2}\, \mathrm{e}^{-\beta H_{\text{osc}}} \cdot \mathrm{Tr}_{\mathbb{C}^2}\, \mathrm{e}^{-\beta H_f} \\ \mathrm{Tr}\,(-1)^F &:= \mathrm{Tr}_{\mathcal{H}}[(-1)^F\, \mathrm{e}^{-\beta H}] = \mathrm{Tr}_{L^2}\, \mathrm{e}^{-\beta H_{\text{osc}}} \cdot \mathrm{Tr}_{\mathbb{C}^2}[(-1)^F\, \mathrm{e}^{-\beta H_F}]. \end{aligned}$$

Now we can calculate the individual parts

$$(10.149)\qquad \mathrm{Tr}_{L^2}\, \mathrm{e}^{-\beta H_{\text{osc}}} = \sum_{n=0}^{\infty} \mathrm{e}^{-\beta(n+\frac{1}{2})|\omega|} = \frac{1}{\mathrm{e}^{\frac{\beta|\omega|}{2}} - \mathrm{e}^{-\frac{\beta|\omega|}{2}}},$$

$$(10.150)\qquad \mathrm{Tr}_{\mathbb{C}^2}\, \mathrm{e}^{-\beta H_f} = \mathrm{e}^{-\frac{\beta\omega}{2}} + \mathrm{e}^{\frac{\beta\omega}{2}},$$

$$(10.151)\qquad \mathrm{Tr}_{\mathbb{C}^2}[(-1)^F\, \mathrm{e}^{-\beta H_f}] = \mathrm{e}^{-\frac{\beta\omega}{2}} - \mathrm{e}^{\frac{\beta\omega}{2}}.$$

Thus

$$(10.152)\qquad \begin{aligned} Z(\beta) = \mathrm{Tr}\, \mathrm{e}^{-\beta H} &= \frac{\mathrm{e}^{\frac{\beta\omega}{2}} + \mathrm{e}^{-\frac{\beta\omega}{2}}}{\mathrm{e}^{\frac{\beta|\omega|}{2}} - \mathrm{e}^{-\frac{\beta|\omega|}{2}}} = \coth(\beta|\omega|/2) \\ \mathrm{Tr}\,(-1)^F\, \mathrm{e}^{-\beta H} &= \frac{\mathrm{e}^{\frac{\beta\omega}{2}} - \mathrm{e}^{-\frac{\beta\omega}{2}}}{\mathrm{e}^{\frac{\beta|\omega|}{2}} - \mathrm{e}^{-\frac{\beta|\omega|}{2}}} = \frac{\omega}{|\omega|} = \pm 1. \end{aligned}$$

Note that the partition function depends on the circumference β of S^1 whereas the supersymmetric index does not.

The independence of the supersymmetric index from β can be exploited to relate it to the computation done for the zero-dimensional QFT. Namely we consider the limit $\beta \to 0$, in which case in the path-integral computation only the time independent modes contribute, and we are left with a finite-dimensional integral that is exactly the same integral we found in the context of the zero-dimensional QFT. This also explains why the Witten index is equal to the partition function for the supersymmetric system considered for the zero-dimensional QFT.

10.3. Perturbative Analysis: First Approach

Let us come back to the potential theory with general superpotential $h(x)$. The semi-classical method can be used to compute the supersymmetric index exactly, thanks to supersymmetry. This also provides the starting point for determining the supersymmetric ground states, not just the index. In the case at hand both the number of ground states and the supersymmetric index have been computed directly and the semi-classical analysis may appear as unnecessary. However, this method is extendable to more general models where exact ground state wave-functions are hard to obtain.

10.3.1. Operator Formalism. As we have seen, the supersymmetric index is unchanged under smooth deformations of the theory. It is convenient to compute the supersymmetric index in the limit where we rescale h according to

$$(10.153)\qquad h(x) \mapsto \lambda h(x), \quad \lambda \gg 1.$$

The Hamiltonian is then given by

$$(10.154)\qquad H = \frac{1}{2}p^2 + \frac{\lambda^2}{2}(h'(x))^2 + \frac{\lambda}{2}h''(x)[\overline{\psi}, \psi].$$

As $\lambda \to \infty$, the potential term becomes large and the lowest energy states become sharply peaked around the lowest values of $(h'(x))^2$. Suppose there is a critical point x_i of $h(x)$ where the potential term vanishes and let us expand the function $h(x)$ there:

$$(10.155)\qquad h(x) = h(x_i) + \frac{1}{2}h''(x_i)(x - x_i)^2 + \frac{1}{6}h'''(x_i)(x - x_i)^3 \cdots.$$

We assume that the critical point is non-degenerate, that is, $h''(x_i) \neq 0$. If we rescale the variable as $(x - x_i) = \frac{1}{\sqrt{\lambda}}(\widetilde{x} - \widetilde{x}_i)$, the expansion becomes

$$h(x) = h(x_i) + \frac{1}{2\lambda}h''(x_i)(\widetilde{x}-\widetilde{x}_i)^2 + \frac{1}{6\lambda^{3/2}}h'''(x_i)(\widetilde{x}-\widetilde{x}_i)^3 + O(\lambda^{-2}). \tag{10.156}$$

This shows that the Hamiltonian is expanded as a power series in $\lambda^{-1/2}$ as

$$\begin{aligned} H =& \lambda\left(\frac{1}{2}\widetilde{p}^2 + \frac{1}{2}h''(x_i)^2(\widetilde{x}-\widetilde{x}_i)^2 + \frac{1}{2}h''(x_i)[\overline{\psi},\psi]\right) \\ &+\lambda^{1/2}(\cdots) + (\cdots) + O(\lambda^{-1/2}), \end{aligned} \tag{10.157}$$

where $\widetilde{p} = -id/d\widetilde{x}$. Thus, we can consider the perturbation theory in $\lambda^{-1/2}$, where the leading term in the Hamiltonian is

$$H_0 = \frac{1}{2}p^2 + \frac{\lambda^2}{2}h''(x_i)^2(x-x_i)^2 + \frac{\lambda}{2}h''(x_i)[\overline{\psi},\psi]. \tag{10.158}$$

This is nothing but the Hamiltonian for the supersymmetric harmonic oscillator with $\omega = h''(x_i)$. Thus, the ground state in the perturbation theory around x_i is given by

$$\Psi_i = \mathrm{e}^{-\frac{\lambda}{2}h''(x_i)(x-x_i)^2}|0\rangle + \cdots \qquad \text{if } h''(x_i) > 0, \tag{10.159}$$

$$\Psi_i = \mathrm{e}^{-\frac{\lambda}{2}|h''(x_i)|(x-x_i)^2}\overline{\psi}|0\rangle + \cdots \qquad \text{if } h''(x_i) < 0, \tag{10.160}$$

where $+\cdots$ represents subleading terms of the power series in $\lambda^{-1/2}$. We can find the subleading terms so that the energy is strictly zero to all orders in $\lambda^{-1/2}$. (To see this, insert the expansion shown in Eq. (10.156) into either of the expressions $\mathrm{e}^{-\frac{\lambda}{2}h(x)}|0\rangle$ or $\mathrm{e}^{\frac{\lambda}{2}h(x)}\overline{\psi}|0\rangle$.) Namely, we have one supersymmetric ground state that is exact in the perturbation theory. The supersymmetric index of this perturbation theory is

$$\mathrm{Tr}\,(-1)^F = \begin{cases} 1 & h''(x_i) > 0, \\ -1 & h''(x_i) < 0. \end{cases} \tag{10.161}$$

If there are N critical points $x_1, \ldots, x_N$, and if all of them are non-degenerate, then there are N approximate supersymmetric ground states $\Psi_1, \ldots, \Psi_N$ that are exact in the perturbation theory around each critical point. Considering the sum of such perturbation theories as a deformation of the actual theory, we can compute the Witten index. It is simply the sum of the index for each perturbation theory and is given by

$$\mathrm{Tr}\,(-1)^F = \sum_{i=1}^{N}\mathrm{sign}(h''(x_i)). \tag{10.162}$$

It is easy to see that this agrees with the exact result obtained earlier. For example, in Case II, the number of x_i with $h''(x_i) > 0$ is greater by one compared to the number of x_i with $h''(x_i) < 0$, and the sum shown in Eq. (10.162) equals 1.

As we have seen, the number of exact supersymmetric ground states is at most 1. Thus, although the above semi-classical analysis reproduces the exact result for the Witten index, it fails for the actual spectrum of supersymmetric ground states. This means that the states $\Psi_1, \ldots, \Psi_N$ are not exactly the supersymmetric ground states of the actual theory. The failure cannot be captured by perturbation theory since the Ψ_i are supersymmetric ground states to *all orders* in the series expansion in $\lambda^{-1/2}$. The effect that gives energy to most of these states is non-perturbative in $\lambda^{-1/2}$. Later in this chapter, we will identify this non-perturbative effect and show how to recover the exact result by taking it into account. The non-perturbative effect is called "quantum tunneling."

10.3.2. Path-Integral Approach — Localization Principle. We next evaluate the Witten index using the path-integral. As we noted earlier, this is done by computing the path-integral on a circle of arbitrary radius (we choose it to be 1),

$$\mathrm{Tr}\,(-1)^F = \int \mathcal{D}X\mathcal{D}\psi\mathcal{D}\overline{\psi}|_P\ \mathrm{e}^{-S_E(X,\psi,\overline{\psi})}, \tag{10.163}$$

where the periodic boundary condition is imposed on the fermions. The Euclidean action is given by

$$S_E = \int_0^{2\pi}\left\{\frac{1}{2}\left(\frac{dx}{d\tau}\right)^2 + \frac{1}{2}(h'(x))^2 + \overline{\psi}\frac{d}{d\tau}\psi + h''(x)\overline{\psi}\psi\right\}d\tau. \tag{10.164}$$

This action is invariant under the supersymmetry transformations

$$\begin{aligned} \delta x &= \epsilon\overline{\psi} - \overline{\epsilon}\psi, \\ \delta\psi &= \epsilon\left(-\frac{dx}{d\tau} + h'(x)\right), \\ \delta\overline{\psi} &= \overline{\epsilon}\left(\frac{dx}{d\tau} + h'(x)\right), \end{aligned} \tag{10.165}$$

which is compatible with the periodic fermionic (and bosonic) boundary conditions.

Recall from our discussion of the zero-dimensional QFT that if the action is invariant under some supersymmetries, the path-integral localizes to regions where the supersymmetric variations of fermionic fields vanish. This follows simply from the integration rules over fermions and applies to any supersymmetric QFT in any dimension. We will thus apply the localization principle to this one-dimensional QFT. By the localization principle, the path-integral is concentrated on the locus where the right-hand side of the fermion variations $\delta\psi$ and $\delta\overline{\psi}$ vanishes. Namely, it is concentrated on

$$\frac{dx}{d\tau} = h'(x) = 0, \tag{10.166}$$

which is given by the constant maps to the critical points $x_1, \dots, x_N$.

The path-integral around the critical point x_i is given by the Gaussian integral, keeping only the quadratic terms in the action. Setting $\xi := x - x_i$, the action in the quadratic approximation is given by

$$S_E^{(i)} = \int_0^{2\pi} \left\{ \frac{1}{2}\xi \left(-\frac{d^2}{d\tau^2} + h''(x_i)^2 \right) \xi + \overline{\psi} \left(\frac{d}{d\tau} + h''(x_i) \right) \psi \right\} d\tau. \tag{10.167}$$

The path-integral around the constant map to x_i is given by

$$\begin{aligned}
\int \mathcal{D}\xi\mathcal{D}\psi\mathcal{D}\overline{\psi}|_P \, e^{-S_E^{(i)}} =& \frac{\det(\partial_\tau + h''(x_i))}{\sqrt{\det(-\partial_\tau^2 + (h''(x_i))^2)}}, \\
=& \frac{\prod_{n\in\mathbb{Z}}(in + h''(x_i))}{\sqrt{\prod_{n\in\mathbb{Z}}(n^2 + (h''(x_i))^2)}} \\
=& \frac{h''(x_i)}{|h''(x_i)|}.
\end{aligned}$$

Summing up the contributions of all the critical points, we obtain

$$\operatorname{Tr}(-1)^F e^{-\beta H} = \sum_{i=1}^{N} \operatorname{sign}(h''(x_i)), \tag{10.168}$$

which is the same result obtained in the operator formalism. Note also that, as before, the non-constant modes along the S^1 (indexed by Fourier mode n) cancel among bosons and fermions and we are left with the constant mode, which thus leads exactly to the computation for the supersymmetric QFT in dimension 0.

Note that the periodic boundary condition for the fermions is crucial for the existence of supersymmetry, as shown by Eq. (10.165), in the path-integral. If we imposed anti-periodic boundary conditions there would be no supersymmetry to begin with and our arguments about localization would not hold. This is the reason the partition function without the insertion of

$(-1)^F$ (i.e., with anti-periodic boundary conditions for fermions) does not localize near the critical points.

10.3.3. Multi-Variable case. Let us consider a supersymmetric potential theory with many variables. We consider a theory of n bosonic and $2n$ fermionic variables x^I, ψ^I, $\overline{\psi}^I$ $(I = 1, \dots, n)$, where ψ^I and $\overline{\psi}^I$ are complex conjugate of each other. The Hamiltonian and the supercharges in this case is a simple generalization of the ones in the single-variable case:

$$\begin{aligned}
H =& \frac{1}{2}\sum_I p_I^2 + \frac{1}{2}(\partial_I h(x))^2 + \frac{1}{2}(\partial_I\partial_J h)[\overline{\psi}^I, \psi^J], \\
Q =& \overline{\psi}^I(ip_I + \partial_I h), \\
\overline{Q} =& \psi^I(-ip_I + \partial_I h),
\end{aligned} \tag{10.169}$$

where $h(x)$ is a function of $x = (x^1, \dots, x^n)$. It is in general difficult to find the supersymmetric ground states. (If $h(x) = \sum_{I=1}^N h(x^I)$, however, we have a decoupled system, and the supersymmetric ground state is the tensor product of the supersymmetric ground states of the single-variable theories.)

We now perform the semi-classical analysis to find the supersymmetric ground states. As before, we rescale $h(x)$ as $\lambda h(x)$ with $\lambda \gg 1$. Assume that the critical points $\{x_1, \cdots, x_N\}$ of $h(X)$ are isolated and non-degenerate. Near each critical point x_i we can choose coordinates $\xi_{(i)}$ such that

$$\begin{aligned}
h(x) =& h(x_i) + \frac{1}{2}\partial_I\partial_J h(x_i)\,(x^I - x_i^I)(x^J - x_i^J) + \cdots \\
=& h(x_i) + \sum_I c_I^{(i)}(\xi_{(i)}^I)^2 + \cdots .
\end{aligned} \tag{10.170}$$

In the large λ limit, the ground state wave-functions are localized near the critical points and the approximate ground states around x_i are given by

$$\Psi_i = e^{-\sum_{I=1}^n \lambda |c_I^{(i)}|(\xi_{(i)}^I)^2} \prod_{J:c_J^{(i)}<0} \overline{\psi}^J|0\rangle. \tag{10.171}$$

Note that the number of $\overline{\psi}^I$'s is $\#\{J|c_J^{(i)} < 0\}$, which is the number of negative eigenvalues of the Hessian $\partial_I\partial_J h$ at x_i. This number is called the *Morse index* of the function $h(x)$ at the critical point x_i. Thus,

$$\text{number of } \overline{\psi}^I\text{'s in } \Psi_i = \text{Morse index of } h(x) \text{ at } x_i =: \mu_i. \tag{10.172}$$

The approximate ground state is bosonic if the Morse index is even and is fermionic if the Morse index is odd. The Witten index of the system is therefore

$$\mathrm{Tr}\,(-1)^F = \sum_{i=1}^{N}(-1)^{\mu_i}, \tag{10.173}$$

As in the single-variable case, it is not necessarily the case that there are as many supersymmetric ground states as the number of critical points of $h(x)$. It is quite likely that some non-perturbative effect lifts some of this degeneracy. As promised before, this will be identified later in this chapter as the quantum tunneling effect.

There are, however, cases where the number of critical points does agree with the number of supersymmetric ground states. For example, if each of the critical points has even Morse index, then all these approximate ground states Ψ_i are really the supersymmetric ground states. This is because lifting of zero energy states to positive energy states is possible only for pairs of bosonic and fermionic states. In particular, in the large λ limit where all other states have large positive energies, the number of supersymmetric ground states is the same as the number of critical points. Likewise, if each of the critical points has odd Morse index, then all the ground states are fermionic and these span the space of supersymmetric ground states, at least in the large λ limit. In the next example we consider a model to which this remark applies.

10.3.4. Complex Case, $n = 2m$ (Landau–Ginzburg Model). Let us consider the case with an even number of variables $n = 2m$ and let us combine the $2m$ bosonic variables $(x^I) = (x^1, y^1, \dots, x^m, y^m)$ into m complex variables

$$z^i = x^i + iy^i,\ \ i = 1, \dots, m. \tag{10.174}$$

We consider the case in which the function $h(x^I)$ is the real part of a holomorphic function $-W(z^i)$ of $(z^i) = (z^1, \dots, z^m)$ (the minus sign here is not essential; it is simply to match convention in later sections):

$$h(x^I) = -\mathrm{Re}\, W(z^i). \tag{10.175}$$

We introduce the complex notation also for the fermions:

$$\psi^i = \psi^{x^i} + i\psi^{y^i},\ \ \overline{\psi}^i = \overline{\psi}^{x^i} + i\overline{\psi}^{y^i}, \tag{10.176}$$

$$\psi^{\bar{\imath}} = \psi^{x^i} - i\psi^{y^i},\ \ \overline{\psi}^{\bar{\imath}} = \overline{\psi}^{x^i} - i\overline{\psi}^{y^i}. \tag{10.177}$$

They are related under the Hermitian conjugation by $(\psi^i)^\dagger = \overline{\psi}^{\bar{\imath}}$ and $(\psi^{\bar{\imath}})^\dagger = \overline{\psi}^i$. The Lagrangian of the system is expressed as

$$\begin{aligned} L = &\sum_{i=1}^{m}\left(|\dot{z}_i|^2 + i\overline{\psi}^i\partial_t\psi^{\bar{\imath}} + i\overline{\psi}^{\bar{\imath}}\partial_t\psi^i - \frac{1}{4}|\partial_i W|^2 \right) \\ &-\frac{1}{2}\sum_{i,j}\left(\partial_i\partial_j W\psi^i\overline{\psi}^j + \partial_{\bar{\imath}}\partial_{\bar{\jmath}}\overline{W}\psi^{\bar{\imath}}\overline{\psi}^{\bar{\jmath}}\right). \end{aligned} \tag{10.178}$$

This theory is the one-dimensional QFT version of the zero-dimensional Landau–Ginzburg theory discussed before. We shall refer to the holomorphic function W as the *superpotential.*

We now assume that W has N critical points $p_1, \dots, p_N$ that are all non-degenerate, $\det \partial_i\partial_j W(p_a) \neq 0$. At each critical point one can expand the holomorphic function $W(z^i)$ in the form

$$W(z) = \sum_{i=1}^{m}(z^i)^2 + O((z^i)^3), \tag{10.179}$$

by an affine change of coordinates if necessary. Since $(x+iy)^2 = x^2 - y^2 + 2ixy$, the function $h(x^I) = \mathrm{Re}\, W(z^i)$ is written as

$$h(x^I) = \sum_{i=1}^{m}\left\{-(x^i)^2 + (y^i)^2\right\} + O((z^i)^3). \tag{10.180}$$

In particular, the Morse index is $\mu = m$. This is true at all critical points. Namely, the N approximate ground states defined around the N critical points of W all have $(-1)^F = (-1)^m$; they are all bosonic or all fermionic. Thus, there is no chance for some of them to be lifted to positive energy states. We see that the number of supersymmetric vacua is at least N and the actual number is also N for a sufficiently large scaling parameter λ.

This system has more symmetry compared to the models we have been studying. As in the zero-dimensional version, it has extended supersymmetry. We recall that in the supersymmetric quantum mechanics considered so far, the supersymmetry transformation has one complex parameter ϵ. In the present model, there are actually two complex fermionic parameters ϵ_+

and ϵ_-, where the transformation rules are

$$\begin{aligned} &\delta z^i = \epsilon_+\overline{\psi}^i - \epsilon_-\psi^i, && \delta \overline{z}^{\bar{i}} = -\bar{\epsilon}_+\psi^{\bar{i}} + \bar{\epsilon}_-\overline{\psi}^{\bar{i}}, \\ &\delta\psi^i = i\bar{\epsilon}_-\dot{z}^i - \epsilon_+\partial_{\bar{i}}\overline{W}, && \delta\overline{\psi}^{\bar{i}} = -i\epsilon_-\dot{\overline{z}}^{\bar{i}} - \bar{\epsilon}_+\partial_i W, \\ &\delta\overline{\psi}^i = -i\bar{\epsilon}_+\dot{z}^i - \epsilon_-\partial_{\bar{i}}\overline{W}, && \delta\psi^{\bar{i}} = i\epsilon_+\dot{\overline{z}}^{\bar{i}} - \bar{\epsilon}_-\partial_i W. \end{aligned} \tag{10.181}$$

If we set $\epsilon_+ = \epsilon_- = \epsilon$, we recover the original supersymmetry. By the Noether procedure, we find the four supercharges $Q_\pm$ and $\overline{Q}_\pm$ that generate the supersymmetry transformations via $\delta\mathcal{O} = [\widehat{\delta}, \mathcal{O}]$ with

$$\widehat{\delta} = i\epsilon_+ Q_- - i\epsilon_- Q_+ - i\bar{\epsilon}_+\overline{Q}_- + i\bar{\epsilon}_-\overline{Q}_+. \tag{10.182}$$

These are expressed as

$$Q_+ = \psi^i p_i - \frac{i}{2}\psi^{\bar{i}}\partial_{\bar{i}}\overline{W},\ Q_- = \overline{\psi}^i p_i + \frac{i}{2}\overline{\psi}^{\bar{i}}\partial_{\bar{i}}\overline{W}, \tag{10.183}$$

$$\overline{Q}_+ = \overline{\psi}^{\bar{i}} p_{\bar{i}} + \frac{i}{2}\overline{\psi}^i\partial_i W,\ \overline{Q}_- = \psi^{\bar{i}} p_{\bar{i}} - \frac{i}{2}\psi^i\partial_i W. \tag{10.184}$$

We note that the ordinary supercharges Q and $\overline{Q}$ are simply the linear combinations $Q = i(Q_- + \overline{Q}_+)$ and $\overline{Q} = -i(\overline{Q}_- + Q_+)$, which is consistent with $\delta_{\epsilon_\pm = \epsilon} = \epsilon Q + \bar{\epsilon}\overline{Q}$. Under an appropriate choice of operator ordering for the Hamiltonian H, these supercharges obey the anti-commutation relations

$$\{Q_\alpha, \overline{Q}_\beta\} = \delta_{\alpha\beta}H, \tag{10.185}$$

$$\{Q_\alpha, Q_\beta\} = \{\overline{Q}_\alpha, \overline{Q}_\beta\} = 0. \tag{10.186}$$

An extremely important fact is that the system can be considered as a supersymmetric quantum mechanics with the supercharges $Q = \overline{Q}_+$; this itself obeys our favorite relation $Q^2 = 0$ and $\{Q, Q^\dagger\} = H$. (The choice of $\overline{Q}_+$ is not essential; any one of the four $\overline{Q}_\pm$ and $Q_\pm$ will do the job.) In particular, one can identify the space of supersymmetric ground states as the $\overline{Q}_+$-cohomology group.

In fact, this last remark enables us to determine the space of ground states exactly. To see this, we first focus on the fermion number operator. The system has, as before, the fermion number symmetry F under which ψ^I and $\overline{\psi}^I$ have opposite charges. One can consider another "fermion number" operator

$$F_V = \sum_{i=1}^{m}(\overline{\psi}^{\bar{i}}\psi^i - \overline{\psi}^i\psi^{\bar{i}}), \tag{10.187}$$

under which $\overline{\psi}^i$ and $\psi^{\bar{i}}$ have the same charge but it is opposite to the charge of $\overline{\psi}^{\bar{i}}$ and ψ^i. This is not a symmetry of the system since the Lagrangian

shown in Eq. (10.178) is not invariant, but there is nothing wrong in considering it as an operator acting on the Hilbert space of states. Now let us consider conjugating $\overline{Q}_+$ with the operator $\sqrt{\lambda}^{F_V}$:

$$\overline{Q}_+ \to \overline{Q}_{+\lambda} = \sqrt{\lambda}^{-F_V}\overline{Q}_+\sqrt{\lambda}^{F_V}. \tag{10.188}$$

Since $\overline{\psi}^{\bar{i}}$ and $\overline{\psi}^i$ have opposite charges under F_V, the effect of cojugations is equivalent to the rescaling of the superpotential $W \to \lambda W$ in the expression of $\overline{Q}_+$ (up to an overall constant multiplication). Since the $\overline{Q}_+$-cohomology and $\overline{Q}_{+\lambda}$-cohomology are isomorphic — under the isomorphism $\sqrt{\lambda}^{F_V}$, the space of supersymmetric ground states is invariant under the rescaling parameter λ — it follows that one can use the result of the semi-classical analysis at large λ, as far as the spectrum of ground states is concerned. Thus, there is a one-to-one correspondence

$$\text{supersymmetric ground states} \longleftrightarrow \text{critical points of } W. \tag{10.189}$$

The ground states all have the same fermion number $(-1)^m$.

The extended supersymmetry has another advantage. Let us consider a correlation function on the circle S^1, where we put periodic boundary condition for fermions,

$$\langle O(\tau_1)\cdots O(\tau_s)\rangle = \int \mathcal{D}z\mathcal{D}\psi\mathcal{D}\overline{\psi}\Big|_P e^{-S(z,\psi,\overline{\psi})}O(\tau_1)\cdots O(\tau_s). \tag{10.190}$$

We note that the (Euclidean) time derivative of z^i is the $\overline{Q}_\pm$ commutator

$$\frac{dz^i}{d\tau} = -i\dot{z}^i = -\{\overline{Q}_+, \psi^i\} = -\{\overline{Q}_-, \overline{\psi}^i\}. \tag{10.191}$$

Thus, if an operator $\mathcal{O}$ commutes with $\overline{Q}_\pm$, $[\overline{Q}_\pm, \mathcal{O}] = 0$, the correlation function $\langle \frac{dz^i}{d\tau}\mathcal{O}\rangle$ vanishes,

$$\left\langle \frac{dz^i}{d\tau}\mathcal{O}\right\rangle = -\langle\{\overline{Q}_+, \psi^i\}\mathcal{O}\rangle = 0. \tag{10.192}$$

It is clear from Eq. (10.181) that a holomorphic combination of the coordinates z^i is $\overline{Q}_\pm$-invariant:

$$[\overline{Q}_\pm, f(z^i)] = 0, \text{ if } \frac{\partial f}{\partial \overline{z}^{\bar{i}}} = 0. \tag{10.193}$$

Thus, the correlation function $\langle \frac{dz^i}{d\tau} f_1(z^i(\tau_1))\cdots f_s(z^i(\tau_s))\rangle$ vanishes. This means that

$$\frac{\partial}{\partial\tau_a}\langle f_1(z^i(\tau_1))\cdots f_s(z^i(\tau_s))\rangle = 0,\ a = 1, \dots, s. \tag{10.194}$$

The correlation function of operators $f_1(z^i(\tau_1)), \ldots, f_s(z^i(\tau_s))$ is independent of the "insertion points" $\tau_1, \ldots, \tau_s$. One can actually push the computation further; the correlator is given by

$$\langle f_1(z^i(\tau_1)) \cdots f_s(z^i(\tau_s)) \rangle = \sum_{a=1}^{N} f_1(p_a) \cdots f_s(p_a), \tag{10.195}$$

where $p_1, \ldots, p_a$ are the critical points of W (assumed to be non-degenerate). This is exactly as in the zero-dimensional case discussed before. In fact, the localization principle tells us that the path-integral localizes on the $\overline{Q}_\pm$ fixed points; the locus where $dz^i/d\tau = 0$ (and $\partial_i W = 0$). This reduces the computation to zero dimensions and gives us Eq. (10.195). Similarly, we can develop the notion of chiral ring, etc., as was done in the context of the zero-dimensional QFT. The fact that the correlation functions of chiral fields do not depend on τ is a hint of the topological nature of this quantum mechanical system. It also implies that the chiral ring is defined without reference to any particular points τ_i.

10.4. Sigma Models

We now move on to supersymmetric systems with more interesting target manifolds. We will see a beautiful relation between the topology of the target manifold and the ground state structure of the supersymmetric sigma model. We also consider turning on superpotentials on the target manifold, viewed as Morse functions on the manifold, which leads to a physical realization of Morse theory.

10.4.1. SQM on a Riemannian Manifold. We consider the supersymmetric quantum mechanics of a particle moving in a Riemannian manifold M of dimension n with metric g. This is the one-dimensional analogue of the supersymmetric non-linear sigma model in $1+1$ dimensions, which will be the main focus of later sections. We assume that M is oriented and compact, although compactness will be relaxed when we later deform the theory by a potential. We denote a (generic) set of local coordinates of M by $x^I = x^1, \ldots, x^n$.

The theory involves n bosonic variables ϕ^I representing the position of the particle and their fermionic partners ψ^I and $\overline{\psi}^I$, which are complex conjugates of each other. More formally, if we denote by $\mathcal{T}$ the one-dimensional manifold parametrized by the time t, the bosonic variables define a map

$$\phi : \mathcal{T} \to M, \tag{10.196}$$

which is represented locally as $x^I \circ \phi = \phi^I$. The fermionic variables define sections

$$\psi, \overline{\psi} \in \Gamma(\mathcal{T}, \phi^* TM \otimes \mathbb{C}), \tag{10.197}$$

which are complex conjugates of each other, where ψ is locally represented by $\psi = \psi^I (\partial/\partial x^I)|_\phi$. The Lagrangian of the system is given by

$$L = \frac{1}{2} g_{IJ} \dot{\phi}^I \dot{\phi}^J + \frac{i}{2} g_{IJ} (\overline{\psi}^I D_t \psi^J - D_t \overline{\psi}^I \psi^J) - \frac{1}{2} R_{IJKL} \psi^I \overline{\psi}^J \psi^K \overline{\psi}^L, \tag{10.198}$$

where

$$D_t \psi^I = \partial_t \psi^I + \Gamma^I_{JK} \partial_t \phi^J \psi^K, \tag{10.199}$$

with Γ^I_{JK} the Christoffel symbol of the Levi–Civita connection. Under the supersymmetry transformations

$$\delta \phi^I = \epsilon \overline{\psi}^I - \overline{\epsilon} \psi^I, \tag{10.200}$$

$$\delta \psi^I = \epsilon (i \dot{\phi}^I - \Gamma^I_{JK} \overline{\psi}^J \psi^K), \tag{10.201}$$

$$\delta \overline{\psi}^I = \overline{\epsilon} (-i \dot{\phi}^I - \Gamma^I_{JK} \overline{\psi}^J \psi^K), \tag{10.202}$$

the action is invariant

$$\delta \int L \, dt = 0, \tag{10.203}$$

and the classical system is supersymmetric. By the Noether procedure, we find the corresponding conserved charges (supercharges)

$$Q = i g_{IJ} \overline{\psi}^I \dot{\phi}^J, \tag{10.204}$$

$$\overline{Q} = -i g_{IJ} \psi^I \dot{\phi}^J. \tag{10.205}$$

The Lagrangian is also invariant under the phase rotation of the fermions

$$\psi^I \to e^{-i\gamma} \psi^I, \ \overline{\psi}^I \to e^{i\gamma} \overline{\psi}^I. \tag{10.206}$$

The corresponding Noether charge is given by

$$F = g_{IJ} \overline{\psi}^I \psi^J. \tag{10.207}$$

Let us quantize the system. The conjugate momenta for ϕ^I and ψ^I are given by $p_I = \partial L/\partial\dot{\phi}^I = g_{IJ}\dot{\phi}^J$ and $\pi_{\psi I} = ig_{IJ}\overline{\psi}^J$ and the canonical (anti-)commutation relations are given by

$$[\phi^I, p_J] = i\delta^I{}_J, \tag{10.208}$$

$$\{\psi^I, \overline{\psi}^J\} = g^{IJ}, \tag{10.209}$$

with all other (anti-)commutators vanishing. In terms of the conjugate momenta p_I the supercharges are given by

$$Q = i\overline{\psi}^I p_I, \quad \overline{Q} = -i\psi^I p_I. \tag{10.210}$$

To find the quantum mechanical expression for the Hamiltonian H, we face the usual problem of operator ordering. Here we fix this ambiguity so that the supersymmetry relation

$$\{Q, \overline{Q}\} = 2H \tag{10.211}$$

holds. We also note that the supercharges Q and $\overline{Q}$ have opposite F-charges

$$[F, Q] = Q, \ [F, \overline{Q}] = -\overline{Q}. \tag{10.212}$$

As a consequence, F commutes with the Hamiltonian

$$[H, F] = 0. \tag{10.213}$$

Namely, F is a conserved charge in the quantum theory. It is easy to see that F generates the phase rotation, as shown by Eq. (10.206). We call this F a fermion number operator.

Quantization is not complete unless we specify the representation of the above algebra of observables. Here there is a natural one. It is represented on the space of differential forms,

$$\mathcal{H} = \Omega(M) \otimes \mathbb{C}, \tag{10.214}$$

equipped with the Hermitian inner product

$$(\omega_1, \omega_2) = \int_M \overline{\omega}_1 \wedge *\omega_2. \tag{10.215}$$

The observables are represented on this Hilbert space as the operators given by

$$\phi^I = x^I\times, \tag{10.216}$$

$$p_I = -i\nabla_I, \tag{10.217}$$

$$\overline{\psi}^I = dx^I\wedge, \tag{10.218}$$

$$\psi^I = g^{IJ} i_{\partial/\partial x^J}, \tag{10.219}$$

where i_V is the operation of contraction of the differential form with the vector field V. If we denote by $|0\rangle$ the vector annihilated by all ψ^I's (as was used in the previous treatment of the representation of the algebra of fermions), we find the following correspondence,

$$|0\rangle \leftrightarrow 1 \tag{10.220}$$

$$\overline{\psi}^I|0\rangle \leftrightarrow dx^I \tag{10.221}$$

$$\overline{\psi}^I\overline{\psi}^J|0\rangle \leftrightarrow dx^I \wedge dx^J \tag{10.222}$$

$$\cdots \leftrightarrow \cdots \tag{10.223}$$

$$\overline{\psi}^1\cdots\overline{\psi}^n|0\rangle \leftrightarrow dx^1 \wedge \cdots \wedge dx^n. \tag{10.224}$$

Since $[F, \overline{\psi}^I] = \overline{\psi}^I$, the fermion number ($F$-charge) of the state corresponding to a p-form is p. Thus the decomposition by form-degree

$$\mathcal{H} = \bigoplus_{p=0}^{n} \Omega^p(M) \otimes \mathbb{C} \tag{10.225}$$

coincides with the grading by the fermion number.

The supercharge Q is then given by

$$Q = i\overline{\psi}^I p_I = dx^I \wedge \nabla_I = dx^I \wedge \frac{\partial}{\partial x^I} = d, \tag{10.226}$$

which is the exterior derivative acting on differential forms. The other supercharge $\overline{Q}$ is defined as the Hermitian conjugate of Q,

$$\overline{Q} = Q^\dagger = d^\dagger. \tag{10.227}$$

The Hamiltonian H is defined so that the supersymmetry relation, Eq. (10.211), holds and is represented as

$$H = \tfrac{1}{2}\{Q, \overline{Q}\} = \tfrac{1}{2}(d\,d^\dagger + d^\dagger\,d) = \tfrac{1}{2}\Delta, \tag{10.228}$$

where Δ is the Laplace–Beltrami operator. Thus, the supersymmetric ground states, or the zero energy states, are simply the harmonic forms

$$\mathcal{H}_{(0)} = \mathcal{H}(M,g) = \bigoplus_{p=0}^{n} \mathcal{H}^p(M,g), \tag{10.229}$$

where $\mathcal{H}(M,g)$ is the space of harmonic forms of the Riemannian manifold (M,g) and $\mathcal{H}^p(M,g)$ is the space of harmonic p-forms.

We recall that the space of supersymmetric ground states can be characterized as the cohomology of the Q-operator. In the present case, since there is a conserved charge F with

$$[F,Q] = Q, \tag{10.230}$$

the Q-complex and the Q-cohomology are graded by the fermion number $F = p$. Since this is the form-degree and Q is identified as the exterior derivative d, the graded Q-cohomology is the de Rham cohomology

$$H^p(Q) = H^p_{DR}(M). \tag{10.231}$$

From the general structure of supersymmetric quantum mechanics, we have

$$\mathcal{H}_{(0)} = \mathcal{H}(M,g) \cong H^\bullet(Q) = H^\bullet_{DR}(M). \tag{10.232}$$

With respect to the F-charge, this refines to

$$\mathcal{H}^p(M,g) \cong H^p_{DR}(M). \tag{10.233}$$

The supersymmetric index is the Euler characteristic of the Q-complex, namely

$$\operatorname{Tr}(-1)^F = \sum_{p=0}^{n}(-1)^p \dim H^p(Q) = \sum_{p=0}^{n}(-1)^p \dim H^p_{DR}(Q) = \chi(M), \tag{10.234}$$

which is the Euler number of the manifold. Here deformation invariance is the familiar statement that the harmonic forms are equal to the de Rham cohomology classes, which are diffeomorophism invariants.

EXERCISE 10.4.1. *Using the independence of Witten index* $\operatorname{Tr}(-1)^F e^{-\beta H}$ *from* β, *derive an expression for the Euler number of a manifold in terms of an integral involving the Riemann curvature tensor over the manifold. In particular, consider the limit* $\beta \to 0$ *of the path-integral, and argue that the finite-action field configurations contributing to the path-integral localize to constant modes independent of time, reducing the path-integral to a zero-dimensional QFT, involving an integration over the manifold (this can also be derived using the localization principle and the supersymmetry transformation of the fermionic fields). Moreover, the fermionic integration brings down Riemann curvature terms from the quartic fermionic term in the action Eq. (10.198) leading to the desired integral over the manifold.*

10.4.2. Deformation by Potential Term. We can modify the Lagrangian by adding a potential term constructed by a real-valued function h on M,

$$h : M \longrightarrow \mathbb{R}. \tag{10.235}$$

The modification is given by addition of

$$\Delta L = -\tfrac{1}{2} g^{IJ}\partial_I h \partial_J h - D_I \partial_J h \overline{\psi}^I \psi^J \tag{10.236}$$

to the Lagrangian, where

$$D_I \partial_J h = \partial_I \partial_J h - \Gamma^K_{IJ}\partial_K h. \tag{10.237}$$

The supersymmetry transformations are modified as

$$\begin{aligned}
\delta\phi^I &= \epsilon\overline{\psi}^I - \bar\epsilon\psi^I, && (10.238)\\
\delta\psi^I &= \epsilon(i\dot\phi^I - \Gamma^I_{JK}\overline{\psi}^J\psi^K + g^{IJ}\partial_J h), && (10.239)\\
\delta\overline{\psi}^I &= \bar\epsilon(-i\dot\phi^I - \Gamma^I_{JK}\overline{\psi}^J\psi^K + g^{IJ}\partial_J h). && (10.240)
\end{aligned}$$

The supercharges are modified accordingly:

$$\begin{aligned}
Q &= \overline{\psi}^I(ig_{IJ}\dot\phi^J + \partial_I h) = \overline{\psi}^I(ip_I + \partial_I h), && (10.241)\\
\overline{Q} &= \psi^I(-ig_{IJ}\dot\phi^J + \partial_I h) = \psi^I(-ip_I + \partial_I h). && (10.242)
\end{aligned}$$

The fermion rotation symmetry $\psi^I \to e^{i\alpha}\psi^I$ and $\overline{\psi}^I \to e^{-i\alpha}\overline{\psi}^I$ is preserved and the conserved charge is again

$$F = g_{IJ}\overline{\psi}^I\psi^J. \tag{10.243}$$

The canonical commutation relation is not modified, and we can use the same representation of the algebra of variables as before. In particular, the Hilbert space of states is the space of differential forms $\Omega^\bullet(M)$. We see that the supercharges are represented as

$$\begin{aligned}
Q &= d + d\phi^I \wedge \partial_I h = d + dh\wedge = e^{-h} d\, e^{h} =: d_h, && (10.244)\\
\overline{Q} &= (d + dh\wedge)^\dagger = e^{h} d^\dagger e^{-h} = d_h^\dagger. && (10.245)
\end{aligned}$$

The Hamiltonian is chosen so that the supersymmetry relation holds, namely

$$H = \frac{1}{2}\{Q, \overline{Q}\} = \frac{1}{2}(d_h d_h^\dagger + d_h^\dagger d_h). \tag{10.246}$$

The space of supersymmetric ground states is isomorphic to the cohomology group of the Q-operator. Since the conserved fermion number F counts the form-degree, and Q has charge 1, the Q-complex and cohomology are graded by the form-degree. However, this Q and the Q before the deformation are related by the similarity transformation

$$Q = \mathrm{e}^{-h} Q_{h=0}\, \mathrm{e}^{h}, \tag{10.247}$$

and the Q-complex is isomorphic to the old one

(10.248)
$$\begin{tikzcd}
0 \arrow[r] & \Omega^0(M) \arrow[r, "d"] \arrow[d, "\mathrm{e}^{-h}"] & \Omega^1(M) \arrow[r, "d"] \arrow[d, "\mathrm{e}^{-h}"] & \cdots \arrow[r, "d"] & \Omega^n(M) \arrow[r, "d"] \arrow[d, "\mathrm{e}^{-h}"] & 0 \\
0 \arrow[r] & \Omega^0(M) \arrow[r, "\mathrm{e}^{-h} d\, \mathrm{e}^{h}"] & \Omega^1(M) \arrow[r, "\mathrm{e}^{-h} d\, \mathrm{e}^{h}"] & \cdots \arrow[r, "\mathrm{e}^{-h} d\, \mathrm{e}^{h}"] & \Omega^n(M) \arrow[r, "\mathrm{e}^{-h} d\, \mathrm{e}^{h}"] & 0.
\end{tikzcd}$$

Therefore,

$$\mathcal{H}^p_{(0)} \cong H^p(Q) \cong H^p(Q_{h=0}) = H^p_{DR}(M), \tag{10.249}$$

In particular, *the dimension of the supersymmetric ground states is independent of the choice of the function h.*

10.4.3. SQM on a Kähler Manifold. We study here the supersymmetric sigma model in the case where the target space M is a Kähler manifold. The focus will be on the extended supersymmetry and two kinds of fermion number operators. The readers do not have to check all these formulae in detail. They follow from the formulae in the non-linear sigma model in $(1+1)$ dimensions, which will be derived systematically in Ch. 12 and Ch. 13.

We recall that a Kähler manifold is a complex manifold with a Hermitian metric g such that the two-form ω defined by $\omega(X, Y) = g(JX, Y)$ is closed. In terms of the local complex coordinates $(z^i) = (z^1, \dots, z^n)$, where n is the complex dimension of M, the Kähler form is written in terms of the metric tensor $g_{i\bar{\jmath}}$ as

$$\omega = i g_{i\bar{\jmath}} dz^i \wedge d\bar{z}^{\bar{\jmath}}. \tag{10.250}$$

The Kähler condition reads as $\partial_i g_{j\bar{k}} = \partial_j g_{i\bar{k}}$ and the Christoffel symbol can be written as $\Gamma^i_{jk} = g^{i\bar{l}} \partial_j g_{k\bar{l}}$.

As before, the sigma model is described by scalar fields ϕ^i and $\overline{\phi^{\bar{\imath}}}$ representing the map $\phi : \mathcal{T} \to M$, and fermions $\psi^i, \psi^{\bar{\imath}}, \overline{\psi}^i, \overline{\psi}^{\bar{\imath}}$ that represent the sections $\psi, \overline{\psi} \in \Gamma(\mathcal{T}, \phi^* TM \otimes \mathbb{C})$. The Lagrangian is as shown in Eq. (10.198). In terms of the complex variables it is expressed as

$$L = g_{i\bar{\jmath}} \dot{\phi}^i \dot{\overline{\phi}}^{\bar{\jmath}} + i g_{i\bar{\jmath}} \overline{\psi}^{\bar{\jmath}} D_t \psi^i + i g_{i\bar{\jmath}} \overline{\psi}^i D_t \psi^{\bar{\jmath}} + R_{i\bar{\jmath}k\bar{l}} \overline{\psi}^i \psi^k \psi^{\bar{\jmath}} \overline{\psi}^{\bar{l}}. \tag{10.251}$$

This system has an extended supersymmetry as in the theory of complex variables with the potential determined by $h = -\mathrm{Re}\, W$. The supersymmetry variation has two complex parameters ϵ_+ and ϵ_- and is given by

$$\begin{aligned}
\delta\phi^i &= \epsilon_+ \overline{\psi}^i - \epsilon_- \psi^i, & \delta\overline{\phi}^{\bar{\imath}} &= -\overline{\epsilon}_+ \psi^{\bar{\imath}} + \overline{\epsilon}_- \overline{\psi}^{\bar{\imath}}, \\
\delta\psi^i &= i\overline{\epsilon}_- \dot{\phi}^i - \epsilon_+ \Gamma^i_{jk} \overline{\psi}^j \psi^k, & \delta\overline{\psi}^{\bar{\imath}} &= -i\epsilon_- \dot{\overline{\phi}}^{\bar{\imath}} - \overline{\epsilon}_+ \Gamma^{\bar{\imath}}_{\bar{\jmath}\bar{k}} \overline{\psi}^{\bar{\jmath}} \psi^{\bar{k}}, \\
\delta\overline{\psi}^i &= -i\overline{\epsilon}_+ \dot{\phi}^i - \epsilon_- \Gamma^i_{jk} \overline{\psi}^j \psi^k, & \delta\psi^{\bar{\imath}} &= i\epsilon_+ \dot{\overline{\phi}}^{\bar{\imath}} - \overline{\epsilon}_- \Gamma^{\bar{\imath}}_{\bar{\jmath}\bar{k}} \overline{\psi}^{\bar{\jmath}} \psi^{\bar{k}}.
\end{aligned} \tag{10.252}$$

By the Noether procedure, we find four supercharges $Q_\pm$ and $\overline{Q}_\pm$,

$$Q_+ = g_{i\bar{\jmath}} \psi^i \dot{\overline{\phi}}^{\bar{\jmath}}, \; Q_- = g_{i\bar{\jmath}} \overline{\psi}^i \dot{\overline{\phi}}^{\bar{\jmath}}, \tag{10.253}$$

$$\overline{Q}_+ = g_{i\bar{\jmath}} \overline{\psi}^{\bar{\jmath}} \dot{\phi}^i, \; \overline{Q}_- = g_{i\bar{\jmath}} \psi^{\bar{\jmath}} \dot{\phi}^i. \tag{10.254}$$

The ordinary supercharges Q and $\overline{Q}$ are simply the linear combinations $Q = i(Q_- + \overline{Q}_+)$ and $\overline{Q} = -i(\overline{Q}_- + Q_+)$. The Lagrangian shown in Eq. (10.251) is invariant under two kinds of phase rotation of fermions:

$$\overline{\psi}^i \mapsto \mathrm{e}^{i(-\alpha+\beta)} \overline{\psi}^i, \; \overline{\psi}^{\bar{\imath}} \mapsto \mathrm{e}^{i(\alpha+\beta)} \overline{\psi}^{\bar{\imath}}, \tag{10.255}$$

$$\psi^i \mapsto \mathrm{e}^{i(\alpha-\beta)} \psi^i, \; \psi^{\bar{\imath}} \mapsto \mathrm{e}^{i(-\alpha-\beta)} \psi^{\bar{\imath}}. \tag{10.256}$$

We call the α and β rotations *vector* and *axial* rotations, respectively. (The names have a (1+1)-dimensional origin.) The corresponding Noether charges are given by

$$F_V = g_{i\bar{\jmath}} (\overline{\psi}^{\bar{\jmath}} \psi^i - \overline{\psi}^i \psi^{\bar{\jmath}}), \; F_A = g_{i\bar{\jmath}} (\overline{\psi}^{\bar{\jmath}} \psi^i + \overline{\psi}^i \psi^{\bar{\jmath}}). \tag{10.257}$$

The fermion number F for a general Riemannian manifold equals F_A, and F_V is the new one present only if M is a complex manifold. In fact, in terms of real coordinates they can be written as $F_V = -i g_{IK} \overline{\psi}^I J^K_L \psi^L$ and $F_A = g_{IK} \overline{\psi}^I \psi^K$, where J^K_L is the matrix for the complex structure.

The canonical commutation relations are expressed in terms of the complex coordinates as

$$[\phi^i, p_j] = i\delta^i{}_j, \; [\overline{\phi}^{\bar{\imath}}, p_{\bar{\jmath}}] = i\delta^{\bar{\imath}}{}_{\bar{\jmath}}, \tag{10.258}$$

$$\{\psi^i, \overline{\psi}^{\bar{\jmath}}\} = g^{\bar{\jmath}i}, \; \{\psi^{\bar{\imath}}, \overline{\psi}^j\} = g^{\bar{\imath}j}, \tag{10.259}$$

where $p_i = \partial L/\partial \dot{\phi}^i = g_{i\bar{j}}\dot{\overline{\phi}}^{\bar{j}}$ and $p_{\bar{j}} = \partial L/\partial \dot{\overline{\phi}}^{\bar{j}} = g_{i\bar{j}}\dot{\phi}^i$. All other (anti-)commutators vanish. The supercharges are now operators

$$Q_+ = \psi^i p_i, \; Q_- = \overline{\psi}^i p_i, \tag{10.260}$$

$$\overline{Q}_+ = \overline{\psi}^{\bar{\jmath}} p_{\bar{\jmath}}, \; \overline{Q}_- = \psi^{\bar{\jmath}} p_{\bar{\jmath}}, \tag{10.261}$$

that generate the supersymmetry transformations in Eq. (10.252) via $\widehat{\delta} = i\epsilon_+ Q_- - i\epsilon_- Q_+ - i\overline{\epsilon}_+\overline{Q}_- + i\overline{\epsilon}_-\overline{Q}_+$. Under the operator ordering for the Hamiltonian H chosen before, these supercharges obey the anti-commutation relations

$$\{Q_\alpha, \overline{Q}_\beta\} = \delta_{\alpha\beta} H, \tag{10.262}$$

$$\{Q_\alpha, Q_\beta\} = \{\overline{Q}_\alpha, \overline{Q}_\beta\} = 0. \tag{10.263}$$

The commutators with the vector and axial fermion numbers F_V and F_A are

$$[F_V, Q_\pm] = -Q_\pm, \; [F_V, \overline{Q}_\pm] = -\overline{Q}_\pm, \tag{10.264}$$

$$[F_A, Q_\pm] = \mp Q_\pm, \; [F_A, \overline{Q}_\pm] = \pm\overline{Q}_\pm. \tag{10.265}$$

As a consequence F_V and F_A are conserved charges:

$$[H, F_V] = [H, F_A] = 0. \tag{10.266}$$

The two fermion numbers commute with each other,

$$[F_V, F_A] = 0, \tag{10.267}$$

and therefore the Hilbert space of states $\mathcal{H} = \Omega(M) \otimes \mathbb{C}$ decomposes with respect to the quantum numbers of F_V and F_A. We note here that

$$\overline{\psi}^i \leftrightarrow dz^i \wedge, \; \overline{\psi}^{\bar{\imath}} \leftrightarrow d\overline{z}^{\bar{\imath}} \wedge, \tag{10.268}$$

$$\psi^i \leftrightarrow g^{\bar{\jmath} i} i_{\partial/\partial \overline{z}^{\bar{\jmath}}}, \; \psi^{\bar{\imath}} \leftrightarrow g^{\bar{\imath} j} i_{\partial/\partial z^j}. \tag{10.269}$$

Thus, by looking at the action of F_V and F_A, as shown in Eqs. (10.255)–(10.256), we see that the state corresponding to a (p, q)-form

$$\eta = \eta_{i_1 \dots i_p \bar{\jmath}_1 \dots \bar{\jmath}_q} dz^{i_1} \wedge \cdots \wedge dz^{i_p} \wedge d\overline{z}^{\bar{\jmath}_1} \wedge \cdots \wedge d\overline{z}^{\bar{\jmath}_q}, \tag{10.270}$$

has F_V charge $-p + q$ and F_A charge $p + q$. Thus, the decomposition with respect to the Hodge degree,

$$\Omega(M) \otimes \mathbb{C} = \bigoplus_{p,q=1}^{n} \Omega^{p,q}(M), \tag{10.271}$$

diagonalizes F_V and F_A:

$$\left. \begin{aligned} F_V &= -p + q \\ F_A &= p + q \end{aligned} \right\} \text{ on } \Omega^{p,q}(M). \tag{10.272}$$

We note that

$$Q_- = \overline{\psi}^i p_i \leftrightarrow dz^i \left(-i\frac{\partial}{\partial z^i}\right) = -i\partial, \tag{10.273}$$

$$\overline{Q}_+ = \overline{\psi}^{\bar{\imath}} p_{\bar{\imath}} \leftrightarrow d\overline{z}^{\bar{\imath}} \left(-i\frac{\partial}{\partial \overline{z}^{\bar{\imath}}}\right) = -i\overline{\partial}, \tag{10.274}$$

$$\overline{Q}_- = Q_-^\dagger \leftrightarrow i\partial^\dagger, \tag{10.275}$$

$$Q_+ = \overline{Q}_+^\dagger \leftrightarrow i\overline{\partial}^\dagger, \tag{10.276}$$

where ∂ and $\overline{\partial}$ are the Dolbeault operators

$$\begin{tikzcd} & \Omega^{p+1,q}(M) \\ \Omega^{p,q}(M) \arrow[ur, "\partial"] \arrow[dr, "\overline{\partial}"'] & \\ & \Omega^{p,q+1}(M). \end{tikzcd}$$

By the commutation relations given by Eq. (10.262) and $\{Q, \overline{Q}\} = 2H$, we find

$$\begin{aligned} H &= \{\overline{Q}_+, \overline{Q}_+^\dagger\} = \overline{\partial}\overline{\partial}^\dagger + \overline{\partial}^\dagger\overline{\partial} =: \Delta_{\overline{\partial}} \\ &= \{Q_-, Q_-^\dagger\} = \partial\partial^\dagger + \partial^\dagger\partial =: \Delta_\partial \\ &= \frac{1}{2}\{Q, Q^\dagger\} = \frac{1}{2}(dd^\dagger + d^\dagger d) = \frac{1}{2}\Delta. \end{aligned} \tag{10.277}$$

That the Laplacians associated with $\overline{\partial}$, ∂ and d agree with each other (up to a factor of 2) is a well-known fact in Kähler geometry. In any case, the space of supersymmetric ground states is the space of harmonic forms

$$\mathcal{H}_{(0)} = \mathcal{H}(M, g) = \bigoplus_{p,q=1}^{n} \mathcal{H}^{p,q}(M, g), \tag{10.278}$$

where $\mathcal{H}^{p,q}(M, g)$ is the space of harmonic (p, q)-forms corresponding to the ground states with vector and axial charges $q_V = -p + q$ and $q_A = p + q$. Note that the ground states of F-charge r correspond to

$$\mathcal{H}^r(M, g) = \bigoplus_{p+q=r} \mathcal{H}^{p,q}(M, g). \tag{10.279}$$

The commutation relations

$$\{\overline{Q}_+, \overline{Q}_+^\dagger\} = H, \tag{10.280}$$

$$\overline{Q}_+^2 = 0, \tag{10.281}$$

show that $\overline{Q}_+$ by itself defines a supersymmetric quantum mechanics. In particular, the space of supersymmetric ground states is identified as the $\overline{Q}_+$-cohomology group. Since

$$[\tfrac{1}{2}(F_V + F_A), \overline{Q}_+] = \overline{Q}_+, \ [\tfrac{1}{2}(-F_V + F_A), \overline{Q}_+] = 0, \tag{10.282}$$

the $\overline{Q}_+$-complex is $(\mathbb{Z} \oplus \mathbb{Z})$-graded and is given by the Dolbeault complex

$$0 \to \Omega^{p,0}(M) \xrightarrow{\overline{\partial}} \Omega^{p,1}(M) \xrightarrow{\overline{\partial}} \cdots \xrightarrow{\overline{\partial}} \Omega^{p,n}(M) \to 0, \tag{10.283}$$

where p is the charge for $\frac{1}{2}(-F_V + F_A)$. Thus, the space of supersymmetric ground states with charge $(q_V, q_A) = (-p+q, p+q)$ is isomorphic to the Dolbeault cohomology group

$$\mathcal{H}^{p,q}(M,g) \cong H^{p,q}_{\overline{\partial}}(M). \tag{10.284}$$

This is also a well-known fact in Kähler geometry or Hodge theory. Similar comments apply for Q_--cohomology, and we have the isomorphism $\mathcal{H}^{p,q}(M,g) \cong H^{p,q}_{\partial}(M)$.

10.4.4. Landau–Ginzburg Model. Suppose there is a non-trivial holomorphic function W on our Kähler manifold M (which is possible only if M is non-compact). To the Lagrangian as shown in Eq. (10.251), we can consider adding the term

$$\Delta L = -\frac{1}{4} g^{i\overline{j}} \partial_i W \partial_{\overline{j}} \overline{W} - \frac{1}{2} D_i \partial_i W \psi^i \overline{\psi}^j - \frac{1}{2} D_{\overline{i}} \partial_{\overline{j}} \overline{W} \psi^{\overline{i}} \overline{\psi^{\overline{j}}}. \tag{10.285}$$

W will be called the superpotential. This system also has extended supersymmetry generated by two complex parameters, where the transformation law is modified by

$$\begin{aligned} \Delta\delta\psi^i &= -\tfrac{1}{2}\epsilon_+ g^{i\overline{j}} \partial_{\overline{j}} \overline{W}, & \Delta\delta\psi^{\overline{i}} &= -\tfrac{1}{2}\overline{\epsilon}_+ g^{\overline{i}j} \partial_j W, \\ \Delta\delta\overline{\psi}^i &= -\tfrac{1}{2}\epsilon_- g^{i\overline{j}} \partial_{\overline{j}} \overline{W}. & \Delta\delta\overline{\psi}^{\overline{i}} &= -\tfrac{1}{2}\overline{\epsilon}_- g^{\overline{i}j} \partial_j W, \end{aligned} \tag{10.286}$$

The expression of the supercharges is modified accordingly. As for the fermion number symmetry, F_V is broken by the added term while F_A remains a symmetry of the system. The quantum Hilbert space is still given

by

$$\mathcal{H} = \bigoplus_{l=1}^{2n} \Omega^l(M), \tag{10.287}$$

on which the supercharges act as

$$\begin{aligned} iQ_- &= \partial + \frac{i}{2}\partial W \wedge, & i\overline{Q}_- &= -(\partial + \frac{i}{2}\partial W \wedge)^\dagger, \\ i\overline{Q}_+ &= \overline{\partial} - \frac{i}{2}\partial W \wedge, & i\mathbb{Q}_+ &= -(\overline{\partial} - \frac{i}{2}\partial W \wedge)^\dagger \end{aligned} \tag{10.288}$$

The space of supersymmetric ground states is isomorphic to, say, the $\overline{Q}_+$-cohomology group.

Although F_V is not a symmetry, one can consider it as an operator acting on the Hilbert space (as $F_V = -p + q$ on (p,q)-forms). As we discussed earlier, conjugation by the operator $\sqrt{\lambda}^{F_V}$ has an effect of rescaling $W \to \lambda W$ in the expression of $\overline{Q}_+$. Since the cohomology is not affected by the conjugation, the spectrum of supersymmetric ground states is invariant under the rescaling of the superpotential W.

Suppose the superpotential W has only non-degenerate critical points $p_1, \ldots, p_N$. In the large λ limit, the ground state wave-functions will be localized at the critical points. Then the behavior of the manifold M away from the critical points is irrelevant, and one can use the earlier analysis done for $M = \mathbb{C}^n$. For each critical point p_i, we obtain the approximate supersymmetric ground state Ψ_i. These states all have the fermion number $F = F_A = n$, and therefore there is no room for tunneling. The exact quantum ground states are in one-to-one correspondence with the critical points of W. In particular, we have shown that the $\overline{Q}_+$-cohomology vanishes except in the middle dimension,

$$H^\ell_{\overline{Q}_+}(M) = \begin{cases} \mathbb{C}^{\#(\text{crit. pts.})} & \ell = n, \\ 0 & \ell \neq n. \end{cases} \tag{10.289}$$

We will explicitly construct the cohomology classes below.

10.4.5. Kähler Manifold with a Holomorphic Vector Field. Let us consider a Kähler manifold M of dimension n that has a holomorphic vector field $v = v^i(z)\frac{\partial}{\partial z^i}$. We will consider a quantum mechanical system whose Hilbert space of states is

$$\mathcal{H} = \bigoplus_{p,q=0}^{n} \Omega^{0,p}(M, \wedge^q T_M), \tag{10.290}$$

with an operator Q_s given by

$$Q_s = \overline{\partial} + s\, v \wedge . \tag{10.291}$$

Here s is a real parameter and $v\wedge$ is the exterior multiplication by v. The operator Q_s is nilpotent, $Q_s^2 = 0$, and if we define the Hamiltonian by $H_s = \frac{1}{2}\{Q_s, Q_s^\dagger\}$, we obtain a supersymmetric quantum mechanics with supercharge Q_s. The space of supersymmetric ground states is of course the Q_s cohomology group.

The first thing to notice is that the dimension of the Q_s-cohomology group is independent of the value of s as long as it is nonzero. To see this, let D be the operator acting as $D = q$ on the subspace $\Omega^{0,*}(M, \wedge^q T_M)$. Then we find $e^{tD} Q_s e^{-tD} = Q_{e^t s}$, and therefore the $Q_{e^t s}$-cohomology group is isomorphic to the Q_s cohomology group. Now let us take the limit $s \to \infty$. Then the ground state wave-function is localized at the zero of v, which we assume to be a smooth submanifold M_0 of M of dimension m. In the strict $s = \infty$ limit, the system reduces to the quantum mechanics on M_0 with supercharge $\overline{\partial}$. The supsersymmetric ground states of the limiting theory are composed of the cohomology classes of the Dolbeault complex on M_0 with values in $\wedge^\bullet T_{M_0}$. Since a zero energy state of the full theory remains as a zero energy state in this limit, we have the inequality

$$\dim \mathcal{H}_{(0)} \leq \dim H^{0,\bullet}_{\overline{\partial}}(M_0, \wedge^\bullet T_{M_0}). \tag{10.292}$$

We will now show that, under certain circumstances, the opposite inequality also holds. Let $N_{M_0/M}$ be the normal bundle $T_M|_{M_0}/T_{M_0}$ of M_0 in M. The assumptions are

(i) a neighborhood of M_0 in M is exactly isomorphic, to a complex manifold, as a neighborhood of the zero section of $N_{M_0/M}$,

(ii) under that isomorphism, v is tangent to the fibres,

(iii) the normal bundle has a trivial determinant bundle, or $c_1(N_{M_0/M}) = 0$.

We will also choose the metric on the neighborhood so that it is induced from a metric on M_0 and a fibre metric. Assumptions (i) and (ii) hold if v generates a $U(1)$ action on M with a simple zero at M_0 so that it can be written as $v = \sum_i a_i z^i \partial/\partial z^i$, where z^i are normal coordinates. Let Ω be the holomorphic section of $N_{M_0/M}$ that exists if (iii) is obeyed. Let us choose a smooth function $f(r)$ such that $f \equiv 1$ for $r < \epsilon$ but $f \equiv 0$ for $r > 2\epsilon$, where ϵ is such that the neighborhood of M_0 in question is in the region

$\|v\|^2 = g_{i\bar{j}} v^i \overline{v}^{\bar{j}} < 3\epsilon$. Let us put

$$\Psi = \sum_{p=0}^{n-m} \pm f^{(p)}(\|v\|^2)\Omega^{i_1 \cdots i_{n-m}} \overline{\partial}(g_{i_1\bar{j}_1}\overline{v}^{\bar{j}_1}) \cdots \overline{\partial}(g_{i_p\bar{j}_p}\overline{v}^{\bar{j}_p}) \frac{\partial}{\partial z^{i_{p+1}}} \wedge \cdots \wedge \frac{\partial}{\partial z^{i_{n-m}}}, \tag{10.293}$$

where $f^{(p)}(r)$ is the pth derivative of $f(r)$. Then under a suitable choice of $\pm$ signs in the above formula, one can show that $(\overline{\partial} + v\wedge)\Psi = 0$.

Let us consider the map

$$\alpha \in \Omega^{0,\bullet}(M_0, \wedge^\bullet T_{M_0}) \mapsto \Psi \wedge \pi^*\alpha \in \Omega^{0,\bullet}(M, \wedge^\bullet T_M), \tag{10.294}$$

where $\pi : N_{M_0/M} \to M_0$ is the projection map. It is easy to see that $\overline{\partial}\alpha = 0$ means $(\overline{\partial} + v\wedge)(\Psi \wedge \pi^*\alpha) = 0$ and also $\Psi \wedge \pi^*\overline{\partial}\beta = \pm(\overline{\partial} + v\wedge)(\Psi \wedge \pi^*\beta)$. Namely, $\overline{\partial}$-closed/exact forms are mapped to $Q_1 = (\overline{\partial} + v\wedge)$-closed/exact forms. Thus, the above defines a map from the Dolbeault cohomology group of M_0 to the Q_1-cohomology group. Furthermore, contracting by the inverse of Ω at M_0, we recover α:

$$\Omega^{-1} \cdot (\Psi \wedge \pi^*\alpha)|_{M_0} = \alpha. \tag{10.295}$$

This shows that the map is an isomorphism, and therefore

$$\mathcal{H}_{(0)} \cong H^{0,\bullet}_{\overline{\partial}}(M_0, \wedge^\bullet T_{M_0}). \tag{10.296}$$

Landau–Ginzburg Model, Revisited. Let us compare the expressions for Q_s in Eq. (10.291) and $\overline{Q}_+$ in Eq. (10.288). They are identical if we replace T_M by T_M^* and v by ∂W. One can therefore apply the above argument to the Landau–Ginzburg model as well. Let M_0 be the subset of M consisting of the critical points of the superpotential W. Then in general we have a bound

$$\dim H^\bullet_{\overline{Q}_+}(M) \leq \dim H^\bullet(M_0). \tag{10.297}$$

Suppose, as before, that W has only non-degenerate critical points, so that M_0 is a set of points, $M_0 = \{p_1, \ldots, p_N\}$. Then the assumptions analogous to (i)(ii)(iii) hold, and one can construct a one-to-one map $H^\bullet(M_0) \to H^\bullet_{\overline{Q}_+}(M)$. Namely, one can explicitly construct the $\overline{Q}_+$-cohomology classes. The result is

$$\Psi_i = \sum_{p=0}^{n} \pm f^{(p)}(\|\partial W\|^2)\epsilon_{i_1 \ldots i_n} \overline{\partial}(g^{i_1\bar{j}_1}\partial_{\bar{j}_1}\overline{W}) \cdots \overline{\partial}(g^{i_p\bar{j}_p}\partial_{\bar{j}_p}\overline{W}) dz^{j_{p+1}} \cdots dz^{i_n}, \tag{10.298}$$

where z_i is a coordinate system near p_i.

10.5. Instantons

Consider the supersymmetric quantum mechanics on a Riemannian manifold (M, g) deformed by a function h as was introduced in Sec. 10.4.2. We consider the case where h is a *Morse function*, namely, all the critical points are isolated and non-degenerate. We denote the critical points by

$$x_1, \dots, x_N. \tag{10.299}$$

Consider rescaling the function h as

$$h \longrightarrow \lambda h, \text{ with } \lambda \gg 1. \tag{10.300}$$

This does not change the number of supersymmetric ground states, as discussed before. The Hamiltonian of the system is

$$H_\lambda = \frac{1}{2}\Delta + \frac{1}{2}\lambda^2 g^{IJ}\partial_I h \partial_J h + \frac{1}{2}\lambda D_I \partial_J h[\overline{\psi}^I, \psi^J]. \tag{10.301}$$

At large λ, low-energy states are localized near the critical points of h, where the potential term $(\lambda^2/2)g^{IJ}\partial_I h \partial_J h$ vanishes. As discussed in the single-variable case, we can consider perturbation theory around each critical point x_i. We can choose coordinates x^I around the critical point x_i such that

$$h = h(x_i) + \sum_{I=1}^{n} c_I (x^I)^2 + O((x^I)^3). \tag{10.302}$$

The coefficients c_I are the eigenvalues of the Hessian of h at the critical point x_i, $\partial_I \partial_J h(x_i)$. The higher-order terms $O((x^I)^3)$ in Eq. (10.302) are subleading in the perturbation theory. The deviation of the metric (from the flat one) around the critical point can also be considered as subleading in the perturbation theory, and one can replace g_{IJ} by $g_{IJ}(x_i)$. For simplicity, we choose it as $g_{IJ}(x_i) = \delta_{IJ}$. (This can be done either by deforming the function h or the metric g_{IJ}; we know that neither affects the Q-cohomology and hence the number of supersymmetric ground states.) Thus the leading order terms of the Hamiltonian in the perturbation theory at x_i are given by

$$H_0(x_i) = \sum_{I=1}^{n} \left\{ \frac{1}{2}p_I^2 + \frac{1}{2}\lambda^2 c_I^2 (x^I)^2 + \frac{1}{2}\lambda c_I [\overline{\psi}^I, \psi^I] \right\}. \tag{10.303}$$

Thus, we find the supersymmetric ground state at leading order in perturbation theory at x_i,

$$\Psi_i^{(0)} = \mathrm{e}^{-\lambda \sum_{I=1}^{n} |c_I|(x^I)^2} \prod_{J:\ c_J<0} \overline{\psi}^J |0\rangle. \tag{10.304}$$

The number of $\overline{\psi}^J$'s that multiply $|0\rangle$ is the Morse index of h at x_i,

$$\mu_i = \#\text{ of negative eigenvalues of the Hessian of } h \text{ at } x_i. \tag{10.305}$$

This shows that the wave-function $\Psi_i^{(0)}$ is a μ_i-form. As in the single-variable case, one can find the modification Ψ_i of $\Psi_i^{(0)}$ so that it remains the zero energy state to all orders in perturbation theory. Since the perturbation theory also preserves the fermion number symmetry F we see that Ψ_i is still a μ_i-form,

$$\Psi_i \in \Omega^{\mu_i}(M) \otimes \mathbb{C}. \tag{10.306}$$

As $\Psi_i^{(0)}$, Ψ_i is supported around and peaked at x_i in the large λ limit (see Fig. 5). Note that Ψ_i is an exact supersymmetric ground state in the perturbation theory.

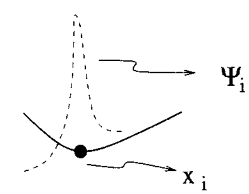


FIGURE 5

bation theory. Other states have diverging energy if we consider the $\lambda \longrightarrow \infty$ limit. Since the number of supersymmetric ground states is independent of λ we see that the number of supersymmetric ground states does not exceed the number of these perturbative zero energy states, namely the number of critical points. However, in general, the perturbative ground states are only approximate ground states in the full theory. This can be seen in the example described below.

EXAMPLE 10.5.1 ($M = S^2$). *Consider the case when the target space is S^2 and h is the height function as in Fig. 6. We find two critical points, one with Morse index $p = 0$ and the other with $p = 2$. Thus, there are two perturbative zero energy states; one is a zero-form and the other one is a two-form.*

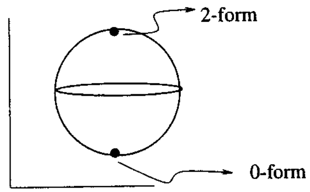


FIGURE 6.

However, we can also consider the deformed sphere such that the height function has many more critical points, as shown in Fig. 7.

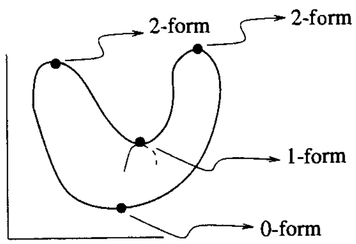


FIGURE 7.

This time we find four critical points, one with Morse index $p = 0$, one with $p = 1$, and two with $p = 2$. Thus, there are four perturbative zero energy states: one zero-form, one one-form and two two-forms. So, there is a discrepancy in the number of perturbative zero energy states between the two theories corresponding to the two different choices of the function h. However, as we have seen, the number of zero energy states of the full theory should not depend on the choice of h. Thus, in either one of the two theories or both, the perturbative ground states are not really the actual ground states of the full theory. For the first choice of h, the two perturbative ground states are both bosonic, $(-1)^F = 1$, and it is impossible for both to become nonzero energy states. Thus, these two perturbative ground states are really the supersymmetric ground states of the full theory. Therefore we see that the number of supersymmetric ground states in the full theory is two. In particular, not all the four perturbative ground states in the second example are exact, and only two linear combinations of them are the actual zero energy states.

Let us come back to the general story. As we have seen explicitly in the above example, it is not necessarily the case that each Ψ_i determines a supersymmetric ground state in the full theory. In other words, it is not necessarily the case that

$$Q\Psi_i = 0 \text{ for all } i. \tag{10.307}$$

Although this holds to all orders in perturbation theory, in general this should somehow be modified in the full theory. Namely, we expect to have an expansion

$$Q\Psi_i = \sum_{j=1}^{N} \Psi_j \langle \Psi_j, Q\Psi_i\rangle + \cdots, \tag{10.308}$$

where $+\cdots$ involves nonzero energy states in perturbation theory. Since these latter states have large energies $\simeq \lambda$, the terms $+\cdots$ are smaller compared to the first N terms by powers of λ^{-1}, and will be omitted henceforth. Thus, what we want to compute is

$$\langle \Psi_j, Q\Psi_i\rangle = \int_M \overline{\Psi}_j \wedge *(d + dh\wedge)\Psi_i. \tag{10.309}$$

Since Ψ_j is a μ_j-form and $Q\Psi_i = (d + dh\wedge)\Psi_i$ is a $(\mu_i + 1)$-form, the above matrix element can be nonzero only if

$$\mu_j = \mu_i + 1. \tag{10.310}$$

We will compute this matrix element using the path-integral formalism.

10.5.1. The Path-Integral Representation. We thus wish to compute non-perturbative corrections to the matrix elements of Q between the perturbative ground states, in the limit of large λ. In this limit the ground state wave-functions are sharply peaked near the critical points of h. In other words, to leading order, the Morse function h, viewed as an operator, acting on the ground state, gives the value of h at the corresponding critical point. This implies that the matrix element of Q between perturbative ground states, to leading order in $1/\lambda$, is equal to

(10.311)
$$\langle \Psi_j, Q\Psi_i\rangle = \frac{1}{h(x_i) - h(x_j) + O(1/\lambda)} \lim_{T\to\infty} \langle \Psi_j,\, e^{-TH}[Q,h]\, e^{-TH}\Psi_i\rangle,$$

where for Ψ_i we can take any function that has non-vanishing overlap with the ith critical point and vanishes at all the others. The operator e^{-TH} as

$T \to \infty$ projects that state to the perturbative ground state corresponding to the ith critical point. The commutator $[Q, h]$ can be expressed as

$$[Q, h] = \partial_I h [Q, \psi^I] = \partial_I h \overline{\psi}^I. \tag{10.312}$$

Thus we have a path-integral expression of the matrix element

(10.313)

$$\lim_{T\to\infty} \langle \Psi_j, \mathrm{e}^{-TH}[Q, h]\, \mathrm{e}^{-TH}\Psi_i\rangle = \int_{\phi(-\infty)=x_i,\ \phi(+\infty)=x_j} \mathcal{D}\phi\mathcal{D}\psi\mathcal{D}\overline{\psi}\ \mathrm{e}^{-S_E}\, \overline{\psi}^I \partial_I h|_{\tau=0}.$$

Here the integration region is the space of fields satisfying the boundary condition that $\phi(-\infty) = x_i$ and $\phi(\infty) = x_j$ and that $d\phi^I/d\tau$, ψ^I and $\overline{\psi}^I$ fall off sufficiently fast as $\tau \to \pm\infty$. The Euclidean action is given by

$$\begin{aligned} S_E = \int_{-\infty}^{\infty} d\tau \bigg\{ &\frac{1}{2} g_{IJ} \frac{d\phi^I}{d\tau}\frac{d\phi^J}{d\tau} + \frac{\lambda^2}{2} g^{IJ}\partial_I h \partial_J h + g_{IJ}\overline{\psi}^I D_\tau \psi^J \\ &+ \lambda D_I \partial_J h \overline{\psi}^I \psi^J + \frac{1}{2} R_{IJKL}\psi^I\overline{\psi}^J\psi^K\overline{\psi}^L \bigg\}. \end{aligned} \tag{10.314}$$

The bosonic part of the action can be written as

$$\begin{aligned} S_{\text{bosonic}} &= \int_{-\infty}^{\infty} d\tau \left(\frac{1}{2}\left|\frac{d\phi^I}{d\tau} \pm \lambda g^{IJ}\partial_J h\right|^2 \mp \lambda \frac{d\phi^I}{d\tau}\partial_I h \right) \\ &= \frac{1}{2}\int_{-\infty}^{\infty} d\tau \left|\frac{d\phi^I}{d\tau} \pm \lambda g^{IJ}\partial_J h\right|^2 \mp \lambda(h(x_j) - h(x_i)). \end{aligned} \tag{10.315}$$

In the above equation we used the boundary condition $\phi(-\infty) = x_i$ and $\phi(+\infty) = x_j$. Thus we see that the configurations that minimize the action are such that

$$\frac{d\phi^I}{d\tau} \pm \lambda g^{IJ}\partial_J h = 0 \text{ if } h(x_j) - h(x_i) \lessgtr 0. \tag{10.316}$$

Such a configuration is called an *instanton*. The name comes from the fact that the transition from x_i to x_j happens at some "instant" (though not really) within the infinite interval of (Euclidean) time $-\infty < \tau < \infty$. We are interested in how many instantons there are. Clearly, an instanton $\phi(\tau)$ is deformed to another instanton by shifting τ: $\phi'(\tau) = \phi(\tau + \delta\tau)$. To see whether there are more deformations, we take the first-order variation of Eq. (10.316). It is straightforward to see that it is given by

$$\mathcal{D}_\pm \delta\phi^I := D_\tau \delta\phi^I \pm \lambda g^{IJ} D_J \partial_K h \delta\phi^K = 0. \tag{10.317}$$

Thus, the number of deformations (including the shift in τ) is given by the dimension of the kernel of the operator $\mathcal{D}_\pm$. We note that the fermion bilinear term in the action in Eq. (10.314) is given by

$$S_{\overline{\psi}\psi} = \int_{-\infty}^{\infty} d\tau\, g_{IJ}\overline{\psi}^I \mathcal{D}_+ \psi^J = -\int_{-\infty}^{\infty} d\tau\, g_{IJ}\mathcal{D}_-\overline{\psi}^I\, \psi^J. \tag{10.318}$$

For the path-integral in Eq. (10.313) to be non-vanishing, since there is a single insertion of $\overline{\psi}$, the number of $\overline{\psi}$ zero modes must be larger than the number of ψ zero modes by 1. Namely, the path-integral is non-vanishing only if

$$\operatorname{Ind}\mathcal{D}_- = -\operatorname{Ind}\mathcal{D}_+ = \dim\operatorname{Ker}\mathcal{D}_- - \dim\operatorname{Ker}\mathcal{D}_+ = 1. \tag{10.319}$$

Localization. In the semi-classical limit the path-integral receives dominant contributions from the configurations where the action is minimized. This is the standard reason to look for instantons (even in non-supersymmetric theories) but in general an instanton merely provides the starting point of the semi-classical approximation. In the supersymmetric quantum mechanics, there is a more fundamental reason to consider instantons – the localization principle. The path-integral picks up contribution *only* around certain instantons and the quadratic approximation at the instantons provides an exact result. In particular, one can see that the path-integral chooses a sign in Eq. (10.316) which was not specified in the previous argument.

The point is that the action S_E and the boundary conditions are invariant under the Euclidean supersymmetry

$$\delta\phi^I = \epsilon\overline{\psi}^I - \overline{\epsilon}\psi^I, \tag{10.320}$$

$$\delta\psi^I = \epsilon\left(-\frac{d\phi^I}{d\tau} + \lambda g^{IJ}\partial_J h - \Gamma^I_{JK}\overline{\psi}^J\psi^K\right), \tag{10.321}$$

$$\delta\overline{\psi}^I = \overline{\epsilon}\left(\frac{d\phi^I}{d\tau} + \lambda g^{IJ}\partial_J h - \Gamma^I_{JK}\overline{\psi}^J\psi^K\right). \tag{10.322}$$

Now, the integrand $[Q, h] = \overline{\psi}^I\partial_I h$ is invariant under the ϵ-supersymmetry (generated by Q):

$$\delta_\epsilon(\overline{\psi}^I\partial_I h) = 0. \tag{10.323}$$

Thus, the path-integral receives a contribution only from δ_ϵ-fixed points. This requires

$$\frac{d\phi^I}{d\tau} = \lambda g^{IJ}\partial_J h. \tag{10.324}$$

This is nothing but one of the instanton equations, Eq. (10.316). Moreover the sign $-$ is chosen and hence the path-integral is non-vanishing only if $h(x_j) > h(x_i)$. Thus, the relevant instanton for the present computation is an ascending gradient flow which starts from x_i and ends on x_j, or a path of steepest ascent from x_i to x_j.

10.5.2. Fermion Zero Modes and Relative Morse Index. As noted above, we are interested in the index of the operator $\mathcal{D}_-$. This index is actually equal to the difference between the Morse index of h at x_j and the one at x_i. Namely,

$$\operatorname{Ind}\mathcal{D}_- = \mu_j - \mu_i. \tag{10.325}$$

Thus, as long as $\mu_j = \mu_i + 1$ (the case we are considering) the condition in Eq. (10.319) for non-vanishing of the path-integral is satisfied.

Eq. (10.325) is actually valid for *any* map $\phi : \mathbb{R} \to M$ such that $\phi(-\infty) = x_i$ and $\phi(\infty) = x_j$. This relation will be important also when we discuss non-linear sigma models in $1+1$ dimensions. It can be proved as follows.

We generalize our definition of the Hessian (which has been defined at the critical points of h as the matrix of the second derivatives) to an arbitrary point x of M. The Hessian H_h at x is defined as the linear map $T_xM \to T_xM$,

$$H_h : v^I \mapsto g^{IJ} D_J \partial_K h\, v^K. \tag{10.326}$$

With respect to an orthonormal frame, H_h is represented as an $n\times n$ symmetric matrix and therefore it can be diagonalized by an orthogonal matrix with real eigenvalues λ_I. Let us consider a trajectory $\phi(\tau)$ such that $\phi(-\infty) = x_i$ and $\phi(\infty) = x_j$. Then the family of matrices $H_h(\phi(\tau))$ defines families of eigenvectors and eigenvalues

$$H_h(\phi(\tau))e_I(\tau) = \lambda_I(\tau)e_I(\tau), \quad -\infty < \tau < \infty. \tag{10.327}$$

The family of eigenvalues $\lambda_I(\tau)$ is called the spectral flow. (We depict in Fig. 8 an example.) We choose $e_I(\tau)$ to define an orthonormal basis of

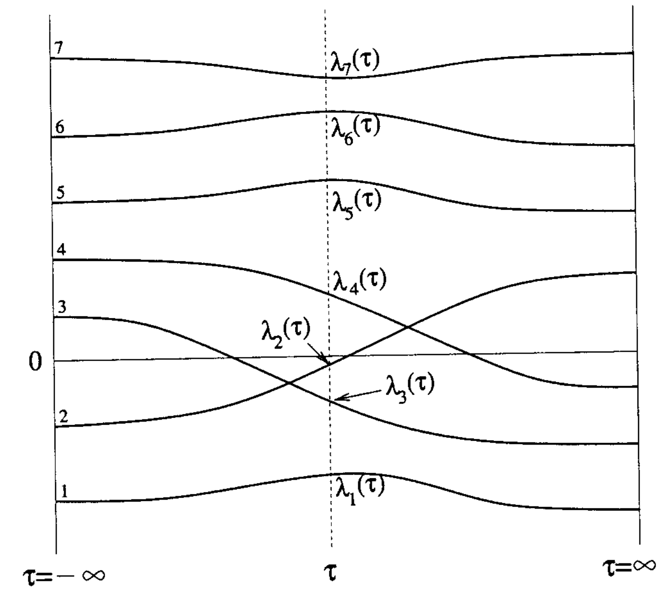


FIGURE 8. The Spectral Flow: An example for the case $\dim M = 7$. The Morse index at x_i is $\mu_i = 2$ whereas that at x_j is $\mu_j = 3$. One of the eigenvalues ($\lambda_2(\tau)$) goes from negative to positive but two of them ($\lambda_3(\tau)$ and $\lambda_4(\tau)$) go from positive to negative

$T_{\phi(\tau)}M$ at each τ. The relative Morse index $\Delta\mu = \mu_j - \mu_i$ counts the net number of eigenvalues that go from positive to negative. Namely,

(10.328)
$$\Delta\mu = \#\{I;\, \lambda_I(-\infty) > 0,\ \lambda_I(\infty) < 0\} - \#\{J;\, \lambda_J(-\infty) < 0,\ \lambda_J(\infty) > 0\}.$$

Let us consider the operators

$$\widetilde{\mathcal{D}}_{\mp} := \frac{d}{d\tau} \mp \begin{pmatrix} \lambda_1(\tau) & & \\ & \ddots & \\ & & \lambda_n(\tau) \end{pmatrix}, \tag{10.329}$$

acting on square-normalizable functions of τ with values in $\mathbb{R}^n$. These are essentially conjugate of each other $\widetilde{\mathcal{D}}_+ = -\widetilde{\mathcal{D}}_-^\dagger$. The equations

$$\widetilde{\mathcal{D}}_{\mp} f_{\mp} = 0 \tag{10.330}$$

are solved by

$$f_{I\mp}(\tau) = e_I \exp\left(\pm \int_0^\tau \lambda_I(\tau')d\tau'\right). \tag{10.331}$$

where e_I is a column vector with 1 at the I-th entry and zero at the others. The solution $f_{I-}(\tau)$ is square-normalizable if and only if $\lambda_I(-\infty) > 0$ and $\lambda_I(\infty) < 0$. Similarly, $f_{J+}(\tau)$ is square-normalizable if and only if $\lambda_J(-\infty) < 0$ and $\lambda_J(\infty) > 0$. Thus, we see that

$$\Delta\mu = \dim \operatorname{Ker} \widetilde{\mathcal{D}}_- - \dim \operatorname{Ker} \widetilde{\mathcal{D}}_+ = \operatorname{Ind} \widetilde{\mathcal{D}}_-. \tag{10.332}$$

The operator $\widetilde{\mathcal{D}}_-$ can be identified as the operator acting on the sections of the bundle ϕ^*TM,

$$\widetilde{\mathcal{D}}_- = \widetilde{D}_\tau - \phi^* H_h, \tag{10.333}$$

where $\widetilde{D}_\tau$ is the connection with respect to which the sections $e_I(\tau)$ are all parallel. On the other hand, we recall that

$$\mathcal{D}_- = D_\tau - \phi^* H_h, \tag{10.334}$$

where D_τ is the operator induced by the Levi–Civita connection of (M, g). Since D_τ and $\widetilde{D}_\tau$ are connections on the same bundle, the index of $\mathcal{D}_-$ and that of $\widetilde{\mathcal{D}}_-$ are the same. Thus, we see that

$$\operatorname{Ind} \mathcal{D}_- = \operatorname{Ind} \widetilde{\mathcal{D}}_- = \Delta\mu, \tag{10.335}$$

which is what we wanted to show.

Genericity Assumption. We make here an assumption that the Morse function h is *generic* in the sense that

$$\operatorname{Ker} \mathcal{D}_+ = 0 \tag{10.336}$$

for any gradient flow (instanton) from x_i to x_j with $\mu_j = \mu_i + 1$. By the relation $\operatorname{Ind} \mathcal{D}_- = \Delta\mu = 1$, each steepest ascent $\gamma^I(\tau)$ has no other deformation than the shift in τ. We denote this one-dimensional modulus by τ_1. Thus, the instanton configuration deformed by τ_1 is

$$\gamma^I_{\tau_1}(\tau) = \gamma^I(\tau + \tau_1). \tag{10.337}$$

τ_1 parametrizes the "position" of the instanton in the infinite interval of Euclidean time.

10.5.3. Evaluation of the Path-Integral. We are finally in a position to evaluate the path-integral. As we have seen above, under the assumption that h is generic, an instanton has a one-dimensional modulus representing the "position" τ_1 of the instanton. By the localization principle, we can exactly evaluate the path-integral in the quadratic approximation.

Changing the variables by $\phi^I = \gamma^I_{\tau_1} + \xi^I$, the action in the quadratic approximation is

$$S_E = \lambda(h(x_j) - h(x_i)) + \int \left(\frac{1}{2}|\mathcal{D}_-\xi|^2 - \mathcal{D}_-\overline{\psi}\psi\right) d\tau, \tag{10.338}$$

where $\mathcal{D}_-$ is the operator acting on the sections of $\gamma^*_{\tau_1}TM$. There is a one-dimensional kernel of $\mathcal{D}_-$ given by

$$\frac{d}{d\tau_1}\gamma^I_{\tau_1} = \frac{d\gamma^I_{\tau_1}}{d\tau}, \tag{10.339}$$

and there is no kernel of $\mathcal{D}_+$. Thus, there is one ξ zero mode, one $\overline{\psi}$ zero mode and no ψ zero mode. The integration variable for the ξ zero mode is τ_1 and we denote by $\overline{\psi}_0$ the variable for the $\overline{\psi}$ zero mode. In particular the variable $\overline{\psi}$ is expanded as

$$\overline{\psi}^I = \frac{d\gamma^I_{\tau_1}}{d\tau}\overline{\psi}_0 + \cdots \tag{10.340}$$

where $+\cdots$ are nonzero mode terms which do not contribute to the path-integral. The nonzero mode path-integral simply gives the ratio of the bosonic and fermionic determinants, which cancels up to sign

$$\frac{\det' \mathcal{D}_-}{\sqrt{\det' \mathcal{D}_-^\dagger \mathcal{D}_-}} = \pm 1 \tag{10.341}$$

The zero mode integrals are

$$\int_{-\infty}^{\infty} d\tau_1 \int d\overline{\psi}_0\, \overline{\psi}_0 \frac{d\gamma^I_{\tau_1}}{d\tau}\partial_I h\Big|_{\tau=0} \tag{10.342}$$

$$= \int_{-\infty}^{\infty} d\tau_1 \frac{d\gamma^I}{d\tau_1}(\tau_1)\partial_I h(\gamma(\tau_1)) = h(x_j) - h(x_i) \tag{10.343}$$

Collecting the two and recovering the classical action factor $\mathrm{e}^{-\lambda(h(x_j)-h(x_i))}$ we obtain the following expression for the contribution of the instanton γ to the path-integral as shown in Eq. (10.313):

$$\pm(h(x_j) - h(x_i))\, \mathrm{e}^{-\lambda(h(x_j)-h(x_i))} \tag{10.344}$$

Summing up the instantons and including the prefactor from Eq. (10.311), we obtain

$$\langle \Psi_j, Q\Psi_i \rangle = \sum_\gamma n_\gamma \, \mathrm{e}^{-\lambda(h(x_j)-h(x_i))} \tag{10.345}$$

where n_γ is $+1$ or -1 depending on the instanton γ.

The sign of n_γ can be determined as follows. The result shown in Eq. (10.345) shows that the integral $\int_M \overline{\Psi_j} \wedge *Q\Psi_i$ receives dominant contributions along the steepest ascents. For each steepest ascent γ, n_γ is 1 or -1 depending on whether the orientation determined by $\overline{\Psi_j} \wedge *Q\Psi_i$ along γ matches with the orientation of M or not. The form Ψ_i defines an orientation of the μ_i-dimensional plane $T^{(-)}_{x_i}M$ of negative eigenmodes of the Hessian of h at x_i. By the spectral flow of the Hessian H_h, this plane can be transported along the steepest ascent and we obtain a sub-bundle T_i^- of γ^*TM with the orientation determined by Ψ_i. Starting with the space of negative eigenmodes of the Hessian at x_j we obtain another sub-bundle T_j^- with the orientation determined by Ψ_j. In the generic situation, only a single eigenvalue goes from positive to negative along the ascent and the eigenmode is the tangent vector v_γ to γ. Then T_i^- is a sub-bundle of T_j^- and the complement is spanned by v_γ. Now, $Q\Psi_i = (d + \lambda dh\wedge)\Psi_i$ defines an orientation of $\mathbb{R}v_\gamma \oplus T_i^-$; it is the one determined by v_γ and Ψ_i. Thus, $n_\gamma = 1$ if this matches with the orientation determined by Ψ_j and $n_\gamma = -1$ otherwise.

10.5.4. Morse–Witten Complex. From what we have seen by the path-integral analysis, we conclude that in the one-instanton approximation

$$Q\Psi_i = \sum_{j:\,\mu_j=\mu_i+1} \Psi_j \sum_\gamma n_\gamma \, \mathrm{e}^{-\lambda(h(x_j)-h(x_i))}. \tag{10.346}$$

The exponential can be eliminated by rescaling the wave-functions Ψ_k. This is the action of the supercharge Q on the perturbative ground states. Since the original supercharge Q is nilpotent, $Q^2 = 0$, it should also be nilpotent when acting on Ψ_i's. Thus, if we define the graded space of perturbative ground states

$$C^\mu := \bigoplus_{\mu_i=\mu} \mathbb{C}\Psi_i, \tag{10.347}$$

we have the cochain complex with the coboundary operator given by the supercharge

$$0 \longrightarrow C^0 \xrightarrow{Q} C^1 \xrightarrow{Q} \cdots \xrightarrow{Q} C^n \xrightarrow{Q} 0. \tag{10.348}$$

The space of supersymmetric ground states is of course the cohomology of this complex. This complex is called Morse–Witten complex.

EXAMPLE 10.5.2 (Example 10.5.1, revisited). *Let us come back to the example of S^2 and examine the case with the second choice of function h as shown in Fig. 7 which is redrawn in Fig. 9. We see that there are two steepest ascents from the critical point A with $\mu = 0$ to the critical point B with $\mu = 1$, γ_1 and γ_2. However, they have opposite orientations and thus*

$$Q\Psi_A = 0. \tag{10.349}$$

From the critical point B, there is one steepest ascent γ_3 to one critical point C with $\mu = 2$ and there is another one γ_4 to another critical point D of $\mu = 2$. If we use the orientation of S^2 for both Ψ_C and Ψ_D, we have

$$Q\Psi_B = \Psi_C - \Psi_D. \tag{10.350}$$

Since there is no critical point of higher Morse index we have

$$Q\Psi_C = Q\Psi_D = 0. \tag{10.351}$$

Thus we obtain

$$\begin{aligned} H^0(Q) &= \mathbb{C}, \\ H^1(Q) &= 0, \\ H^2(Q) &= \mathbb{C}. \end{aligned} \tag{10.352}$$

This is indeed the correct cohomology of S^2.

The Relation $Q^2 = 0$. As mentioned above, the nilpotency relation $Q^2 = 0$ should hold for the supercharge Q. However, it may not be obvious in the realization given by Eq. (10.346). We have seen that it is indeed the case in the above example. Actually, one can show explicitly that $Q^2 = 0$ holds in general, as long as M is a finite-dimensional manifold. What we need to show is that, for x_i and x_j such that $\mu_j = \mu_i + 2$, we have

$$\sum_{k:\mu_k=\mu_i+1} \left(\sum_{\gamma:i\to k} n_\gamma \right) \left(\sum_{\gamma':k\to j} n_{\gamma'} \right) = 0. \tag{10.353}$$

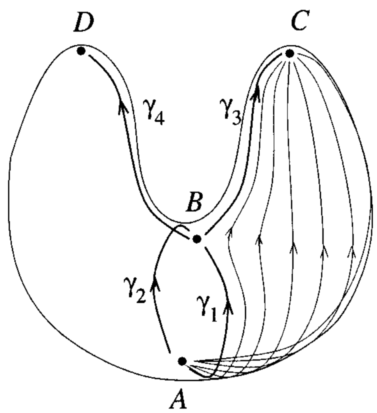


FIGURE 9. The Gradient Flow Lines

Here γ runs over the gradient flow lines from x_i to x_k and γ' runs over the gradient flows from x_k to x_j. To show this, we consider the space of gradient flow lines from a critical point x_a to another critical point x_b:
(10.354)
$$\mathcal{M}(x_a,x_b)=\left\{\phi:\mathbb{R}\to M\;\middle|\;\begin{array}{l}\frac{d\phi^I}{d\tau}=\lambda g^{IJ}\partial_J h,\\ \lim\limits_{\tau\to-\infty}\phi(\tau)=x_a,\ \lim\limits_{\tau\to+\infty}\phi(\tau)=x_b.\end{array}\right\}\Big/\mathbb{R},$$

where $/\mathbb{R}$ means modding out by the shift in τ. This is a manifold of dimension $\mu(x_b)-\mu(x_a)-1$. The choice of orientation of the negative subspace $T^{(-)}_{x_a}M$ determines an orientation of $T^{(+)}_{x_a}M$, and these determine an orientation of all $\mathcal{M}(x_a,x_b)$. For the x_i and x_j with $\mu_j-\mu_i=2$, $\mathcal{M}(x_i,x_j)$ is a one-dimensional oriented manifold. The boundary of $\mathcal{M}(x_i,x_j)$ consists of "broken flow lines" where breaking occurs at the critical points x_k with $\mu_k=\mu_i+1$. Namely, we have

$$\partial\mathcal{M}(x_i,x_j)=\bigcup_{x_k:\mu_k=\mu_i+1}\mathcal{M}(x_i,x_k)\times\mathcal{M}(x_k,x_j),\tag{10.355}$$

and one can show that this holds including the orientation. On the other hand, for x_i and x_k with $\mu_k=\mu_i+1$, $\mathcal{M}(x_i,x_k)$ is a discrete set of oriented points consisting of gradient flow lines from x_i to x_k, and it is easy to see that n_γ determines the orientation of the point represented by γ. Namely,

we see that

$$\#\mathcal{M}(x_i,x_k)=\sum_{\gamma:i\to k}n_\gamma.\tag{10.356}$$

Since the number of boundary points is equal to zero, $\#\partial\mathcal{M}(x_i,x_j)=0$, Eq. (10.355) yields what we wanted, Eq. (10.353).

In the above example of $M=S^2$, we have $\mathcal{M}(A,B)=\{\gamma_1,\gamma_2\}$, $\mathcal{M}(B,C)=\{\gamma_3\}$ and $\mathcal{M}(B,D)=\{\gamma_4\}$. The one-dimensional space $\mathcal{M}(\mathbf{A},C)$ consists of the thin lines as depicted in Fig. 9. It is easy to see that there are two boundary lines which are the broken lines $\gamma_1\#\gamma_3$ and $\gamma_2\#\gamma_3$. This indeed shows that $\partial\mathcal{M}(A,C)=\mathcal{M}(A,B)\times\mathcal{M}(B,C)$.

10.5.5. Bott–Morse Function. In the above discussion, we have assumed that h has only non-degenerate and therefore isolated critical points. It is a natural question to ask what happens if this condition is relaxed. Here we briefly comment on the case where h admits critical manifolds of dimension >0 but h is still non-degenerate in the normal direction. Such a function is called *Bott–Morse*. Let M_i $(i=1,\dots,N)$ be the connected components of the critical point set of h. By the Bott–Morse assumption, M_i is a smooth submanifold of M, where the Hessian of h has zero eigenvalues only in the direction tangent to M_i. We define the Morse index μ_i of M_i to be the number of negative eigenvalues of the Hessian.

The spectrum of a supersymmetric ground state is invariant under the rescaling $h\to\lambda h$ and we consider, as before, the large λ limit. Then the ground state wave-function is localized at the critical point set $\cup_i M_i$. We first focus on one component, say M_1. Near each point of M_1, the analysis decomposes into two parts — directions normal to M_1 and directions tangent to M_1. In the normal directions, a zero energy state is a μ_1-form, which is a volume form on the negative eigenspace of the Hessian. If the bundle over M_1 of the negative eigenspaces is orientable, then these normal μ_1-forms glue together to make a globally defined μ_1-form Ψ_1. In what follows, we assume that this is the case although the other case can be treated with a slight modification. In the tangent directions, the Hamiltonian is essentially the Laplacian, and the harmonic forms are the zero energy states. Thus, the perturbative ground states localized at M_1 are of the form $\omega_\alpha\wedge\Psi_1$ where ω_α are harmonic forms on M_1. Collecting together the states from all M_i's, we obtain $\sum_i\dim H^\bullet(M_i)$ perturbative ground states. In the strict $\lambda\to\infty$

limit, these, and only these, are the exact zero energy states. Since the true zero energy state for finite λ remains also the zero energy state in the $\lambda \to \infty$ limit, we obtain the inequality $\dim \mathcal{H}_{(0)} \leq \sum_i \dim H^\bullet(M_i)$. Since $\mathcal{H}_{(0)} \cong H^\bullet(M)$ this means

$$\dim H^\bullet(M) \leq \sum_i \dim H^\bullet(M_i). \tag{10.357}$$

So much for the perturbative analysis. As in the case where h is non-degenerate, these approximate ground states may be lifted to have nonzero energy by quantum tunneling or instanton effects. A way to incorporate tunneling has been found by Austin and Braam, which we describe here. (The derivation by the physics analysis is left as an exercise for the readers.) We denote by R_p the union of critical submanifolds of Morse index p, and we assume that there is no ascending gradient flow from R_p to R_q if $p > q$. Let $\mathcal{M}(R_p, R_q)$ be the space of ascending gradient flow lines from R_p to R_q. For each gradient flow line $\phi : \mathbb{R} \to M$, we have the initial point $\phi(-\infty) \in R_p$ and final point $\phi(+\infty) \in R_q$. This defines the initial and final maps

$$i_p^q : \mathcal{M}(R_p, R_q) \to R_p,$$
$$f_p^q : \mathcal{M}(R_p, R_q) \to R_q.$$

Now we put

$$C^r = \bigoplus_{p+q=r} \Omega^p(R_q), \tag{10.358}$$

and define the operator $Q : C^r \to C^{r+1}$ by $\sum_{s\geq 0} Q_s$ where

$$Q_s : \Omega^p(R_q) \to \Omega^{p-s+1}(R_{q+s}), \tag{10.359}$$
$$\omega \mapsto \begin{cases} d\omega & s = 0, \\ (-1)^p (f_q^{q+s})_* (i_q^{q+s})^* \omega & \text{otherwise}. \end{cases}$$

Here $(i_q^{q+s})^*$ is the pull-back of forms from R_p to $\mathcal{M}(R_q, R_{q+s})$ and $(f_q^{q+s})_*$ is the integration along the fiber of the final point projection $\mathcal{M}(R_q, R_{q+s}) \to R_{q+s}$. Then $(C^\bullet, Q)$ defines a complex. This complex is actually filtered,

$$0 \subset \cdots \subset C_1^\bullet \subset C_0^\bullet = C^\bullet$$

with

$$C_k^r = \bigoplus_{q\geq k} \Omega^{r-q}(R_q).$$

The quotient $GC_k^\bullet = C_k^\bullet / C_{k+1}^\bullet$ is equal to the de Rham complex of R_k. We can apply the method of spectral sequences to compute the cohomology of the complex $(C^\bullet, Q)$. The E_1 term is then given by $E_1^{k,r} = H^{r-k}(R_k)$. This is the space of approximate zero energy states obtained by the perturbation theory. The cohomology of the full complex $(C^\bullet, Q)$ is isomorphic to the space of exact zero energy states. This is how the instanton effect is taken into account.

10.5.6. Moment Map for $U(1)$ Actions. In certain cases, the problem of finding the supersymmetric ground states simply reduces to the computation of cohomology of the individual critical manifolds M_i. Such is the case where h is the moment map on a $U(1)$-invariant Kähler manifold.

Let M be a Kähler manifold with a $U(1)$ action that preserves both the metric and the complex structure. Then, the Kähler form ω is $U(1)$-invariant. A *moment map* h associated with the $U(1)$ action is a function on M such that the one-form dh is equal to $i_v\omega$ where v is a vector field on M that generates the $U(1)$ action. (Note that $di_v\omega = L_v\omega - i_v d\omega = 0 - 0 = 0$ because ω is a $U(1)$-invariant closed form. Thus one can find a function h such that $i_v\omega = dh$, at least locally. The assumption here is that h solves this equation globally.) The critical points of h are the fixed points of the $U(1)$ action. The Bott–Morse assumption is automatically satisfied for h, where M_i are components of the fixed point manifold.

The reduction of the problem can be shown as follows. One can find $U(1)$-invariant tubular neighborhoods U_i of M_i which do not intersect with one another. For each i, we choose a $U(1)$-invariant smooth function h_i supported on U_i which is a Morse function when restricted on M_i (with non-degenerate critical points only). Let us then replace the function h by

$$h_\epsilon = h + \epsilon \sum_{i=1}^N h_i. \tag{10.360}$$

The standard conjugation argument shows that this replacement does not affect the spectrum of supersymmetric ground states. For a sufficiently small ϵ, the function h_ϵ has isolated non-degenerate critical points only, and all of them are $U(1)$-fixed points, namely, in $\cup_i M_i$. Since the critical points are all non-degenerate, the supersymmetric ground states are the cohomology classes of the standard Morse–Witten complex. We now show that the coboundary operator of the complex receives contributions only from gradient flows that lies inside M_i's. We recall that non-trivial contributions come only from isolated gradient flows ("isolated" except for the shift of the domain parameter). Let γ be a gradient flow of h_ϵ from a critical point p to a critical point q that lies (partly) outside $\cup_i M_i$. A flow from $p \in M_i$ to $q \in M_j$ with $i \neq j$ must always be of this kind. Then, its $U(1)$-rotations are also gradient flows of h_ϵ from p to q, and they make a non-trivial one parameter family. Thus, γ is not isolated and cannot contribute in the coboundary operator. This shows that the Morse–Witten complex splits into the indivisual ones for M_i defined by the Morse function$h_i|_{M_i}$. In particular, the cohomology group splits into the sum of the cohomology groups of M_i's.

A Morse function is said to be *perfect* if the coboundary operator (10.346) is trivial, namely, if the perturbative ground states Ψ_i at the critical points x_i are all exact ground states. This notion of perfectness can be generalized to Bott–Morse functions in a obvious way. What we have shown above is that the moment map assicated with a $U(1)$ action on a Kähler manifold is a perfect Bott–Morse function.

10.5.7. Application to Quantum Field Theory. Later, we will apply this method to quantum field theories in $(1+1)$ dimensions, which can be considered roughly as quantum mechanics with infinitely many degrees of freedom. In that setting we will need to consider an infinite-dimensional manifold M. There are two main subtle points associated with the infinite-dimensionality. One is that the definition of Morse index is not obvious. As we will see, the spectrum of the Hessian is not bounded from below nor from above. This problem will be partially solved by some kind of regularization, but sometimes the Morse index can be defined only up to addition of some connstant. This is related to an anomaly of fermion number conservation. Another, and more serious, problem is that the relation $Q^2 = 0$ is not automatic. Sometimes it fails because of the failure of Eq. (10.355), which would mean that the supersymmetry algebra itself is anomalous, and one would not be able to consider the "Q-complex". Such a phenomenon does not happen, fortunately, for the theory of closed strings, but will happen for open strings.

CHAPTER 11

Free Quantum Field Theories in $1+1$ Dimensions

As already mentioned, the higher the dimension of the QFT, the more complicated it will be. We will be interested mainly in the case of QFTs in two dimensions, the topic to which we now turn. In this chapter we will be dealing mainly with the simplest two-dimensional QFTs, those that are "free" in dimension 2. By free, we mean that the action is quadratic in the field variables. This is the case where everything can be done exactly and explicitly and serves as a good introduction to more complicated two-dimensional QFTs which we will deal with later. Moreover, as in quantum mechanics, they play an important role as the starting point of perturbation theory or semi-classical approximation in a more general interactive theory. In supersymmetric theories, some quantities are determined exactly using quadratic approximation of the theory, and the role of free field theories is even more important.

There is another reason to single out free theories: the sigma model with target a circle of radius R provides an example of a free theory. It turns out that this example is already rich enough to exhibit a duality that is an equivalence between the sigma model on a circle of radius R and that of radius $1/R$. This is known as T-duality. In the supersymmetric setting, T-duality is the basic example of mirror symmetry, as will be studied in later sections. We will see that mirror symmetry is in a sense the refinement of T-duality.

11.1. Free Bosonic Scalar Field Theory

11.1.1. Classical Theory. We start our study of quantum field theory in $1+1$ dimensions with the free theory of a single scalar field x. We formulate the theory on the cylinder $\Sigma = \mathbb{R} \times S^1$ where $\mathbb{R}$ is parametrized by the time t and S^1 is parametrized by the spatial coordinate s of period

2π, $s \equiv s + 2\pi$. The action for the scalar field $x = x(t,s)$ is given by

$$S = \frac{1}{2\pi}\int_\Sigma L\, dt\, ds = \frac{1}{4\pi}\int_\Sigma \left((\partial_t x)^2 - (\partial_s x)^2\right) dt\, ds. \tag{11.1}$$

This theory can also be considered as a sigma model, where x defines a map of the worldsheet Σ to the target space $\mathbb{R}$. The Euler–Lagrange equation is given by

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial s^2}\right) x = 0. \tag{11.2}$$

This is solved by

$$x(t,s) = f(t-s) + g(t+s), \tag{11.3}$$

where f and g are arbitrary functions. The part $f(t-s)$ represents a configuration moving to the right, whereas $g(t+s)$ represents the left-moving configuration, both at the speed of light. These two motions do not interfere with each other. This is the decoupling of the right- and left- moving modes, which is a special property of massless fields in $1+1$ dimensions.

The action is invariant under the shift in x

$$\delta x = \alpha, \tag{11.4}$$

where α is a constant. One can find the corresponding conserved charges by following the Noether procedure. This time, we let the variation parameter depend on both temporal and spatial coordinates, $\alpha \to \alpha(t,s)$. Then the action varies as

$$\delta S = \frac{1}{2\pi}\int_\Sigma \partial_\mu \alpha\, j^\mu\, dt\, ds, \tag{11.5}$$

where

$$\begin{cases} j^t = \partial_t x, \\ j^s = -\partial_s x. \end{cases} \tag{11.6}$$

For a classical configuration that extremizes the action, this current j^μ obeys the conservation equation

$$\partial_\mu j^\mu = 0. \tag{11.7}$$

In particular, the charge

$$p = \frac{1}{2\pi}\int_{S^1} j^t\, ds \tag{11.8}$$

is a constant of motion. Since the shift in x can be considered as the translation of the target space $\mathbb{R}$, the conserved charge p can be interpreted as the target space momentum. The action is also invariant under worldsheet space-time translations

$$\delta_\alpha x = \alpha^\mu \partial_\mu x. \tag{11.9}$$

The conserved currents are

$$\begin{cases} T^t_{\ t} = \frac{1}{2}\left((\partial_t x)^2 + (\partial_s x)^2\right), \\ T^s_{\ t} = -\partial_s x \partial_t x, \end{cases} \quad \begin{cases} T^t_{\ s} = \partial_s x \partial_t x, \\ T^s_{\ s} = -\frac{1}{2}\left((\partial_t x)^2 + (\partial_s x)^2\right) \end{cases} \tag{11.10}$$

and the conserved charges are

$$H = \frac{1}{2\pi}\int_{S^1} T^t_{\ t}\, ds = \frac{1}{2\pi}\int_{S^1} \frac{1}{2}\left((\partial_t x)^2 + (\partial_s x)^2\right) ds, \tag{11.11}$$

$$P = \frac{1}{2\pi}\int_{S^1} T^t_{\ s}\, ds = \frac{1}{2\pi}\int_{S^1} \partial_t x \partial_s x\, ds. \tag{11.12}$$

These are respectively the Hamiltonian and momentum of the system.

Let us consider the Fourier expansion of $x(t,s)$ along S^1:

$$x(t,s) = x_0(t) + \sum_{n\neq 0} x_n(t) e^{ins}. \tag{11.13}$$

Since $x(t,s)$ is real-valued, $x_0(t)$ is real and $x_{-n}(t)$ is the complex conjugate of $x_n(t)$, $(x_n(t))^* = x_{-n}(t)$. The action is then expressed as

$$S = \int dt \left\{ \frac{1}{2}(\dot{x}_0)^2 + \sum_{n=1}^{\infty}(|\dot{x}_n|^2 - n^2|x_n|^2)\right\}. \tag{11.14}$$

We see from this expression that this free theory consists of infinitely many decoupled systems; a single real scalar x_0 without a potential, and a complex scalar x_n with the harmonic oscillator potential $U = n^2|x_n|^2$, where n varies over $\{1,2,3,\ldots\}$. In this way we have reduced the difficulty of dealing with a theory in 1+1 dimensions, to a theory in 1 dimension, but with infinitely many degrees of freedom.

11.1.2. Quantization. Let us quantize this system. In principle, we should obtain as the Hilbert space a suitable space of functions on the loop space of $\mathbb{R}$. The fact that we have decomposed the system to infinitely many degrees of freedom already will lead to the appropriate notion of function space by considering the infinite tensor product of the Hilbert spaces of each of the decoupled systems. We have already analyzed all the constituent theories, so we can borrow the results. We first consider the sector of the real scalar x_0. The conjugate momentum for x_0 is $p_0 = \dot{x}_0$ and there is a momentum eigenstate $|k\rangle_0$ for each k;

$$p_0|k\rangle_0 = k|k\rangle_0. \tag{11.15}$$

This is also the energy $k^2/2$ eigenstate of the Hamiltonian

$$H_0 = \frac{1}{2}p_0^2. \tag{11.16}$$

Let us next consider the nth harmonic oscillator, x_n. x_n is a complex variable and decomposes into two real variables x_{1n} and x_{2n} defined by $x_n = (x_{1n} + ix_{2n})/\sqrt{2}$. As usual, one can define the operators $a_{in} = (p_{in}/\sqrt{n} - i\sqrt{n}x_{in})/\sqrt{2}$ and $a_{in}^\dagger = (p_{in}/\sqrt{n} + i\sqrt{n}x_{in})/\sqrt{2}$ for $i = 1, 2$, where $p_{in} = \dot{x}_{in}$. These obey the commutation relations $[a_{in}, a_{jn}^\dagger] = \delta_{i,j}$, $[a_{in}, a_{jn}] = [a_{in}^\dagger, a_{jn}^\dagger] = 0$. The Hamiltonian is given by

$$H_n = n\left(a_{1n}^\dagger a_{1n} + \frac{1}{2}\right) + n\left(a_{2n}^\dagger a_{2n} + \frac{1}{2}\right). \tag{11.17}$$

Now, let us change the variables as $\alpha_n = \sqrt{n/2}(a_{1n} + ia_{2n})$, $\alpha_{-n} = \alpha_n^\dagger = \sqrt{n/2}(a_{1n}^\dagger - ia_{2n}^\dagger)$, $\widetilde{\alpha}_n = \sqrt{n/2}(a_{1n} - ia_{2n})$ and $\widetilde{\alpha}_{-n} = \widetilde{\alpha}_n^\dagger = \sqrt{n/2}(a_{1n}^\dagger + ia_{2n}^\dagger)$ where we take $n \geq 1$ here. These new operators satisfy the relations

$$[\alpha_n, \alpha_{-n}] = [\widetilde{\alpha}_n, \widetilde{\alpha}_{-n}] = n, \ [\alpha_n, \widetilde{\alpha}_{\pm n}] = [\alpha_{-n}, \widetilde{\alpha}_{\pm n}] = 0. \tag{11.18}$$

Thus, α_{-n} and $\widetilde{\alpha}_{-n}$ are the creation operators while α_n and $\widetilde{\alpha}_n$ are the annihilation operators. In terms of these variables the Hamiltonian H_n is expressed as

$$H_n = \alpha_{-n}\alpha_n + \widetilde{\alpha}_{-n}\widetilde{\alpha}_n + n. \tag{11.19}$$

We define $|0\rangle_n$ as the vector annihilated by α_n and $\widetilde{\alpha}_n$. This is a ground state for the Hamiltonian H_n, with energy n. A general energy eigenstate is constructed by multiplying powers of creation operators α_{-n} and $\widetilde{\alpha}_{-n}$ acting on $|0\rangle$.

The Hilbert space of the total system is a tensor product of the Hilbert spaces of these constituent theories. Let us define the state

$$|k\rangle := |k\rangle_0 \otimes \bigotimes_{n=1}^{\infty} |0\rangle_n. \tag{11.20}$$

Then a general state is constructed by multiplying the powers of α_{-n} and $\widetilde{\alpha}_{-n}$ for various n. The Hamiltonian is the sum

$$\begin{aligned} H =& H_0 + \sum_{n=1}^{\infty} H_n \\ =& \frac{1}{2}p_0^2 + \sum_{n=1}^{\infty}(\alpha_{-n}\alpha_n + \widetilde{\alpha}_{-n}\widetilde{\alpha}_n + n) \\ =& \frac{1}{2}p_0^2 + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n + \sum_{n=1}^{\infty}\widetilde{\alpha}_{-n}\widetilde{\alpha}_n - \frac{1}{12} \end{aligned} \tag{11.21}$$

where we have used the zeta function regularization to sum up the ground state oscillation energies of the infinitely many harmonic oscillator systems:

$$\sum_{n=1}^{\infty} n = \zeta(-1) = -\frac{1}{12}. \tag{11.22}$$

The worldsheet momentum is

$$\begin{aligned} P =& \frac{1}{2\pi}\int_{S^1} \partial_t x \partial_s x ds = \sum_{n+m=0} im\dot{x}_n x_m \\ =& -\sum_{n=1}^{\infty}\alpha_{-n}\alpha_n + \sum_{n=1}^{\infty}\widetilde{\alpha}_{-n}\widetilde{\alpha}_n, \end{aligned} \tag{11.23}$$

where we used the relation $x_n = (\widetilde{\alpha}_{-n} - \alpha_n)/(\sqrt{2}in)$ and $\dot{x}_n = (\widetilde{\alpha}_{-n} + \alpha_n)/\sqrt{2}$ which can be derived by tracing the definition of α_n and $\widetilde{\alpha}_n$. The target space momentum is simply

$$p = \frac{1}{2\pi}\int_{S^1} \dot{x}\, ds = \dot{x}_0 = p_0. \tag{11.24}$$

The state

$$\prod_{n=1}^{\infty}(\alpha_{-n})^{m_n}(\widetilde{\alpha}_{-n})^{\widetilde{m}_n}|k\rangle \tag{11.25}$$

has the following worldsheet energy and momentum

$$H = \frac{k^2}{2} + \sum_{n=1}^{\infty} n(m_n + \widetilde{m}_n) - \frac{1}{12}, \tag{11.26}$$

$$P = \sum_{n=1}^{\infty} n(-m_n + \widetilde{m}_n), \tag{11.27}$$

and also has the target space momentum

$$p = k. \tag{11.28}$$

The state $|0\rangle = |k = 0\rangle$ is the unique ground state with the ground state energy

$$E_0 = -\frac{1}{12}, \tag{11.29}$$

and target space momentum $p = 0$.

We note that

$$[H, x_0] = -ip_0,\ [H, p_0] = 0, \tag{11.30}$$

$$[H, \alpha_n] = -n\alpha_n,\ [H, \widetilde{\alpha}_n] = -n\widetilde{\alpha}_n. \tag{11.31}$$

Thus, we have

$$x_0(t) = \mathrm{e}^{iHt} x_0\, \mathrm{e}^{-iHt} = x_0 + tp_0, \tag{11.32}$$

$$\alpha_n(t) = \mathrm{e}^{iHt} \alpha_n\, \mathrm{e}^{-iHt} = \mathrm{e}^{-int}\alpha_n, \tag{11.33}$$

$$\widetilde{\alpha}_n(t) = \mathrm{e}^{iHt} \widetilde{\alpha}_n\, \mathrm{e}^{-iHt} = \mathrm{e}^{-int}\widetilde{\alpha}_n. \tag{11.34}$$

Since $x_n = (\widetilde{\alpha}_{-n} - \alpha_n)/(\sqrt{2}in)$ we obtain

$$x(t, s) = x_0 + tp_0 + \frac{i}{\sqrt{2}} \sum_{n \neq 0} \frac{1}{n} (\alpha_n\, \mathrm{e}^{-in(t-s)} + \widetilde{\alpha}_n\, \mathrm{e}^{-in(t+s)}). \tag{11.35}$$

Note that this is the most general solution to the equation of motion, Eq. (11.2), that is compatible with the periodicity $x(t, s + 2\pi) = x(t, s)$. Also, we now see that α_n are the right-moving modes and $\widetilde{\alpha}_n$ are the left-moving modes. Eq. (11.35) is consistent with

$$[P, x_0] = 0,\ [P, p_0] = 0, \tag{11.36}$$

$$[P, \alpha_n] = n\alpha_n,\ [P, \widetilde{\alpha}_n] = -n\widetilde{\alpha}_n. \tag{11.37}$$

11.1.3. Vertex Operators. In Eqs. (11.21)–(11.23), which express the Hamiltonian and momentum, the annihilation operators α_n, $\widetilde{\alpha}_n$ $(n > 0)$ appear to the right of the creation operators α_{-n}, $\widetilde{\alpha}_{-n}$. This is called the *normal ordering*. We introduce the symbol $:(-):$ to indicate the normal ordering. For example, for $n \geq 1$,

$$:\alpha_{-n}\alpha_n: = :\alpha_n\alpha_{-n}: = \alpha_{-n}\alpha_n,\quad :\widetilde{\alpha}_{-n}\widetilde{\alpha}_n: = :\widetilde{\alpha}_n\widetilde{\alpha}_{-n}: = \widetilde{\alpha}_{-n}\widetilde{\alpha}_n. \tag{11.38}$$

Also, we extend it to the zero modes x_0, p_0 so that

$$:x_0 p_0: = :p_0 x_0: = x_0 p_0. \tag{11.39}$$

Then it is straightforward to see that

(11.40)
$$x(t_1, s_1)x(t_2, s_2) = :x(t_1, s_1)x(t_2, s_2): - it_1 + \frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{n}\left((z_2/z_1)^n + (\widetilde{z}_2/\widetilde{z}_1)^n\right),$$

where $z_j = \mathrm{e}^{i(t_j - s_j)}$ and $\widetilde{z}_j = \mathrm{e}^{i(t_j + s_j)}$. The infinite sum is oscillatory and ambiguous. From now on, we assume an infinitesimal Wick rotation $t \to \mathrm{e}^{-i\epsilon}t$ with $\epsilon > 0$ (the complete Wick rotation $\epsilon = \pi/2$ would lead to $\widetilde{z}_i = \bar{z}_i$). If $t_1 > t_2$, we have $|z_2/z_1| < 1$, $|\widetilde{z}_2/\widetilde{z}_1| < 1$, and the sum is convergent to $-\frac{1}{2}\log(1 - z_2/z_1) - \frac{1}{2}\log(1 - \widetilde{z}_2/\widetilde{z}_1)$. This convergence shows that

$$\mathrm{T}\Big[x(t_1, s_1)x(t_2, s_2)\Big] = :x(t_1, s_1)x(t_2, s_2): - \frac{1}{2}\log[(z_1 - z_2)(\widetilde{z}_1 - \widetilde{z}_2)], \tag{11.41}$$

where $\mathrm{T}[A(t_1, s_1)B(t_2, s_2)]$ is the time ordered product, which is $A(1)B(2)$ if $t_1 > t_2$ and $B(2)A(1)$ if $t_2 > t_1$.

The normal ordered operator for $\exp(ikx)$ is expressed as

(11.42)
$$:\exp\Big(ikx(t, s)\Big):$$
$$= \mathrm{e}^{ik\frac{i}{\sqrt{2}}\sum_{n=1}^{\infty} \frac{-1}{n}(\alpha_{-n}z^n + \widetilde{\alpha}_{-n}\widetilde{z}^n)}\, \mathrm{e}^{ikx_0}\, \mathrm{e}^{iktp_0}\, \mathrm{e}^{ik\frac{i}{\sqrt{2}}\sum_{n=1}^{\infty}\frac{1}{n}(\alpha_n z^{-n} + \widetilde{\alpha}_n\widetilde{z}^{-n})}.$$

It acts on the vacuum $|0\rangle$ as

$$:\mathrm{e}^{ikx(t,s)}:|0\rangle = \mathrm{e}^{ik\frac{i}{\sqrt{2}}\sum_{n=1}^{\infty}\frac{-1}{n}(\alpha_{-n}z^n + \widetilde{\alpha}_{-n}\widetilde{z}^n)}\, \mathrm{e}^{ikx_0}|0\rangle. \tag{11.43}$$

Since e^{ikx_0} increases the target space momentum p by k, we have $\mathrm{e}^{ikx_0}|0\rangle = |k\rangle$. This can also be seen by noting that $|k\rangle$ is represented by the wave-function $\Psi_k(x) = \mathrm{e}^{ikx}$ while the operator e^{ikx_0} is represented by the multiplication by e^{ikx}. This latter representation also shows that $\langle k_1|k_2\rangle = 2\pi\delta(k_1 - k_2)$. If we take the limit $t \to -\infty$, we have $|z| \to 0$ and $:\mathrm{e}^{ikx(t,s)}:|0\rangle$ converges to

$$:\mathrm{e}^{ikx(t,s)}:|0\rangle \xrightarrow{t \to -\infty} \mathrm{e}^{ikx_0}|0\rangle = |k\rangle. \tag{11.44}$$

Thus, this operation increases the momentum by k. The operator shown in Eq. (11.42) is called the *vertex operator* of (target space) momentum k.

It is easy to compute the two-point correlation function of the vertex operators.

$$\begin{aligned}(11.45)\qquad &\langle e^{ik_1x(t_1,s_1)}\, e^{ik_2x(t_2,s_2)}\rangle \\ &= \langle 0|\mathrm{T}\Big[:e^{ik_1x(t_1,s_1)}::e^{ik_2x(t_2,s_2)}:\Big]|0\rangle \\ &= 2\pi\delta(k_1+k_2)[(z_1-z_2)(\tilde{z}_1-\tilde{z}_2)]^{\frac{k_1k_2}{2}}.\end{aligned}$$

11.1.4. Partition Function. Let us now compute the partition function of the system. As we have seen in quantum mechanics, the partition function can be defined as

$$Z(\beta) = \mathrm{Tr}\, e^{-\beta H}. \tag{11.46}$$

This partition function corresponds to evaluating the path-integral where the worldsheet is the Euclidean cylinder of length β with the two boundaries identified. Thus the worldsheet in this case is a rectangular torus with sides 2π and β. Actually this is not the most general thing we can do. We can also try to evaluate the path-integral on a torus which is not rectangular but is skewed as shown in Fig. 1. This corresponds to shifting one end of the

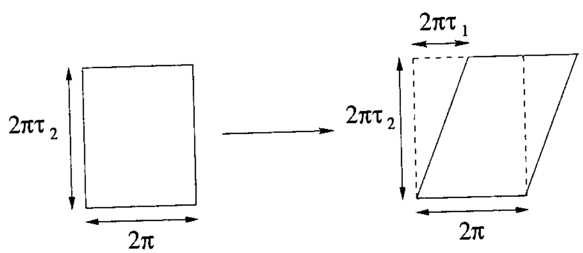


FIGURE 1

cylinder by $2\pi\tau_1$ before identifying it with the other end. (We also rename the length as $\beta \to 2\pi\tau_2$.) In the operator language this operation of rotating corresponds to inserting the translation operator $e^{-2\pi i\tau_1 P}$ in the trace,

$$Z(\tau_1,\tau_2) = \mathrm{Tr}\; e^{-2\pi i\tau_1 P}\, e^{-2\pi\tau_2 H}. \tag{11.47}$$

Let us define

$$H_R := \frac{1}{2}(H-P) = \frac{1}{4}p_0^2 + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - \frac{1}{24}, \tag{11.48}$$

$$H_L := \frac{1}{2}(H+P) = \frac{1}{4}p_0^2 + \sum_{n=1}^{\infty}\tilde{\alpha}_{-n}\tilde{\alpha}_n - \frac{1}{24}, \tag{11.49}$$

which involve left-moving and right-moving nonzero modes respectively. Then the partition function can be written as

$$\begin{aligned}Z(\tau,\bar{\tau}) &= \mathrm{Tr}\; e^{2\pi i\tau H_R}\, e^{-2\pi i\bar{\tau}H_L}, \\ &= \mathrm{Tr}\; q^{H_R}\bar{q}^{H_L}\end{aligned} \tag{11.50}$$

where

$$\tau = \tau_1 + i\tau_2, \tag{11.51}$$

and $q = e^{2\pi i\tau}$. Recall that the Hilbert space is the tensor product of Hilbert spaces of infinitely many decoupled systems — the free particle system of zero modes and right-moving and left-moving harmonic oscillator modes of frequency n. Denoting the respective Hilbert spaces by $\mathcal{H}_0$, $\mathcal{H}_n^R$ and $\mathcal{H}_n^L$, we obtain the factorized form of the partition function

$$Z(\tau,\bar{\tau}) = (q\bar{q})^{-1/24}\mathrm{Tr}_{\mathcal{H}_0}(q\bar{q})^{p_0^2/4}\prod_{n=1}^{\infty}\mathrm{Tr}_{\mathcal{H}_n^R}q^{\alpha_{-n}\alpha_n}\mathrm{Tr}_{\mathcal{H}_n^L}\bar{q}^{\tilde{\alpha}_{-n}\tilde{\alpha}_n}, \tag{11.52}$$

where the prefactor $(q\bar{q})^{-1/24} = e^{-2\pi\tau_2(-1/12)}$ comes from the regularized zero point oscillation energy of the infinitely many harmonic oscillator systems, as shown in Eq. (11.22). It is easy to evaluate each factor;

$$\mathrm{Tr}_{\mathcal{H}_n^R}q^{\alpha_{-n}\alpha_n} = \sum_{k=0}^{\infty}q^{nk} = \frac{1}{1-q^n}, \tag{11.53}$$

$$\mathrm{Tr}_{\mathcal{H}_n^L}\bar{q}^{\tilde{\alpha}_{-n}\tilde{\alpha}_n} = \frac{1}{1-\bar{q}^n}, \tag{11.54}$$

$$\mathrm{Tr}_{\mathcal{H}_0}(q\bar{q})^{p_0^2/4} = \mathrm{Tr}_{\mathcal{H}_0}\, e^{-2\pi\tau_2 H_0} = V\int_{-\infty}^{+\infty}\frac{dp}{2\pi}\, e^{-2\pi\tau_2(\frac{1}{2}p^2)} = \frac{V}{2\pi}\frac{1}{\sqrt{\tau_2}}. \tag{11.55}$$

In the last part, V stands for the cut-off volume in order to make the partition function finite. Putting all these factors together we obtain

$$\begin{aligned} Z(\tau,\bar{\tau}) &= (q\bar{q})^{-1/24}\frac{V}{2\pi}\frac{1}{\sqrt{\tau_2}}\prod_{n=1}^{\infty}\left|\frac{1}{1-q^n}\right|^2 \\ &= \frac{V}{2\pi}\frac{1}{\sqrt{\tau_2}}|\eta(\tau)|^{-2}, \end{aligned} \tag{11.56}$$

where $\eta(\tau)$ is the Dedekind eta function

$$\eta(\tau) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n). \tag{11.57}$$

Using the modular transformation properties of the eta function

$$\eta(\tau+1) = e^{\pi i/12}\eta(\tau), \quad \eta(-1/\tau) = (-i\tau)^{1/2}\eta(\tau), \tag{11.58}$$

one sees that the partition function is invariant under the differomorphisms on T^2 acting on τ as

$$\tau \longmapsto \frac{a\tau+b}{c\tau+d}. \tag{11.59}$$

This is as it should be, and can be viewed as another confirmation of the regularization procedure we used. (Note in particular that the leading power of q comes from the zeta function regularization, and without the correct factor the modular invariance would be lost.) Note also that the partition function does not depend on the area of the worldsheet torus, but only depends on its complex structure. This is a feature of conformal theories. As we will discuss in more detail later, sigma models for generic target manifolds do not lead to conformal theories.

11.2. Sigma Model on Torus and T-duality

11.2.1. Sigma Model on S^1. Now consider the case where the target space is a circle S^1 of radius R instead of the real line. The theory is described by a single scalar field x which is periodic with period $2\pi R$:

$$x \equiv x + 2\pi R. \tag{11.60}$$

The classical action is still given by Eq. (11.1). As in the case of the real line, space-time translations and target space translations are symmetries of the system. The corresponding Noether charges H, P and p are expressed again by Eqs. (11.11), (11.12) and (11.8).

Unlike in case of the real line, since the circle has discrete Fourier modes (as we have studied in Sec. 10.1.1) the target space momentum is quantized in units of $1/R$:

$$p = l/R, \ l \in \mathbb{Z}. \tag{11.61}$$

Also, the target space coordinate x is not single-valued but is a periodic variable of period $2\pi R$. This means that there are topologically non-trivial field configurations in the theory which are classified by the winding number m defined by

$$x(s+2\pi) = x(s) + 2\pi m R. \tag{11.62}$$

As we have seen, the conserved current for the momentum is

$$\begin{cases} j^t = \partial_t x, \\ j^s = -\partial_s x. \end{cases} \tag{11.63}$$

One can find another current

$$\begin{cases} j^t_w = \partial_s x, \\ j^s_w = -\partial_t x, \end{cases} \tag{11.64}$$

which satisfies the "conservation equation" $\partial_\mu j^\mu_w = 0$ (this is not an equation of motion, but an identity, like the Bianchi identity $dF = 0$ for electromagnetism). The corresponding "charge" is

$$w = \frac{1}{2\pi}\int_{S^1} j^t_w ds = \frac{1}{2\pi}(x(2\pi) - x(0)) = mR \tag{11.65}$$

in the sector with winding number m. Thus, w is the topological charge that counts the winding number.

The Hilbert space $\mathcal{H}$ is decomposed into sectors labelled by two integers — momentum l and winding number m:

$$\mathcal{H} = \bigoplus_{(l,m)\in\mathbb{Z}\oplus\mathbb{Z}} \mathcal{H}_{(l,m)}. \tag{11.66}$$

The subspace $\mathcal{H}_{(l,m)}$ is the space with $p = l/R$ and $w = mR$ and contains a basic element

$$|l,m\rangle, \tag{11.67}$$

which is annihilated by α_n and $\tilde{\alpha}_n$ with $n > 0$. The space $\mathcal{H}_{(l,m)}$ is constructed by acting on $|l,m\rangle$ with the powers of the creation operators α_{-n} and $\tilde{\alpha}_{-n}$.

We denote by p_0 and w_0 the operators counting the momentum and the winding number

$$p_0|l,m\rangle = \frac{l}{R}|l,m\rangle,\quad w_0|l,m\rangle = mR|l,m\rangle. \tag{11.68}$$

The operator $e^{i\frac{l}{R}x_0}$ shifts the momentum. There should also be operators that shift the winding number. We denote them by $e^{imR\widehat{x}_0}$ so that

$$e^{i\frac{l_1}{R}x_0}|l,m\rangle = |l+l_1,m\rangle,\qquad e^{im_1R\widehat{x}_0}|l,m\rangle = |l,m+m_1\rangle. \tag{11.69}$$

The operators $x_0, p_0, \widehat{x}_0, w_0$ have the commutation relations

$$[x_0,p_0]=i,\ [\widehat{x}_0,w_0]=i, \tag{11.70}$$

while other commutators vanish. Let us denote

$$p_R = \frac{1}{\sqrt{2}}(p_0-w_0),\quad p_L = \frac{1}{\sqrt{2}}(p_0+w_0). \tag{11.71}$$

Then the field $x(t,s)$ decomposes as the sum $x_R(t-s)+x_L(t+s)$ of right-moving and left-moving fields that commute with each other;

$$x_R(t-s) = \frac{x_0-\widehat{x}_0}{2} + \frac{1}{\sqrt{2}}(t-s)p_R + \frac{i}{\sqrt{2}}\sum_{n\neq 0}\frac{1}{n}\alpha_n\, e^{-in(t-s)}, \tag{11.72}$$

$$x_L(t+s) = \frac{x_0+\widehat{x}_0}{2} + \frac{1}{\sqrt{2}}(t+s)p_L + \frac{i}{\sqrt{2}}\sum_{n\neq 0}\frac{1}{n}\widetilde{\alpha}_n\, e^{-in(t+s)}. \tag{11.73}$$

We note that the derivatives

$$\frac{1}{\sqrt{2}}(\partial_t-\partial_s)x = p_R + \sum_{n\neq 0}\alpha_n\, e^{-in(t-s)}, \tag{11.74}$$

$$\frac{1}{\sqrt{2}}(\partial_t+\partial_s)x = p_L + \sum_{n\neq 0}\widetilde{\alpha}_n\, e^{-in(t+s)}, \tag{11.75}$$

define currents that measure the charges p_R and p_L respectively. The world-sheet Hamiltonian H and momentum P are given by

$$H_R = \frac{1}{2}(H-P) = \frac{1}{2}p_R^2 + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - \frac{1}{24}, \tag{11.76}$$

$$H_L = \frac{1}{2}(H+P) = \frac{1}{2}p_L^2 + \sum_{n=1}^{\infty}\widetilde{\alpha}_{-n}\widetilde{\alpha}_n - \frac{1}{24}. \tag{11.77}$$

We see that there is a unique ground state $|0,0\rangle$ and the ground state energy is again

$$E_0 = -\frac{1}{12}. \tag{11.78}$$

The computation of the partition function is similar to the case of the sigma model on $\mathbb{R}$ except for the summation over the zero modes. Instead of the divergent factor $V/2\pi\sqrt{\tau_2}$ coming from the zero mode integral, we have the discrete sum over the momentum l and winding number m corresponding to the decomposition shown in Eq. (11.66). Namely, we have

$$Z(\tau,\overline{\tau};R) = \frac{1}{|\eta(\tau)|^2}\sum_{(l,m)\in\mathbb{Z}\oplus\mathbb{Z}} q^{\frac{1}{4}(l/R-mR)^2}\overline{q}^{\frac{1}{4}(l/R+mR)^2}. \tag{11.79}$$

The factor $|\eta(\tau)|^{-2}$ comes from the oscillator modes in precisely the same way as in the case of the sigma model on $\mathbb{R}$.

11.2.2. T-duality. We see that the partition function is invariant under the replacement $R\mapsto 1/R$:

$$Z(\tau,\overline{\tau};1/R) = Z(\tau,\overline{\tau};R). \tag{11.80}$$

The full spectrum is also invariant as long as we interchange the quantum numbers associated with the winding and the momentum as well, $l\leftrightarrow m$. Namely, there is an isomorphism of our Hilbert space $\mathcal{H}$ to the Hilbert space $\widehat{\mathcal{H}}$ of the sigma model on S^1 of radius $1/R$, under which

$$\mathcal{H}_{(l,m)} \longrightarrow \widehat{\mathcal{H}}_{(m,l)}. \tag{11.81}$$

This corresponds to the exchange of operators

$$(p_R,p_L)\mapsto(-\widehat{p}_R,\widehat{p}_L). \tag{11.82}$$

This symmetry of the theory is called $R\to 1/R$ duality or *T-duality*. Since p_R and p_L are the conserved charges, it is natural to expect that the corresponding currents given by Eqs. (11.75)–(11.74) also transform in the same way. Thus, we expect that T-duality maps the currents as

$$(\partial_t\pm\partial_s)x\mapsto\pm(\partial_t\pm\partial_s)\widehat{x}, \tag{11.83}$$

or in terms of the Fourier modes

$$\alpha_n\mapsto-\widehat{\alpha}_n,\quad \widetilde{\alpha}_n\mapsto\widehat{\widetilde{\alpha}}_n. \tag{11.84}$$

Finally, since $\widehat{x}_0$ generates the shift of m, which is the momentum of the T-dual theory, it can be identified as the zero mode of the coordinate $\widehat{x}$. To summarize, we have found

$$\widehat{x}(t,s) = -x_R(t-s)+x_L(t+s). \tag{11.85}$$

Everything we said here, including the point expressed by Eq. (11.84), can also be derived using the path-integral method.

Path-integral Derivation Let us formulate the theory on a Riemann surface Σ of genus g. We put a (Euclidean) metric $h = h_{\mu\nu}d\sigma^\mu d\sigma^\nu$ on Σ where $(\sigma^\mu) = (\sigma^1, \sigma^2)$ are local coordinates. We use a variable $\phi = x/R$ which is periodic with period 2π. The action is then written as

$$S_\varphi = \frac{1}{4\pi}\int_\Sigma R^2 h^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi\sqrt{h}\,d^2\sigma. \tag{11.86}$$

This action can also be obtained from the following action for φ and a one-form field $\mathcal{B}_\mu$:

$$S' = \frac{1}{2\pi}\int_\Sigma \frac{1}{2R^2}h^{\mu\nu}\mathcal{B}_\mu\mathcal{B}_\nu\sqrt{h}d^2\sigma + \frac{i}{2\pi}\int_\Sigma \mathcal{B}\wedge d\varphi. \tag{11.87}$$

Completing the square with respect to $\mathcal{B}_\mu$, which is solved by

$$\mathcal{B} = iR^2 * d\varphi, \tag{11.88}$$

and integrating it out, we obtain the action for the sigma model, as shown in Eq. (11.86).

EXERCISE 11.2.1. *Verify this claim.*

If, changing the order of integration, we first integrate over the scalar field φ, we obtain a constraint $d\mathcal{B} = 0$. This constraint is solved by

$$\mathcal{B} = d\vartheta_0 + \sum_{i=1}^{2g} a_i\omega^i, \tag{11.89}$$

where ϑ_0 is a real scalar field, ω_i $(i = 1, \ldots, 2g)$ are closed one-forms that represent a basis of $H^1(\Sigma, \mathbb{R}) \cong \mathbb{R}^{2g}$, and the a_i's are real numbers. One can choose the $2g$ one-forms ω^i such that there are one-cycles γ_i representing a basis of $H_2(\Sigma, \mathbb{Z}) \cong \mathbb{Z}^{2g}$ with

$$\int_{\gamma_i}\omega_j = \delta_{i,j}. \tag{11.90}$$

Then $\int_\Sigma \omega^i \wedge \omega^j = J^{ij}$ is a non-degenerate matrix with integral entries whose inverse is also an integral matrix. Integration over φ actually yields constraints on the a_j's as well. Recall that φ is a periodic variable of period 2π. This means that φ does not have to come back to its original value when circling along non-trivial one-cycles in Σ, but comes back to itself up to 2π shifts. If φ shifts by $2\pi n_i$ along the cycle γ_i, $d\varphi$ has an expansion like Eq. (11.89) with the coefficient $2\pi n_i$ for ω^i. Thus, for a general configuration of φ we have

$$d\varphi = d\varphi_0 + \sum_{i=1}^{2g} 2\pi n_i\omega^i, \tag{11.91}$$

where φ_0 is a single-valued function on Σ. Now, integration over φ means integration over the function φ_0 and summation over the integers n_i's. Integration over φ_0 yields the constraint $d\mathcal{B} = 0$ which is solved by Eq. (11.89). What about the summation over the n_i's? To see this we substitute in $\int \mathcal{B}\wedge d\varphi$ for $\mathcal{B}$ from Eq. (11.89);

$$\int_\Sigma \mathcal{B}\wedge d\varphi = 2\pi\sum_{i,j} a_i J^{ij} n_j. \tag{11.92}$$

Now, noting that J^{ij} is a non-degenerate integral matrix with an integral inverse and using the fact that $\sum_n e^{ian} = 2\pi\sum_m \delta(a - 2\pi m)$, we see that summation over n_i constrains the a_i's to be integer multiples of 2π;

$$a_i = 2\pi m_i, \quad m_i \in \mathbb{Z}. \tag{11.93}$$

Inserting this into Eq. (11.89), we see that $\mathcal{B}$ can be written as

$$\mathcal{B} = d\vartheta, \tag{11.94}$$

where now ϑ is a periodic variable of period 2π. Now, inserting this into the original action we obtain

$$S_\vartheta = \frac{1}{4\pi}\int_\Sigma \frac{1}{R^2}h^{\mu\nu}\partial_\mu\vartheta\partial_\nu\vartheta\sqrt{h}d^2x \tag{11.95}$$

which is an action for a sigma model with target space an S^1 of radius $1/R$. Thus, we have shown that the sigma model with target S^1 of radius R is equivalent to the model with radius $1/R$. Namely, we have shown the $R \to 1/R$ duality or T-duality using the path-integral method. The above path-integral manipulation is called a *duality transformation* and can also be applied to massless fields (including vector fields or higher-rank anti-symmetric tensor fields) in arbitrary dimensions.

Comparing Eq. (11.88) with Eq. (11.94), we obtain the relation

$$R\,d\varphi = i\frac{1}{R} * d\vartheta. \tag{11.96}$$

Since $Rd\varphi$ and $iR * d\varphi$ are the conserved currents in the original system that measure momentum and winding number respectively, Eq. (11.96) means that *momentum and winding number are exchanged under the* $R \to 1/R$ *duality.* This is exactly what we saw above in the operator formalism. In particular, Eq. (11.96) is nothing but (the Euclidean version of) Eq. (11.83). Eq. (11.96) explicitly shows that *equations of motion and Bianchi identities are exchanged.* This is a general property of duality transformations.

The vertex operator

$$\exp(i\vartheta) \tag{11.97}$$

that creates a unit momentum in the dual theory must be equivalent to an operator that creates a unit winding number in the original theory. This can be confirmed by the following path-integral manipulation. Let us consider the insertion of

$$\exp\left(-i\int_p^q \mathcal{B}\right) \tag{11.98}$$

in the system with the action shown in Eq. (11.87), where the integration is along a path τ emanating from p and ending on q. Then using Eq. (11.94) we see that

$$\exp\left(-i\int_p^q \mathcal{B}\right) = \mathrm{e}^{-i\vartheta(q)}\,\mathrm{e}^{i\vartheta(p)}. \tag{11.99}$$

On the other hand, the insertion of $\mathrm{e}^{-i\int_p^q \mathcal{B}}$ changes the $\mathcal{B}$-linear term in Eq. (11.87). We note that $\int_p^q \mathcal{B}$ can be expressed as $\int_\Sigma \mathcal{B}\wedge\omega$, where ω is a one-form with delta function support along the path τ. This ω can be written as $\omega = d\theta_\tau$ where θ_τ is a multi-valued function on Σ that jumps by 1 when crossing the path τ. Now, the modification of the action from Eq. (11.87) can be written as

$$\frac{i}{2\pi}\int_\Sigma \mathcal{B}\wedge d\varphi \longrightarrow \frac{i}{2\pi}\int_\Sigma \mathcal{B}\wedge d\varphi + i\int_p^q \mathcal{B} = \frac{i}{2\pi}\int_\Sigma \mathcal{B}\wedge d(\varphi + 2\pi\theta_\tau). \tag{11.100}$$

Integrating out $\mathcal{B}_\mu$, we obtain the action shown in Eq. (11.86) with φ replaced by $\varphi' = \varphi + 2\pi\theta_\tau$. Note that φ' jumps by 2π when crossing the path τ which starts and ends on p and q. In particular, it has winding number 1 and -1 around p and q respectively. Comparing with Eq. (11.99), we see that the insertion of $\mathrm{e}^{i\vartheta}$ creates the unit winding number in the original system.

11.2.3. Sigma Model on T^2. Now consider the case when the target space is a rectangular torus $T^2 = S^1_{R_1}\times S^1_{R_2}$ where R_1 and R_2 are the radii of the two circles. Since the theory consists of the sigma models on circles that are decoupled from each other, the Hilbert space is a tensor product of the constituent theories. One can replace the parameters R_1, R_2 of the theory by the area and the complex structure of the torus

$$A = \mathrm{area}/(2\pi)^2 = R_1R_2,\ \ \sigma = iR_1/R_2. \tag{11.101}$$

By T-duality, inverting the radius of one of the circles leaves the theory invariant but it changes the area and the complex structure of the torus. This actually interchanges A and the imaginary part of the complex structure σ. For instance, if we dualize on the second circle we have the transformation

$$(A, \mathrm{Im}\,\sigma) = (R_1R_2, R_1/R_2) \mapsto (A', \mathrm{Im}\,\sigma') = (R_1/R_2, R_1R_2). \tag{11.102}$$

In other words, the shape (complex structure) and the size (Kähler structure) of the target torus are exchanged under this duality. In the above discussion, we considered a rectangular torus where the complex structure is pure imaginary. More generally, the complex structure is parametrized by a complex number

$$\sigma = \sigma_1 + i\sigma_2 \tag{11.103}$$

whose real part σ_1 is a periodic parameter of period 1 that corresponds to deviation from the rectangular torus. On the other hand, the area is a single real parameter. Thus, it appears that the exchange under T-duality of the complex structure and the area fails in the general case. But this is misleading: one can consider deforming the theory by assigning a phase factor

$$\exp\left(i\int_\Sigma x^*B\right) \tag{11.104}$$

in the path-integral. Here x is considered as a map from the worldsheet to the target space T^2 and B is the cohomology class

$$B \in H^2(T^2, \mathbb{R}). \tag{11.105}$$

For instance, the path-integral representation of the partition function of the deformed theory is given by

$$Z = \int \mathcal{D}x\ \mathrm{e}^{-S}\,\mathrm{e}^{i\int_\Sigma x^*B}. \tag{11.106}$$

Since $H^2(T^2,\mathbb{R})$ is one-dimensional, we can represent the B-field as a number, which we will denote by B as well. We also note that $e^{i\int_\Sigma x^*B} = 1$ for any x if $B = 2\pi n$ for some integer n. Thus, we should consider B as a periodic variable of period 2π. We define the complexified area ρ by

$$\rho = \frac{B}{2\pi} + iA. \tag{11.107}$$

Then one can show that T-duality on one of the circles exchanges the complexified area ρ and the complex structure σ. It is a good exercise to show that the partition function is invariant under this exchange.

EXERCISE 11.2.2. *Compute the partition function of the theory on T^2 with a B-field, and show that it is invariant under the interchange of σ and ρ.*

11.3. Free Dirac Fermion

Another important example of a free QFT is the theory of free Dirac fermions. A Dirac fermion is an anti-commuting complex spinor field. (we could also consider the case of real fermions — called Majorana fermions — which have half as many degrees of freedom as the one we will be studying here). In $(1+1)$-dimensional Minkowski space, the generators of the Clifford algebra $(e^t)^2 = -(e^s)^2 = 1$, $e^te^s = -e^se^t$, are represented by 2×2 matrices

$$\gamma^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \gamma^s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{11.108}$$

The Dirac fermion is represented by a column vector

$$\psi = \begin{pmatrix} \psi_- \\ \psi_+ \end{pmatrix}. \tag{11.109}$$

The action is given by

$$\begin{aligned} S =& \frac{1}{2\pi}\int_\Sigma i\overline{\psi}\gamma^\mu\partial_\mu\psi\ dt\,ds \\ =& \frac{1}{2\pi}\int_\Sigma \left(i\overline{\psi}_-(\partial_t+\partial_s)\psi_- + i\overline{\psi}_+(\partial_t-\partial_s)\psi_+\right) dt\,ds, \end{aligned} \tag{11.110}$$

where $\overline{\psi} = \psi^\dagger\gamma^t$ and $\overline{\psi}_\pm = \psi_\pm^\dagger$. Here Σ is the worldsheet which we take again to be $\mathbb{R}\times S^1$. The equation of motion is the Dirac equation $\gamma^\mu\partial_\mu\psi = 0$, namely

$$(\partial_t+\partial_s)\psi_- = 0,\ (\partial_t-\partial_s)\psi_+ = 0. \tag{11.111}$$

These equations are solved by

$$\psi_-(t,s) = f(t-s),\ \psi_+(t,s) = g(t+s). \tag{11.112}$$

Thus, ψ_- is a right-moving field and ψ_+ is a left-moving field.

The action is invariant under the phase rotations of the fermions

$$V : \psi_\pm \mapsto e^{-i\alpha}\psi_\pm, \tag{11.113}$$

$$A : \psi_\pm \mapsto e^{\mp i\beta}\psi_\pm. \tag{11.114}$$

We call them the *vector rotation* and the *axial rotation* respectively. By the Noether procedure, we find the corresponding conserved currents

$$V : \begin{cases} j_V^t = \overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+, \\ j_V^s = \overline{\psi}_-\psi_- - \overline{\psi}_+\psi_+, \end{cases} \quad A : \begin{cases} j_A^t = -\overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+, \\ j_A^s = -\overline{\psi}_-\psi_- - \overline{\psi}_+\psi_+, \end{cases} \tag{11.115}$$

and conserved charges

$$F_V = \frac{1}{2\pi}\int_{S^1} j_V^t\, ds = \frac{1}{2\pi}\int_{S^1}\left(\overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+\right) ds, \tag{11.116}$$

$$F_A = \frac{1}{2\pi}\int_{S^1} j_A^t\, ds = \frac{1}{2\pi}\int_{S^1}\left(-\overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+\right) ds. \tag{11.117}$$

We call these the vector and axial fermion numbers. The action is invariant under the space-time translations. We find the conserved currents

(11.118)
$$\begin{cases} T^t_{\ t} = -i\overline{\psi}_-\partial_s\psi_- + i\overline{\psi}_+\partial_s\psi_+, \\ T^s_{\ t} = i\overline{\psi}_-\partial_t\psi_- - i\overline{\psi}_+\partial_t\psi_+, \end{cases} \quad \begin{cases} T^t_{\ s} = i\overline{\psi}_-\partial_s\psi_- + i\overline{\psi}_+\partial_s\psi_+, \\ T^s_{\ s} = -i\overline{\psi}_-\partial_t\psi_- - i\overline{\psi}_+\partial_t\psi_+, \end{cases}$$

and the conserved charges

$$H = \frac{1}{2\pi}\int_{S^1}\left(-i\overline{\psi}_-\partial_s\psi_- + i\overline{\psi}_+\partial_s\psi_+\right) ds, \tag{11.119}$$

$$P = \frac{1}{2\pi}\int_{S^1}\left(i\overline{\psi}_-\partial_s\psi_- + i\overline{\psi}_+\partial_s\psi_+\right) ds. \tag{11.120}$$

Let us now expand the fields in the Fourier modes on S^1. We notice at this stage that we have not specified the boundary condition on S^1. We consider here a periodic boundary condition for both ψ_+ and ψ_-. (Other choices will be considered separately.) Then the fields are expanded as

$$\psi_- = \sum_{n\in\mathbb{Z}}\psi_n(t)\,e^{ins},\ \overline{\psi}_- = \sum_{n\in\mathbb{Z}}\overline{\psi}_n(t)\,e^{ins}, \tag{11.121}$$

$$\psi_+ = \sum_{n\in\mathbb{Z}}\widetilde{\psi}_n(t)\,e^{-ins},\ \overline{\psi}_+ = \sum_{n\in\mathbb{Z}}\overline{\widetilde{\psi}}_n(t)\,e^{-ins}. \tag{11.122}$$

Since $\overline{\psi}_\pm = \psi_\pm^\dagger$, the modes are related by

$$\overline{\psi}_n = \psi_{-n}^\dagger,\ \overline{\widetilde{\psi}}_n = \widetilde{\psi}_{-n}^\dagger. \tag{11.123}$$

In terms of these variables, the action is expressed as

$$S = \int \left[\sum_{n\in\mathbb{Z}} i\overline{\psi}_{-n}(\partial_t + in)\psi_n + \sum_{n\in\mathbb{Z}} i\overline{\widetilde{\psi}}_{-n}(\partial_t + in)\widetilde{\psi}_n\right] dt, \tag{11.124}$$

and we see that the system consists of infinitely many fermionic systems that are decoupled from each other.

11.3.1. Quantization. Let us quantize the system. From the form of the action, we find the anti-commutation relations

$$\{\psi_n, \overline{\psi}_m\} = \delta_{n+m,0},\ \{\widetilde{\psi}_n, \overline{\widetilde{\psi}}_m\} = \delta_{n+m,0}, \tag{11.125}$$

with all other anti-commutators vanishing. For each n, the algebra of $\psi_n, \overline{\psi}_{-n}$ is represented in a two-state vector space. As in the case of the free boson, we construct the total Hilbert space based on the product of the ground states of the constituent theories. We can read off from the action given in Eq. (11.124) that the Hamiltonian for the $\psi_n, \overline{\psi}_{-n}$ sector is given by

$$H_{n(+)} = n\overline{\psi}_{-n}\psi_n. \tag{11.126}$$

The ground state $|0\rangle_n$ is the one with $\psi_n|0\rangle_n = 0$ for $n > 0$ and $\overline{\psi}_{-n}|0\rangle_n = 0$ for $n < 0$. For $n = 0$, both of the states have the same energy and we choose one of them, say the one with $\overline{\psi}_0|0\rangle_0 = 0$. On the other hand, the Hamiltonian for the $\widetilde{\psi}_n, \overline{\widetilde{\psi}}_{-n}$ sector is given by

$$H_{n(-)} = n\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n. \tag{11.127}$$

The ground state $\widetilde{|0\rangle}_n$ is the one with $\widetilde{\psi}_n|0\rangle_n = 0$ for $n > 0$ and $\overline{\widetilde{\psi}}_{-n}|0\rangle_n = 0$ for $n < 0$. For $n = 0$ both states have the same energy and we choose one of them, say the one with $\widetilde{\psi}_0|0\rangle_0 = 0$. Thus, we define a state $|0\rangle$ of the total Hilbert space as the tensor product of these states

$$|0\rangle = \bigotimes_{n\in\mathbb{Z}^+} |0\rangle_n \otimes \widetilde{|0\rangle}_n. \tag{11.128}$$

This state is annihilated by the positive frequency modes

$$\psi_n|0\rangle = \overline{\psi}_n|0\rangle = \widetilde{\psi}_n|0\rangle = \overline{\widetilde{\psi}}_n|0\rangle = 0,\ n = 1, 2, 3, \ldots, \tag{11.129}$$

and also (by the choice made above)

$$\overline{\psi}_0|0\rangle = \widetilde{\psi}_0|0\rangle = 0. \tag{11.130}$$

Then the Hamiltonian is expressed as

$$\begin{aligned} H &= \sum_{n\in\mathbb{Z}} \left(n\overline{\psi}_{-n}\psi_n + n\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n\right) \\ &= \sum_{n=1}^{\infty} \left(n\overline{\psi}_{-n}\psi_n - n(-\psi_{-n}\overline{\psi}_n + 1) + n\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n - n(-\widetilde{\psi}_{-n}\overline{\widetilde{\psi}}_n + 1)\right) \\ &= \sum_{n\in\mathbb{Z}} \left(n{:}\overline{\psi}_{-n}\psi_n{:} + n{:}\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n{:}\right) + \frac{1}{6}, \end{aligned} \tag{11.131}$$

where ${:}\overline{\psi}_{-n}\psi_n{:}$ is a short hand notation for

$$\begin{cases} \overline{\psi}_{-n}\psi_n & n > 0, \\ -\psi_n\overline{\psi}_{-n} & n < 0, \end{cases} \tag{11.132}$$

and we have used the zeta function regularization to obtain $\sum_{n=1}^{\infty}(-2n) = \frac{1}{6}$. The ground state energy is

$$E_0 = \frac{1}{6}. \tag{11.133}$$

Note that the ground states are degenerate; the four states below are all ground states with energy $E_0 = \frac{1}{6}$;

$$\begin{array}{ccc} & \psi_0\overline{\widetilde{\psi}}_0|0\rangle & \\ \psi_0|0\rangle & & \overline{\widetilde{\psi}}_0|0\rangle \\ & |0\rangle. & \end{array} \tag{11.134}$$

The expression for the momentum is easier to obtain:

$$P = \sum_{n\in\mathbb{Z}} \left(-n{:}\overline{\psi}_{-n}\psi_n{:} + n{:}\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n{:}\right). \tag{11.135}$$

The vector and axial fermion numbers can be expressed as

$$
\begin{aligned}
F_V &= \sum_{n\in\mathbb{Z}} \left(\overline{\psi}_{-n}\psi_n + \overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n\right) \\
&= \overline{\psi}_0\psi_0 + \overline{\widetilde{\psi}}_0\widetilde{\psi}_0 + \sum_{n\neq 0}\left(:\overline{\psi}_{-n}\psi_n: + :\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n:\right) - 1
\end{aligned}
\tag{11.136}
$$

$$
\begin{aligned}
F_A &= \sum_{n\in\mathbb{Z}} \left(-\overline{\psi}_{-n}\psi_n + \overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n\right) \\
&= -\overline{\psi}_0\psi_0 + \overline{\widetilde{\psi}}_0\widetilde{\psi}_0 + \sum_{n\neq 0}\left(-:\overline{\psi}_{-n}\psi_n: + :\overline{\widetilde{\psi}}_{-n}\widetilde{\psi}_n:\right)
\end{aligned}
\tag{11.137}
$$

where the term of -1 in Eq. (11.136) comes from the sum $\sum_{n=1}^{\infty} 2 = -1$ which is again obtained by zeta function regularization. It is straightforward to compute these charges for the four ground states in Eq. (11.134). The result is

$$
F_V : \begin{array}{ccc} & 0 & \\ -1 & & 1 \\ & 0 & \end{array}, \quad F_A : \begin{array}{ccc} & 1 & \\ 0 & & 0 \\ & -1 & \end{array}. \tag{11.138}
$$

It is straightforward to see that

$$
[H, \psi_n] = -n\psi_n, \quad [H, \overline{\psi}_n] = -n\overline{\psi}_n, \tag{11.139}
$$

$$
[H, \widetilde{\psi}_n] = -n\widetilde{\psi}_n, \quad [H, \overline{\widetilde{\psi}}_n] = -n\overline{\widetilde{\psi}}_n. \tag{11.140}
$$

Applying these to Eqs. (11.121)– (11.122), we find

$$
\psi_- = \sum_{n\in\mathbb{Z}} \psi_n\, \mathrm{e}^{-in(t-s)}, \quad \overline{\psi}_- = \sum_{n\in\mathbb{Z}} \overline{\psi}_n\, \mathrm{e}^{-in(t-s)}, \tag{11.141}
$$

$$
\psi_+ = \sum_{n\in\mathbb{Z}} \widetilde{\psi}_n\, \mathrm{e}^{-in(t+s)}, \quad \overline{\psi}_+ = \sum_{n\in\mathbb{Z}} \overline{\widetilde{\psi}}_n\, \mathrm{e}^{-in(t+s)}. \tag{11.142}
$$

This shows that ψ_- and $\overline{\psi}_-$ are indeed right-moving fields and ψ_+ and $\overline{\psi}_+$ are left-moving fields.

11.3.2. Dirac's Sea. By construction, the state $|0\rangle$ is a lowest energy state and is therefore stable. There is a useful and insightful reinterpretation of this fact. (In this discussion, n stands for a *positive* integer.) The first of the commutation relations as shown in Eq. (11.139) can be interpreted as follows: ψ_{-n} $(n > 0)$ is the creation operator of a fermion e_n of positive energy n, whereas ψ_n is the creation operator of a fermion e_{-n} of negative energy, $-n$. The fact that $\psi_n|0\rangle = 0$ can then be interpreted as saying that *the state $|0\rangle$ is filled with the negative energy fermions, e_{-n}*. By fermion statistics (or Pauli's exclusion principle), a fermionic state cannot be occupied by two or more particles. Acting by $\overline{\psi}_{-n} = \psi_n^\dagger$ on $|0\rangle$ can be interpreted as removing the negative energy fermion e_{-n} or as creating a *hole*. This hole can further be interpreted as a positive energy *anti-fermion $\overline{e}_n$* of opposite charge for F_V and F_A. On the other hand, the fact that $\psi_{-n}|0\rangle \neq 0$ simply means that *the state $|0\rangle$ is not filled with the positive energy fermion e_n*. Similarly for the left-moving modes. Here we interpret the state $|0\rangle$ as occupied by the negative energy fermions, $\overline{\tilde{e}}_{-n}$ (with creation operator $\overline{\widetilde{\psi}}_n$), and empty for the positive energy fermions, $\overline{\tilde{e}}_n$ (with creation operator $\overline{\widetilde{\psi}}_{-n}$). Acting by $\psi_{-n} = (\overline{\widetilde{\psi}}_n)^\dagger$ removes $\overline{\tilde{e}}_{-n}$, creating a hole or the anti-particle $\tilde{e}_n$ of positive energy and opposite charge.

From this point of view, the state $|0\rangle$ can be interpreted as

$$
|0\rangle = \left(\prod_{n=1}^{\infty} \psi_n \overline{\widetilde{\psi}}_n\right) |0\rangle', \tag{11.143}
$$

where $|0\rangle'$ is the state that is empty for all negative and positive energy fermions. The state $|0\rangle$ is filled with all the negative energy particles and is therefore stable. One can make a hole but that costs positive energy, or it can be interpreted as creation of a positive energy anti-particle.

This point of view is due to P. A. M. Dirac and has many applications in various fields. (We will shortly encounter one of them.) The state $|0\rangle$ filled with negative energy states is called Dirac's sea, and the point of view that an anti-particle is considered as a hole is called Dirac's hole theory.

11.3.3. Twisted Boundary Conditions. As promised, we consider here the case where the fields are not periodic but obey the twisted boundary conditions

$$
\psi_-(t, s + 2\pi) = \mathrm{e}^{2\pi i a}\psi_-(t, s), \tag{11.144}
$$

$$
\psi_+(t, s + 2\pi) = \mathrm{e}^{-2\pi i \widetilde{a}}\psi_+(t, s). \tag{11.145}
$$

The periodicity of $\overline{\psi}_\pm$ follows from these condition by complex conjugation. The redefined fields $\psi'_-(t,s) = \mathrm{e}^{-ias}\psi_-(t,s)$ and $\psi'_+(t,s) = \mathrm{e}^{i\widetilde{a}s}\psi_+(t,s)$ are periodic, but the action for them (obtained by inserting $\psi_\pm$ into Eq. (11.110)) is

$$
S = \frac{1}{2\pi}\int_\Sigma \left(i\overline{\psi}'_-(\partial_t + \partial_s + ia)\psi'_- + i\overline{\psi}'_+(\partial_t - \partial_s + i\widetilde{a})\psi'_+\right) dt\, ds. \tag{11.146}
$$

This is the action for a Dirac fermion coupled to flat $U(1)$ gauge fields on S^1, with holonomies $\mathrm{e}^{2\pi i a}$ and $\mathrm{e}^{-2\pi i \widetilde{a}}$ for the right- and the left-movers respectively. Thus, twisting the boundary condition is equivalent to coupling to flat gauge fields (without changing the boundary condition).

The fields obeying the twisted boundary condition are expanded as

$$\psi_- = \sum_{r\in\mathbb{Z}+a} \psi_r(t)\,\mathrm{e}^{irs}, \quad \overline{\psi}_- = \sum_{r'\in\mathbb{Z}-a} \overline{\psi}_{r'}(t)\,\mathrm{e}^{ir's}, \tag{11.147}$$

$$\psi_+ = \sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}} \widetilde{\psi}_{\widetilde{r}}(t)\,\mathrm{e}^{-i\widetilde{r}s}, \quad \overline{\psi}_+ = \sum_{\widetilde{r}'\in\mathbb{Z}-\widetilde{a}} \overline{\widetilde{\psi}}_{\widetilde{r}'}(t)\,\mathrm{e}^{-i\widetilde{r}'s} \tag{11.148}$$

where $\psi_r^\dagger = \overline{\psi}_{-r}$ and $\widetilde{\psi}_{\widetilde{r}}^\dagger = \overline{\widetilde{\psi}}_{-\widetilde{r}}$. In terms of these modes, the action is written as

$$S = \int \left[\sum_{r\in\mathbb{Z}+a} i\overline{\psi}_{-r}(\partial_t + ir)\psi_r + \sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}} i\overline{\widetilde{\psi}}_{-\widetilde{r}}(\partial_t + i\widetilde{r})\widetilde{\psi}_{\widetilde{r}}\right] dt. \tag{11.149}$$

Quantization of the system proceeds as before, starting with

$$\{\psi_r, \overline{\psi}_{r'}\} = \delta_{r+r',0}, \quad \{\widetilde{\psi}_r, \overline{\widetilde{\psi}}_{\widetilde{r}'}\} = \delta_{\widetilde{r}+\widetilde{r}',0}. \tag{11.150}$$

The Hamiltonian is given by

$$H = \sum_{r\in\mathbb{Z}+a} r\,\overline{\psi}_{-r}\psi_r + \sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}} \widetilde{r}\,\overline{\widetilde{\psi}}_{-\widetilde{r}}\widetilde{\psi}_{\widetilde{r}}. \tag{11.151}$$

The state $|0\rangle_{a,\widetilde{a}}$ annihilated by

$$\psi_r\ (r\geq 0), \quad \overline{\psi}_{r'}\ (r'>0), \quad \widetilde{\psi}_{\widetilde{r}}\ (\widetilde{r}\geq 0), \quad \overline{\widetilde{\psi}}_{\widetilde{r}'}\ (\widetilde{r}'>0), \tag{11.152}$$

is a ground state. It is the unique ground state if $a \neq 0$ and $\widetilde{a} \neq 0$, but there are other ground state(s) if $a = 0$ or $\widetilde{a} = 0$. (For the case $a = \widetilde{a} = 0$ — the periodic boundary condition we studied earlier — the state $|0\rangle_{0,0}$ is equal to $\psi_0|0\rangle$ among the four ground states, as shown by Eq. (11.134).) The ground state energy is given by

$$E_0(a,\widetilde{a}) = \sum_{\substack{r\in\mathbb{Z}+a\\ r<0}} r + \sum_{\substack{\widetilde{r}\in\mathbb{Z}+\widetilde{a}\\ \widetilde{r}<0}} \widetilde{r}. \tag{11.153}$$

To evaluate this, we define the zeta function $\zeta(s,x) = \sum_{n=0}^{\infty}(n+x)^{-s}$ by analytic continuation from the region $\mathrm{Re}(s) > 1$. It is known that (see

Appendix 11.4)

$$\zeta(-1,x) = \frac{1}{24} - \frac{1}{2}\left(x - \frac{1}{2}\right)^2, \tag{11.154}$$

$$\zeta(0,x) = -x + \frac{1}{2}. \tag{11.155}$$

The ground state energy is $-\zeta(-1,1-a) - \zeta(-1,1-\widetilde{a}) = -\frac{1}{12} + \frac{1}{2}(\frac{1}{2}-a)^2 + \frac{1}{2}(\frac{1}{2}-\widetilde{a})^2$ if $0 < a < 1$ and $0 < \widetilde{a} < 1$. More generally it is

$$E_0(a,\widetilde{a}) = -\frac{1}{12} + \frac{1}{2}\left(a - [a] - \frac{1}{2}\right)^2 + \frac{1}{2}\left(\widetilde{a} - [\widetilde{a}] - \frac{1}{2}\right)^2. \tag{11.156}$$

In particular we find

$$E_0(0,0) = -\frac{1}{12} + \frac{1}{8} + \frac{1}{8} = \frac{1}{6}, \tag{11.157}$$

$$E_0(\tfrac{1}{2},\tfrac{1}{2}) = -\frac{1}{12}, \tag{11.158}$$

where the former recovers Eq. (11.133). As in Eq. (11.132), let us define

$$:\overline{\psi}_{-r}\psi_r: = \begin{cases} \overline{\psi}_{-r}\psi_r & r \geq 0, \\ -\psi_r\overline{\psi}_{-r} & r < 0, \end{cases} \quad :\overline{\widetilde{\psi}}_{-\widetilde{r}}\widetilde{\psi}_{\widetilde{r}}: = \begin{cases} \overline{\widetilde{\psi}}_{-\widetilde{r}}\widetilde{\psi}_{\widetilde{r}} & \widetilde{r} \geq 0, \\ -\widetilde{\psi}_{\widetilde{r}}\overline{\widetilde{\psi}}_{-\widetilde{r}} & \widetilde{r} < 0. \end{cases} \tag{11.159}$$

It is arranged so that they annihilate $|0\rangle_{a,\widetilde{a}}$. (We call such an operator ordering *the normal ordering with respect to the ground state* $|0\rangle_{a,\widetilde{a}}$.) Then the Hamiltonian H and momentum P are given by

$$H_R = \frac{1}{2}(H-P) = \sum_{r\in\mathbb{Z}+a} r\,:\overline{\psi}_{-r}\psi_r: + \frac{1}{2}\left(a - [a] - \frac{1}{2}\right)^2 - \frac{1}{24}, \tag{11.160}$$

$$H_L = \frac{1}{2}(H+P) = \sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}} \widetilde{r}\,:\overline{\widetilde{\psi}}_{-\widetilde{r}}\widetilde{\psi}_{\widetilde{r}}: + \frac{1}{2}\left(\widetilde{a} - [\widetilde{a}] - \frac{1}{2}\right)^2 - \frac{1}{24}. \tag{11.161}$$

The vector and axial fermion numbers are given by

$$F_R = \frac{1}{2}(F_V - F_A) = \sum_{r\in\mathbb{Z}+a} :\overline{\psi}_{-r}\psi_r: + a - [a] - \frac{1}{2}, \tag{11.162}$$

$$F_L = \frac{1}{2}(F_V + F_A) = \sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}} :\overline{\widetilde{\psi}}_{-\widetilde{r}}\widetilde{\psi}_{\widetilde{r}}: + \widetilde{a} - [\widetilde{a}] - \frac{1}{2}, \tag{11.163}$$

where we have used

$$\sum_{\substack{r\in\mathbb{Z}+a\\ r<0}} 1 = \zeta(0, 1-(a-[a])) = a - [a] - \frac{1}{2}. \tag{11.164}$$

At $a = 0$ or $\widetilde{a} = 0$, the ground state energy and momentum are not smooth and the ground state fermion numbers are not even continuous. This is not because the energy and momentum or the fermion numbers are non-smooth or discontinuous, but because the family of vacua $|0\rangle_{a,\widetilde{a}}$ is discontinuous at $a=0$ and $\widetilde{a}=0$. To see this, let us move a from small positive values to small negative values (we ignore $\widetilde{a}$ in the present discussion, say by fixing it at $\widetilde{a}=\frac{1}{2}$). For $a>0$, $|0\rangle_a$ is the unique ground state with (right-moving) energy $H_R=\frac{1}{2}(a-\frac{1}{2})^2-\frac{1}{24}$ and fermion number $F_R=a-\frac{1}{2}$. As a approaches 0 from above, another state $\overline{\psi}_{-a}|0\rangle_a$ comes close in energy but is separate in fermion number — it has $H_R=\frac{1}{2}(a+\frac{1}{2})^2-\frac{1}{24}$ and $F_R=a+\frac{1}{2}$. At $a=0$, the two have the same energy but different fermion numbers. As a is decreased below 0, the latter state becomes the unique ground state, which is newly denoted as $|0\rangle_a$. The flow of the energy and fermion number is depicted in Fig. 2. This flow is called the *spectral flow*.

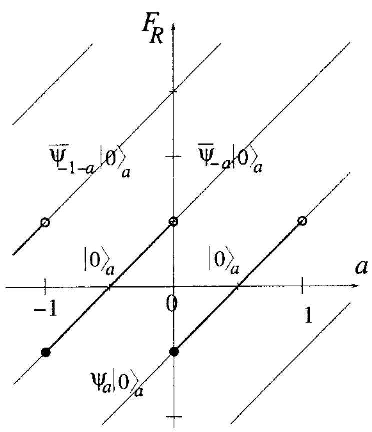


FIGURE 2. The Spectral Flow

The fermion obeying periodic (resp. anti-periodic) boundary condition is said to be in the *Ramond* sector (resp. *Neveu–Schwarz* sector), often abbreviated as R-sector or NS-sector. For example, the Dirac fermions with $(a,\widetilde{a})=(0,0),(0,\frac{1}{2}),(\frac{1}{2},0)$ and $(\frac{1}{2},\frac{1}{2})$ are in R-R, R-NS, NS-R, and NS-NS sectors respectively. These boundary conditions are allowed also for *Majorana fermions*, i.e., fermions constrained by the reality condition $\overline{\psi}_\pm=\psi_\pm$.

11.3.4. Partition Functions. We compute here the torus partition function of the system. We consider the torus of modular parameter

$\tau=\tau_1+i\tau_2$, namely, the space of coordinate $\zeta=(s+it)/2\pi$ with the identification $\zeta\equiv\zeta+1\equiv\zeta+\tau$ (t is now the Euclidean time). We assume that the fields obey the boundary conditions

$$\psi_-(t,s)=\mathrm{e}^{-2\pi ia}\psi_-(t,s+2\pi)=\mathrm{e}^{-2\pi ib}\psi_-(t+2\pi\tau_2,s+2\pi\tau_1),\tag{11.165}$$

$$\psi_+(t,s)=\mathrm{e}^{2\pi i\widetilde{a}}\psi_+(t,s+2\pi)=\mathrm{e}^{2\pi i\widetilde{b}}\psi_+(t+2\pi\tau_2,s+2\pi\tau_1).\tag{11.166}$$

Such a system corresponds to the periodic Dirac fermion on the torus, whose right- and left-movers are coupled to the flat gauge potentials

$$A^{0,1}=2\pi i\frac{b-\tau a}{2\tau_2}d\overline{\zeta},\quad \widetilde{A}^{1,0}=2\pi i\frac{\widetilde{b}-\overline{\tau}\widetilde{a}}{2\tau_2}d\zeta.\tag{11.167}$$

The partition function is represented as a trace in the space of states. Looking at the periodicity in $s\to s+2\pi$, we see that we can use the Hilbert space and operators developed in the previous section. The Euclidean time evolution $t\to t+2\pi\tau_2$, represented by the operator $\mathrm{e}^{-2\pi\tau_2H}$, induces the space translation $s\to s-2\pi\tau_1$ represented by $\mathrm{e}^{-2\pi i\tau_1P}$, together with the phase rotation of the fields represented by $\mathrm{e}^{-2\pi ibF_R+2\pi i\widetilde{b}F_L}$. Thus, we see that the partition function is represented as

$$Z=\mathrm{Tr}\left(\mathrm{e}^{-2\pi i(b-\frac{1}{2})F_R+2\pi i(\widetilde{b}-\frac{1}{2})F_L}\,\mathrm{e}^{-2\pi i\tau_1P}\,\mathrm{e}^{-2\pi\tau_2H}\right),\tag{11.168}$$

where the shift of b by $1/2$ is the standard one associated with the anti-commutativity of the fermions. We recall that the eigenvalues of F_R and F_L are respectively $a-1/2$ and $\widetilde{a}-1/2$ modulo integers. Thus, the partition function is periodic under integer shifts of $b,\widetilde{b}$ if we require $(\widetilde{a},\widetilde{b})=\pm(a,b)$. Such is the case when $\widetilde{A}^{1,0}=\mp(A^{0,1})^\dagger$, namely, when the system can be considered as a periodic Dirac fermion ψ with the left- and the right-movers coupled the same flat connection,

$$\begin{aligned}S&=\frac{1}{2\pi}\int i\overline{\psi}\gamma^\mu(\partial_\mu+iA_\mu)\psi\,dsd\tau,\\ A&=\frac{\pi i}{\tau_2}\left[(b-\tau a)d\overline{\zeta}-(b-\overline{\tau}a)d\zeta\right].\end{aligned}\tag{11.169}$$

Since $\mathrm{e}^{-2\pi i\tau_1P}\,\mathrm{e}^{-2\pi\tau_2H}=q^{H_R}\overline{q}^{H_L}$, the partition function has a left-right factorized form. Let us assume $0\le a<1$ (or replace a by $a-[a]$) and let us put $a':=a-1/2$ and $b':=b-1/2$. The right-moving part can be computed as

(11.170)
$$\begin{aligned}Z^R_{[a,b]}(\tau) =&q^{-\frac{1}{24}+\frac{1}{2}(a')^2}\,\mathrm{e}^{-2\pi i b'a'}\prod_{r\in\mathbb{Z}+a}\mathrm{Tr}_{\mathcal{H}_r} q^{r:\overline{\psi}_{-r}\psi_r:}\,\mathrm{e}^{-2\pi i(b-\frac{1}{2}):\overline{\psi}_{-r}\psi_r:}\\ =&q^{-\frac{1}{24}+\frac{1}{2}(a')^2}\,\mathrm{e}^{-2\pi i b'a'}\prod_{n=1}^{\infty}(1-q^{n-1+a}\,\mathrm{e}^{-2\pi i b})(1-q^{n-a}\,\mathrm{e}^{2\pi i b})\\ =&-q^{\frac{a^2}{2}}\,\mathrm{e}^{-2\pi i b'a}\frac{\vartheta\left[\begin{smallmatrix}1/2\\1/2\end{smallmatrix}\right](b-\tau a,\tau)}{\eta(\tau)}\\ =&\frac{\vartheta\left[\begin{smallmatrix}a'\\-b'\end{smallmatrix}\right](0,\tau)}{\eta(\tau)},\end{aligned}$$

where

(11.171) $$\vartheta\left[\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}\right](v,\tau) := \sum_{n\in\mathbb{Z}} q^{\frac{1}{2}(n+\alpha)^2}\,\mathrm{e}^{2\pi i(v+\beta)(n+\alpha)}.$$

See Appendix 11.4 for some properties of the theta functions. One can show the property

(11.172) $$Z^R_{[a,b]} = Z^R_{[a+1,b]} = -\mathrm{e}^{2\pi i a} Z^R_{[a,b+1]} = Z^R_{[-a,1-b]}.$$

For the case $(\widetilde{a},\widetilde{b}) = (a,b)$ or $(\widetilde{a},\widetilde{b}) = (-a,1-b)$ which realizes the system given by Eq. (11.169), the full partition function is

(11.173) $$Z_{[a,b]}(\tau,\overline{\tau}) = \left|Z^R_{[a,b]}(\tau)\right|^2.$$

From the properties in Eq. (11.172), it is indeed periodic under $a\to a+1$ and $b\to b+1$. The partition function has to be independent of the choice of the coordinates. Note that the coordinate transformations inducing $\tau\to\tau+1$ and $\tau\to -1/\tau$ are accompanied by the transformations of the holonomy $(a,b)\to(a,b+a)$ and $(a,b)\to(b,-a)$ respectively. One can show that the functions $Z^R_{[a,b]}(\tau)$ obey the modular transformation properties

(11.174) $$Z^R_{[a,b+a]}(\tau+1) = \mathrm{e}^{-\pi i(a^2-1/6)} Z^R_{[a,b]}(\tau),$$

(11.175) $$Z^R_{[b,-a]}(-1/\tau) = \mathrm{e}^{-2\pi i(-a)'b'} Z^R_{[a,b]}(\tau).$$

This shows that the full partition function is indeed invariant under the transformations $(\tau,a,b)\to(\tau+1,a,b+a)$ and $(\tau,a,b)\to(-1/\tau,b,-a)$.

Let us consider the case $a(=\widetilde{a})=0$, which corresponds to the periodic Dirac fermion on S^1. The ordinary partition function is Eq. (11.173) with $b(=\widetilde{b}) = 1/2$ while the one for $b(=\widetilde{b})=0$ corresponds to $\mathrm{Tr}(-1)^F q^{H_R}\overline{q}^{H_L}$, where $(-1)^F = \mathrm{e}^{\pi i F_A}$ which is also $-\mathrm{e}^{\pi i F_V}$ if $a=\widetilde{a}=0$. (Note that the eigenvalues of F_A and F_V are integers if $a=\widetilde{a}=0$.) We find from the second line of Eq. (11.170) that

(11.176) $$\mathrm{Tr}\, q^{H_R}\overline{q}^{H_L} = 4(q\overline{q})^{1/12}\prod_{n=1}^{\infty}(1+q^n)^2(1+\overline{q}^n)^2,$$

(11.177) $$\mathrm{Tr}\,(-1)^F q^{H_R}\overline{q}^{H_L} = (1-1)^2(q\overline{q})^{1/12}\prod_{n=1}^{\infty}(1-q^n)^2(1-\overline{q})^2 = 0.$$

The q-expansion of the partition function starts with $4(q\overline{q})^{1/12}$, reflecting the fact that there are four ground states given by Eq. (11.134) with energy $E_0 = 1/6$. The vanishing of $\mathrm{Tr}(-1)^F q^{H_R}\overline{q}^{H_L}$ is because two of them are $(-1)^F$ even and two of them are $(-1)^F$ odd, as shown by Eq. (11.138).

11.3.5. Boson–Fermion Equivalence. The partition functions for the special values $(a,b) = (\widetilde{a},\widetilde{b}) = (0,0),(0,\frac{1}{2}),(\frac{1}{2},0)$ and $(\frac{1}{2},\frac{1}{2})$ are given by

(11.178) $$Z_{[0,0]} = \frac{1}{|\eta(\tau)|^2}\left|\sum_{n\in\mathbb{Z}}(-1)^n q^{\frac{1}{2}(n-\frac{1}{2})^2}\right|^2 = 0,$$

(11.179) $$Z_{[0,\frac{1}{2}]} = \frac{1}{|\eta(\tau)|^2}\left|\sum_{n\in\mathbb{Z}} q^{\frac{1}{2}(n-\frac{1}{2})^2}\right|^2,$$

(11.180) $$Z_{[\frac{1}{2},0]} = \frac{1}{|\eta(\tau)|^2}\left|\sum_{n\in\mathbb{Z}}(-1)^n q^{\frac{1}{2}n^2}\right|^2,$$

(11.181) $$Z_{[\frac{1}{2},\frac{1}{2}]} = \frac{1}{|\eta(\tau)|^2}\left|\sum_{n\in\mathbb{Z}} q^{\frac{1}{2}n^2}\right|^2.$$

Since this set of $(a,b) = (\widetilde{a},\widetilde{b})$ is invariant under $(a,b)\to(a,b+a)$ and $(a,b)\to(b,-a)$, the sum of the above partition functions is invariant under the modular transformations $\tau\to\tau+1$ and $\tau\to -1/\tau$. This sum can be considered as the sum over the (left-right correlated) spin structures of the

Dirac fermion on the torus. One half of the sum can be expressed as

$$
\begin{aligned}
(11.182)\quad & \frac{1}{2}\left(Z_{[0,0]}+Z_{[0,\frac{1}{2}]}+Z_{[\frac{1}{2},0]}+Z_{[\frac{1}{2},\frac{1}{2}]}\right)\\
&=\frac{1}{|\eta(\tau)|^2}\sum_{n+\tilde{n}\in 2\mathbb{Z}} q^{\frac{1}{2}(n-\frac{1}{2})^2}\overline{q}^{\frac{1}{2}(\tilde{n}-\frac{1}{2})^2}+\frac{1}{|\eta(\tau)|^2}\sum_{n+\tilde{n}\in 2\mathbb{Z}} q^{\frac{1}{2}n^2}\overline{q}^{\frac{1}{2}\tilde{n}^2}\\
&=\frac{1}{|\eta(\tau)|^2}\sum_{(l,m)\in\mathbb{Z}\oplus\mathbb{Z}} q^{\frac{1}{2}(\frac{l}{2}-m)^2}\overline{q}^{\frac{1}{2}(\frac{l}{2}+m)^2}
\end{aligned}
$$

Comparing with Eq. (11.79), we find that this is nothing but the partition function for the sigma model on the torus of radius $R=\sqrt{2}$ (or $R=1/\sqrt{2}$ by $R\to 1/R$ duality). Namely

$$
(11.183)\qquad \frac{1}{2}\left(Z_{[0,0]}+Z_{[0,\frac{1}{2}]}+Z_{[\frac{1}{2},0]}+Z_{[\frac{1}{2},\frac{1}{2}]}\right)=Z(R=\sqrt{2}).
$$

We note that the first two and the last two terms can be identified as the following traces over the R-R and NS-NS sectors:

$$
(11.184)\qquad \frac{1}{2}\left(Z_{[0,0]}+Z_{[0,\frac{1}{2}]}\right)=\mathrm{Tr}_{\text{R-R}}\left[\left(\frac{(-1)^F+1}{2}\right)q^{H_R}\overline{q}^{H_L}\right],
$$

$$
(11.185)\qquad \frac{1}{2}\left(Z_{[\frac{1}{2},0]}+Z_{[\frac{1}{2},\frac{1}{2}]}\right)=\mathrm{Tr}_{\text{NS-NS}}\left[\left(\frac{(-1)^F+1}{2}\right)q^{H_R}\overline{q}^{H_L}\right].
$$

Here again, $(-1)^F=\mathrm{e}^{\pi i F_A}$ which is the same as $-\mathrm{e}^{\pi i F_V}$ on the R-R sector and $\mathrm{e}^{\pi i F_V}$ on the NS-NS sector. The operator $\frac{(-1)^F+1}{2}$ is a projection operator onto $(-1)^F$ even states, which we call a vector-like *GSO projection.* Thus, Eq. (11.182) can be considered as the partition function $\mathrm{Tr}\, q^{H_R}\overline{q}^{H_L}$ for the system of Dirac fermions where only $(-1)^F$ even states in the R-R and NS-NS sectors are kept. We call the latter system the Dirac fermion with vector-like GSO projection. Then Eq. (11.183) shows that *the Dirac fermion with vector-like GSO projection is equivalent to the sigma model on the circle of radius* $R=\sqrt{2}$.boson–fermion equivalence This is called *boson–fermion equivalence* which is a feature peculiar to $1+1$ dimensions. In fact boson–fermion equivalence holds in more general (interacting) theories [**56, 185**].

To see the correspondence in more detail, let us compute the partition function with weight $\mathrm{e}^{-2\pi i(bF_R-\tilde{b}F_L)}$ of the system with vector-like GSO projection. It is easy to find that

$$
\begin{aligned}
(11.186)\quad & \mathrm{Tr}\left(\mathrm{e}^{-2\pi i bF_R+2\pi i\tilde{b}F_L}q^{H_R}\overline{q}^{H_L}\right)\\
&=\frac{1}{|\eta(\tau)|^2}\sum_{(l,m)\in\mathbb{Z}\oplus\mathbb{Z}} q^{\frac{1}{2}(\frac{l}{2}-m)^2}\overline{q}^{\frac{1}{2}(\frac{l}{2}+m)^2}\,\mathrm{e}^{-2\pi i b(\frac{l}{2}-m)+2\pi i\tilde{b}(\frac{l}{2}+m)}.
\end{aligned}
$$

Comparing with the similar weighted partition function of the $R=\sqrt{2}$ sigma model on S^1, we find that the quantum numbers of the two theories are related as $F_R=p_R=\frac{1}{2}l-m$ and $F_L=p_L=\frac{1}{2}l+m$, where l and m are the momentum and the winding number of the target circle. In other words, we find

$$
(11.187)\qquad \begin{aligned} F_V&=l,\\ F_A&=2m.\end{aligned}
$$

Note that F_A has even eigenvalues because of the GSO projection. Since $(-1)^F=-\mathrm{e}^{\pi i F_V}$ on R-R and $(-1)^F=\mathrm{e}^{\pi i F_V}$ on NS-NS sectors, the R-R sector (resp. NS-NS sector) corresponds to odd (resp. even) momentum states of the S^1 sigma model with $R=\sqrt{2}$;

$$
(11.188)\qquad \mathcal{H}_{\text{R-R}}\cong\bigoplus_{\substack{l:\text{odd}\\ m\in\mathbb{Z}}}\mathcal{H}_{(l,m)},\qquad \mathcal{H}_{\text{NS-NS}}\cong\bigoplus_{\substack{l:\text{even}\\ m\in\mathbb{Z}}}\mathcal{H}_{(l,m)}.
$$

From Eq. (11.187), we find the correspondence between the conserved currents to be

$$
(11.189)\qquad \overline{\psi}_-\psi_-\longleftrightarrow\frac{1}{\sqrt{2}}(\partial_t-\partial_s)x,
$$

$$
(11.190)\qquad \overline{\psi}_+\psi_+\longleftrightarrow\frac{1}{\sqrt{2}}(\partial_t+\partial_s)x.
$$

Note that the fields $\psi_\pm$ and $\overline{\psi}_\pm$ by themselves are $(-1)^F$ odd and are not operators of the GSO projected theory, but the products of an even number of them (and their derivatives) are. The above currents $\overline{\psi}_\pm\psi_\pm$ are examples of such operators. The operator $\overline{\psi}_+\overline{\psi}_-$ has $(F_V,F_A/2)=(2,0)$ and thus creates two units of momentum while preserving the winding number. The field $\overline{\psi}_+\psi_-$ has $(F_V,F_A/2)=(0,1)$ and creates one unit of winding number while preserving the momentum. These suggest the correspondence

$$
(11.191)\qquad \overline{\psi}_+\overline{\psi}_-\longleftrightarrow\,:\mathrm{e}^{2ix/R}:\,=\,:\mathrm{e}^{\sqrt{2}ix}:,
$$

$$
(11.192)\qquad \overline{\psi}_+\psi_-\longleftrightarrow\,:\mathrm{e}^{i\tilde{x}/R^{-1}}:\,=\,:\mathrm{e}^{\sqrt{2}i\tilde{x}}:,
$$

where $\widehat{x}$ is the field of the T-dual theory (sigma model on S^1 of radius $1/R = 1/\sqrt{2}$). The vertex operators $:\mathrm{e}^{ilx/R}:$ with l odd exchange R-R and NS-NS sectors, and cannot be represented as the polynomials in the (derivatives of) $\psi_\pm$, $\overline{\psi}_\pm$. Such operators are called *spectral flow* operators since the mode expansion of the fields "flows" from the one with $r \in \mathbb{Z}$ to the one with $r \in \mathbb{Z} + \frac{1}{2}$ and vice versa.

11.4. Appendix

11.4.1. Zeta Functions. Let us define

$$\zeta(s, x) = \sum_{n=0}^{\infty} \frac{1}{(x+n)^s} \tag{11.193}$$

by analytic continuation from the region $\mathrm{Re}\,(s) > 1$ where the series is convergent. Riemann's zeta function is the special case $\zeta(s) = \zeta(s, 1)$. For the special values $s = -m = 0, -1, -2, \ldots$, it is given by

$$\zeta(-m, x) = -\frac{\phi'_{m+2}(x)}{(m+1)(m+2)} \tag{11.194}$$

where $\phi_n(x)$ are Bernoulli polynomials defined by

$$t\frac{\mathrm{e}^{xt} - 1}{\mathrm{e}^t - 1} = \sum_{n=1}^{\infty} \frac{t^n}{n!}\phi_n(x). \tag{11.195}$$

For example,

$$\zeta(0, x) = -\frac{\phi'_2(x)}{2} = -x + \frac{1}{2}, \tag{11.196}$$

$$\zeta(-1, x) = -\frac{\phi'_3(x)}{6} = -\frac{1}{2}\left(x^2 - x + \frac{1}{6}\right) = -\frac{1}{2}\left(x - \frac{1}{2}\right)^2 + \frac{1}{24}. \tag{11.197}$$

11.4.2. Theta Functions. Here we collect some properties of the theta functions. Let us define, for $q = \mathrm{e}^{2\pi i\tau}$ (with $\mathrm{Im}\,\tau > 0$),

$$\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right](v, \tau) := \sum_{n\in\mathbb{Z}} q^{\frac{1}{2}(n+\alpha)^2}\,\mathrm{e}^{2\pi i(v+\beta)(n+\alpha)}. \tag{11.198}$$

They have the periodicity $\vartheta\left[\begin{smallmatrix}\alpha+1\\ \beta\end{smallmatrix}\right] = \mathrm{e}^{-2\pi i\alpha}\vartheta\left[\begin{smallmatrix}\alpha\\ \beta+1\end{smallmatrix}\right] = \vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right]$. They also have periodicity in $v \to v + 1$ and $v + \tau$;

$$\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right](v + 1, \tau) = \mathrm{e}^{2\pi i\alpha}\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right](v, \tau), \tag{11.199}$$

$$\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right](v + \tau, \tau) = \mathrm{e}^{-2\pi i(v+\beta)}\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right](v, \tau). \tag{11.200}$$

Theta functions have the modular transformation property

$$\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right](v, \tau + 1) = \mathrm{e}^{-\pi i(\alpha^2+\alpha)}\vartheta\left[\begin{smallmatrix}\alpha\\ \beta+\alpha+\frac{1}{2}\end{smallmatrix}\right](v, \tau), \tag{11.201}$$

$$\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right]\left(\frac{v}{\tau}, -\frac{1}{\tau}\right) = (-i\tau)^{\frac{1}{2}}\,\mathrm{e}^{\pi i v^2/\tau + 2\pi i\alpha\beta}\vartheta\left[\begin{smallmatrix}\beta\\ -\alpha\end{smallmatrix}\right](v, \tau). \tag{11.202}$$

Theta functions have the following product formulae for the values $(\alpha, \beta) = (0, 0), (0, \frac{1}{2}), (\frac{1}{2}, 0)$ and $(\frac{1}{2}, \frac{1}{2})$:

$$\vartheta\left[\begin{smallmatrix}0\\ 0\end{smallmatrix}\right](v, \tau) = \prod_{n=1}^{\infty}(1 - q^n)(1 + zq^{n-\frac{1}{2}})(1 + z^{-1}q^{n-\frac{1}{2}}), \tag{11.203}$$

$$\vartheta\left[\begin{smallmatrix}0\\ 1/2\end{smallmatrix}\right](v, \tau) = \prod_{n=1}^{\infty}(1 - q^n)(1 - zq^{n-\frac{1}{2}})(1 - z^{-1}q^{n-\frac{1}{2}}), \tag{11.204}$$

$$\vartheta\left[\begin{smallmatrix}1/2\\ 1/2\end{smallmatrix}\right](v, \tau) = iq^{\frac{1}{8}}\,\mathrm{e}^{\pi i v}\prod_{n=1}^{\infty}(1 - q^n)(1 - zq^n)(1 - z^{-1}q^{n-1}), \tag{11.205}$$

$$\vartheta\left[\begin{smallmatrix}1/2\\ 0\end{smallmatrix}\right](v, \tau) = q^{\frac{1}{8}}\,\mathrm{e}^{\pi i v}\prod_{n=1}^{\infty}(1 - q^n)(1 + zq^n)(1 + z^{-1}q^{n-1}). \tag{11.206}$$

CHAPTER 12

$\mathcal{N} = (2,2)$ Supersymmetry

In our discussion of supersymmetric QFTs in dimensions 0 and 1, we have presented actions which possess fermionic symmetries. We did not present any systematic discussion of how one arrives at such actions. We will remedy this gap in this section and the next. We develop the notion of superspace which, in addition to the usual bosonic coordinates, contains fermionic coordinates (as many as the number of supersymmetries). We will also generalize the notion of fields to superfields. Supersymmetry is realized on the superspace by translations in the fermionic directions. Writing down actions which are coordinate invariant in the superspace sense will thus naturally lead to supersymmetric actions.

Here we will mainly consider supersymmetric field theories in 1+1 dimensions with four real supercharges (or two complex supercharges), two with positive chirality and two with negative chirality. This is called $\mathcal{N} = (2,2)$ supersymmetry, and is relevant for mirror symmetry. By reduction to 1 and 0 dimensions, one obtains the actions discussed in the previous sections for the case with four supercharges. One can also develop superspace techniques for the case with two supercharges. That will be recorded in Appendix 12.5.

12.1. Superfield Formalism

We start our discussion by providing a systematic way to obtain supersymmetric Lagrangians. This involves introducing superspace and superfields.

12.1.1. Superspace and Superfields. We consider a field theory on $\mathbb{R}^2$ with time and space coordinates

$$x^0 = t,\ x^1 = s. \tag{12.1}$$

We take the flat Minkowski metric $\eta_{00}=-1$, $\eta_{11}=1$ and $\eta_{01}=0$. Besides these bosonic coordinates we introduce four fermionic coordinates

$$\theta^+,\,\theta^-,\,\overline{\theta}^+,\,\overline{\theta}^-. \tag{12.2}$$

These are complex fermionic coordinates which are related to each other by complex conjugation, $(\theta^{\pm})^*=\overline{\theta}^{\pm}$. The indices $\pm$ stand for the spin (or chirality) under a Lorentz transformation. Namely, a Lorentz transformation acts on the bosonic and fermionic coordinates as

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \to \begin{pmatrix} \cosh\gamma & \sinh\gamma \\ \sinh\gamma & \cosh\gamma \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix}, \tag{12.3}$$

$$\theta^{\pm} \;\to\; \mathrm{e}^{\pm\gamma/2}\theta^{\pm},\;\overline{\theta}^{\pm} \;\to\; \mathrm{e}^{\pm\gamma/2}\overline{\theta}^{\pm}. \tag{12.4}$$

The fermionic coordinates anti-commute with each other, $\theta^\alpha\theta^\beta=-\theta^\beta\theta^\alpha$, $\overline{\theta}^\alpha\overline{\theta}^\beta=-\overline{\theta}^\beta\overline{\theta}^\alpha$, and $\theta^\alpha\overline{\theta}^\beta=-\overline{\theta}^\beta\theta^\alpha$. The $(2,2)$ *superspace* is the space with the coordinates x^0,x^1, $\theta^{\pm}$, $\overline{\theta}^{\pm}$.

Superfields are functions defined on the superspace. They can be Taylor expanded in monomials in $\theta^{\pm}$ and $\overline{\theta}^{\pm}$.

$$\begin{aligned}\mathcal{F}(x^0,x^1,\theta^+,\theta^-,\overline{\theta}^+,\overline{\theta}^-) =& f_0(x^0,x^1)+\theta^+f_+(x^0,x^1)\\ &+\theta^-f_-(x^0,x^1)+\overline{\theta}^+f'_+(x^0,x^1)\\ &+\overline{\theta}^-f'_-(x^0,x^1)+\theta^+\theta^-f_{+-}(x^0,x^1)+\cdots.\end{aligned} \tag{12.5}$$

(Superfields are to supersymmetry what N-vector fields are to $SO(N)$ symmetry – a convenient organizational scheme.) Since any of the fermionic coordinates squares to zero, $(\theta^{\pm})^2=(\overline{\theta}^{\pm})^2=0$, there are at most $2^4=16$ nonzero terms in the expansion. A superfield Φ is bosonic if $[\theta^\alpha,\Phi]=0$ and is fermionic if $\{\theta^\alpha,\Phi\}=0$. We introduce some differential operators on the superspace,

$$\mathcal{Q}_{\pm} \;=\; \frac{\partial}{\partial\theta^{\pm}}+i\overline{\theta}^{\pm}\partial_{\pm}, \tag{12.6}$$

$$\overline{\mathcal{Q}}_{\pm} \;=\; -\frac{\partial}{\partial\overline{\theta}^{\pm}}-i\theta^{\pm}\partial_{\pm}. \tag{12.7}$$

Here $\partial_{\pm}$ are differentiations by $x^{\pm}:=x^0\pm x^1$:

$$\partial_{\pm}=\frac{\partial}{\partial x^{\pm}}=\frac{1}{2}\left(\frac{\partial}{\partial x^0}\pm\frac{\partial}{\partial x^1}\right). \tag{12.8}$$

These differential operators satisfy the anti-commutation relations

$$\{\mathcal{Q}_{\pm},\overline{\mathcal{Q}}_{\pm}\}=-2i\partial_{\pm}, \tag{12.9}$$

with all other anti-commutators vanishing. We define another set of differential operators

$$D_{\pm} \;=\; \frac{\partial}{\partial\theta^{\pm}}-i\overline{\theta}^{\pm}\partial_{\pm}, \tag{12.10}$$

$$\overline{D}_{\pm} \;=\; -\frac{\partial}{\partial\overline{\theta}^{\pm}}+i\theta^{\pm}\partial_{\pm}, \tag{12.11}$$

which anti-commute with $\mathcal{Q}_{\pm}$ and $\overline{\mathcal{Q}}_{\pm}$, i.e., $\{D_{\pm},\mathcal{Q}_{\pm}\}=0$, etc. These obey similar anti-commutation relations

$$\{D_{\pm},\overline{D}_{\pm}\}=2i\partial_{\pm}, \tag{12.12}$$

with all other anti-commutators vanishing.

EXERCISE 12.1.1. *We have discussed the notion of superspace adapted to the signature* $(1,1)$. *Generalize this to the Euclidean signature. In particular show that the* $\pm$ *index on* $x^{\pm}$ *and* $\theta^{\pm},\overline{\theta}^{\pm}$ *distinguishes holomorphic versus anti-holomorphic supercoordinates.*

Vector R-rotations and *axial R-rotations* of a superfield are defined by

$$\mathrm{e}^{i\alpha F_V}:\mathcal{F}(x^\mu,\theta^{\pm},\overline{\theta}^{\pm}) \;\mapsto\; \mathrm{e}^{i\alpha q_V}\mathcal{F}(x^\mu,\,\mathrm{e}^{-i\alpha}\theta^{\pm},\,\mathrm{e}^{i\alpha}\overline{\theta}^{\pm}), \tag{12.13}$$

$$\mathrm{e}^{i\beta F_A}:\mathcal{F}(x^\mu,\theta^{\pm},\overline{\theta}^{\pm}) \;\mapsto\; \mathrm{e}^{i\beta q_A}\mathcal{F}(x^\mu,\,\mathrm{e}^{\mp i\beta}\theta^{\pm},\,\mathrm{e}^{\pm i\beta}\overline{\theta}^{\pm}), \tag{12.14}$$

where q_V and q_A are numbers called vector R-charge and axial R-charge of $\mathcal{F}$. The transformations given by Eqs. (12.13)–(12.14) induce transformations of the constituent fields of $\mathcal{F}$.

A *chiral superfield* Φ is a superfield that satisfies the equations

$$\overline{D}_{\pm}\Phi=0. \tag{12.15}$$

If Φ_1 and Φ_2 are chiral superfields, the product $\Phi_1\Phi_2$ is also a chiral superfield. A general chiral superfield Φ has the form

$$\Phi(x^\mu,\theta^{\pm},\overline{\theta}^{\pm})=\phi(y^{\pm})+\theta^\alpha\psi_\alpha(y^{\pm})+\theta^+\theta^-F(y^{\pm}), \tag{12.16}$$

where $y^{\pm}=x^{\pm}-i\theta^{\pm}\overline{\theta}^{\pm}$. The complex conjugate of a chiral superfield Φ obeys the condition

$$D_{\pm}\overline{\Phi}=0 \tag{12.17}$$

and is called an *anti-chiral superfield.*

EXERCISE 12.1.2. *Show that a chiral superfield can be expanded as shown in Eq. (12.16).*

A *twisted chiral superfield* U is a superfield that satisfies

$$\overline{D}_+ U = D_- U = 0. \tag{12.18}$$

If U_1 and U_2 are twisted chiral superfields, the product $U_1 U_2$ is also a twisted chiral superfield. A general twisted chiral superfield U has the form

$$U(x^\mu, \theta^\pm, \overline{\theta}^\pm) = v(\widetilde{y}^\pm) + \theta^+ \overline{\chi}_+(\widetilde{y}^\pm) + \overline{\theta}^- \chi_-(\widetilde{y}^\pm) + \theta^+ \overline{\theta}^- E(\widetilde{y}^\pm), \tag{12.19}$$

where $\widetilde{y}^\pm = x^\pm \mp i\theta^\pm \overline{\theta}^\pm$. The complex conjugate $\overline{U}$ of a twisted chiral superfield U obeys the condition

$$D_+ \overline{U} = \overline{D}_- \overline{U} = 0 \tag{12.20}$$

and is called a *twisted anti-chiral superfield.*

12.1.2. Supersymmetric Actions. We now construct action functionals of superfields that are invariant under the transformation

$$\delta = \epsilon_+ \mathcal{Q}_- - \epsilon_- \mathcal{Q}_+ - \overline{\epsilon}_+ \overline{\mathcal{Q}}_- + \overline{\epsilon}_- \overline{\mathcal{Q}}_+. \tag{12.21}$$

Let us first consider the functional of the superfields $\mathcal{F}_i$ of the form

$$\int d^2x\, d^4\theta\, K(\mathcal{F}_i) = \int d^2x\, d\theta^+ d\theta^- d\overline{\theta}^- d\overline{\theta}^+\, K(\mathcal{F}_i), \tag{12.22}$$

where $K(-)$ is an arbitrary differentiable function of the $\mathcal{F}_i$'s. This is invariant under the variation δ. For example, let us look at the term proportional to ϵ_+;

$$\int d^2x d^4\theta\, \epsilon_+ (\mathcal{Q}_- \mathcal{F}_i) \frac{\partial K}{\partial \mathcal{F}_i} = \int d^2x d^4\theta\, \epsilon_- \left(\frac{\partial}{\partial \theta^-} + i\overline{\theta}^- \partial_- \right) K(\mathcal{F}_i). \tag{12.23}$$

The integration over $d^4\theta$ is nonzero only if we have $\theta^+\theta^-\overline{\theta}^+\overline{\theta}^-$. Therefore the first term is zero since the integrand does not have θ^- because of the derivative $\partial/\partial\theta^-$. The second term is a total derivative and vanishes after integration over d^2x. Vanishing of the coefficients of ϵ_+ and $\overline{\epsilon}_\pm$ can be seen in a similar way. The functional of the form shown in Eq. (12.22) is called a *D-term.*

We next consider the functional of chiral superfields Φ_i of the form

$$\int d^2x d^2\theta\, W(\Phi_i) = \int d^2x\, d\theta^- d\theta^+\, W(\Phi_i)\Big|_{\overline{\theta}^\pm = 0}, \tag{12.24}$$

where $W(\Phi_i)$ is a holomorphic function of the Φ_i's. This is also invariant under the variation δ. Let us first look at the coefficient of $\epsilon_\pm$:

$$\pm \int d^2x\, d\theta^- d\theta^+\, \epsilon_\pm \left(\frac{\partial}{\partial \theta^\mp} + i\overline{\theta}^\mp \partial_\mp \right) W(\Phi_i) \Bigg|_{\overline{\theta}^\pm = 0}. \tag{12.25}$$

The first term vanishes for the standard reason. The second term vanishes because we put $\overline{\theta}^\pm = 0$ (or since it is a total derivative). Let us next look at the coefficient of $\overline{\epsilon}_\pm$. For this we note that $\overline{\mathcal{Q}}_\pm = \overline{D}_\pm - 2i\theta^\pm \partial_\pm$. Then the variation is

$$\mp \int d^2x\, d\theta^- d\theta^+\, \overline{\epsilon}_\pm \left(\overline{D}_\mp - 2i\theta^\pm \partial_\pm \right) W(\Phi_i) \Bigg|_{\overline{\theta}^\pm = 0}. \tag{12.26}$$

The first term in the integrand $\overline{D}_\mp W(\Phi_i)$ is zero because Φ_i are chiral superfields and $W(\Phi_i)$ is a holomorphic function (it does not contain $\overline{\Phi}_i$). The second integral vanishes because it is a total derivative in x^μ. The functional of the form shown in Eq. (12.24) is called an *F-term.*

We finally consider the functional of twisted chiral superfields U_i of the form

$$\int d^2x\, d^2\widetilde{\theta}\, \widetilde{W}(U_i) = \int d^2x\, d\overline{\theta}^- d\theta^+\, \widetilde{W}(U_i)\Big|_{\overline{\theta}^+ = \theta^- = 0}, \tag{12.27}$$

where $\widetilde{W}(U_i)$ is a holomorphic function of the U_i's. By a similar argument as in the case of the F-term, one can see that this functional is invariant under δ. The functional of the form shown in Eq. (12.27) is called a *twisted F-term.*

12.1.3. Some Superfield Calculus. We present some calculus on superspace, some of which will be used in later sections. However, the reader can skip these exercises in the first reading and return when they are needed.

The basic element of the superfield calculus is the analogue of Poincaré's lemma in the ordinary calculus. Suppose $\mathcal{F}$ is a superfield that decays rapidly at infinity in (x^0, x^1)-space. Then

LEMMA 12.1.1 (Poincaré's Lemma). *$D_+\mathcal{F} = 0$ implies $\mathcal{F} = D_+\mathcal{G}$ for some superfield $\mathcal{G}$. The same is true for the differential operators D_-, $\overline{D}_+$ and $\overline{D}_-$.*

PROOF. $D_+\mathcal{F} = 0$ implies $\overline{D}_+ D_+ \mathcal{F} = 0$. Using the anti-commutation relation from Eq. (12.12), we find $2i\partial_+\mathcal{F} = D_+\overline{D}_+\mathcal{F}$. Since $\mathcal{F}$ decays rapidly at infinity, we can integrate this relation as

$$2i\mathcal{F} = \int_{-\infty}^{x^+} D_+\overline{D}_+\mathcal{F}\,dx'^+ = D_+\int_{-\infty}^{x^+}\overline{D}_+\mathcal{F}\,dx'^+. \tag{12.28}$$

We thus obtain $\mathcal{F} = D_+\mathcal{G}$ where $\mathcal{G}$ is the superfield $\frac{1}{2i}\int_{-\infty}^{x^+}\overline{D}_+\mathcal{F}dx'^+$. This is what we wanted to show. □

The reasoning used in the proof is sufficient to show the following.

1. $\overline{D}_+\overline{D}_-\mathcal{F} = 0$ implies $\mathcal{F} = \mathcal{G}_+ + \mathcal{G}_-$ for some superfields $\mathcal{G}_\pm$ such that $\overline{D}_+\mathcal{G}_+ = 0$ and $\overline{D}_-\mathcal{G}_- = 0$. Similar results hold for D_+D_-, $\overline{D}_+D_-$ and $D_+\overline{D}_-$.
2. A chiral superfield Φ can be written as $\Phi = \overline{D}_+\overline{D}_-\mathcal{E}$ for some superfield $\mathcal{E}$. If U is a twisted chiral superfield it can be written as $U = \overline{D}_+D_-\mathcal{V}$.
3. $\overline{D}_+\overline{D}_-\mathcal{F} = D_+D_-\mathcal{F} = 0$ implies $\mathcal{F} = U_1 + \overline{U_2}$ for some twisted chiral superfields U_i. For the equation $\overline{D}_+D_-\mathcal{F} = D_+\overline{D}_-\mathcal{F} = 0$, we have $\mathcal{F} = \Phi_1 + \overline{\Phi_2}$ for some chiral superfields Φ_i.

EXERCISE 12.1.3. *Prove the above statements.*

Let us consider the integral

$$\int d^2x\, d^4\theta AB \tag{12.29}$$

where A and B are arbitrary superfields. It is easy to see that the extremum of this integral with respect to the variation of A is attained only by $B=0$.

EXERCISE 12.1.4. *Show that if A is restricted to be a chiral superfield, then the extrema are attained by B with $\overline{D}_+\overline{D}_-B = 0$.*

12.2. Basic Examples

Here we present basic examples of classical $(2,2)$ supersymmetric field theories. One is a theory of a single chiral superfield and the other is a theory of a single twisted chiral superfield.

12.2.1. Theory of a Chiral Superfield. We first consider a supersymmetric action for a single chiral superfield Φ. As noted above, the superfield Φ has the following θ-expansion

$$\begin{aligned}\Phi &= \phi(y^\pm) + \theta^\alpha\psi_\alpha(y^\pm) + \theta^+\theta^-F(y^\pm)\\ &= \phi - i\theta^+\overline{\theta}^+\partial_+\phi - i\theta^-\overline{\theta}^-\partial_-\phi - \theta^+\theta^-\overline{\theta}^-\overline{\theta}^+\partial_+\partial_-\phi\\ &\quad +\theta^+\psi_+ - i\theta^+\theta^-\overline{\theta}^-\partial_-\psi_+ + \theta^-\psi_- - i\theta^-\theta^+\overline{\theta}^+\partial_+\psi_- + \theta^+\theta^-F,\end{aligned} \tag{12.30}$$

where in the last equality we have further expanded $y^\pm = x^\pm - i\theta^\pm\overline{\theta}^\pm$ at $x^\pm$. The θ-expansion of the anti-chiral superfield $\overline{\Phi}$ is easily obtained by complex conjugation of Eq. (12.30);

$$\begin{aligned}\overline{\Phi} &= \overline{\phi} + i\theta^+\overline{\theta}^+\partial_+\overline{\phi} + i\theta^-\overline{\theta}^-\partial_-\overline{\phi} - \theta^+\theta^-\overline{\theta}^-\overline{\theta}^+\partial_+\partial_-\overline{\phi}\\ &\quad -\overline{\theta}^+\overline{\psi}_+ - i\overline{\theta}^+\theta^-\overline{\theta}^-\partial_-\overline{\psi}_+ - \overline{\theta}^-\overline{\psi}_- - i\overline{\theta}^-\theta^+\overline{\theta}^+\partial_+\overline{\psi}_- + \overline{\theta}^-\overline{\theta}^+\overline{F}.\end{aligned} \tag{12.31}$$

Note that $(\psi_1\psi_2)^* = \psi_2^*\psi_1^*$ for fermionic variables/coordinates.

Now let us compute the D-term

$$S_{\rm kin} = \int d^2x\, d^4\theta\,\overline{\Phi}\Phi. \tag{12.32}$$

The integration $\int d^4\theta\,\overline{\Phi}\Phi$ amounts to extracting the coefficient of $\theta^4 = \theta^+\theta^-\overline{\theta}^-\overline{\theta}^+$ in the θ-expansion of $\overline{\Phi}\Phi$. By a straightforward computation we have

$$\begin{aligned}\overline{\Phi}\Phi\Big|_{\theta^4} &= -\overline{\phi}\partial_+\partial_-\phi + \partial_+\overline{\phi}\partial_-\phi + \partial_-\overline{\phi}\partial_+\phi - \partial_+\partial_-\overline{\phi}\phi\\ &\quad + i\overline{\psi}_+\partial_-\psi_+ - i\partial_-\overline{\psi}_+\psi_+ + i\overline{\psi}_-\partial_+\psi_- - i\partial_+\overline{\psi}_-\psi_- + |F|^2.\end{aligned} \tag{12.33}$$

Here again, the derivatives of fields appear due to the changing of variables from y to x and doing the Taylor expansion around $\theta = 0$. By partial integration, the action takes the form

$$S_{\rm kin} = \int d^2x\left(|\partial_0\phi|^2 - |\partial_1\phi|^2 + i\overline{\psi}_-(\partial_0+\partial_1)\psi_- + i\overline{\psi}_+(\partial_0-\partial_1)\psi_+ + |F|^2\right). \tag{12.34}$$

Thus, we have obtained the standard kinetic term for the complex scalar field ϕ and the Dirac fermion fields $\psi_\pm$, $\overline{\psi}_\pm$. Note also that the field F has no kinetic term. Such a field is often called an *auxiliary field* (such as in the path-integral derivation of T-duality). Next let us compute the F-term

$$S_W = \int d^2x\, d^2\theta\, W(\Phi) + c.c. \tag{12.35}$$

for a holomorphic function $W(\Phi)$ of Φ. This holomorphic function is called a *superpotential*. The integral $\int d^2\theta W(\Phi)$ amounts to extracting the coefficient of $\theta^2 = \theta^+\theta^-$ in the θ-expansion of $W(\Phi)$. It is straightforward to see

$$W(\Phi)\Big|_{\theta^2} = W'(\phi)F - W''(\phi)\psi_+\psi_-. \tag{12.36}$$

Thus, the F-term is

(12.37)
$$S_W = \int d^2x \left(W'(\phi)F - W''(\phi)\psi_+\psi_- + \overline{W}'(\overline{\phi})\overline{F} - \overline{W}''(\overline{\phi})\overline{\psi}_-\overline{\psi}_+ \right).$$

Now let us consider the sum of S_{kin} and S_W as the total action;

(12.38)
$$S = S_{\text{kin}} + S_W.$$

By completing the square of F, we obtain the following action

(12.39)
$$\begin{aligned} S = \int d^2x \Big(&|\partial_0\phi|^2 - |\partial_1\phi|^2 - |W'(\phi)|^2 + i\overline{\psi}_-(\partial_0+\partial_1)\psi_- \\ &+ i\overline{\psi}_+(\partial_0-\partial_1)\psi_+ - W''(\phi)\psi_+\psi_- \\ &- \overline{W}''(\overline{\phi})\overline{\psi}_-\overline{\psi}_+ + |F+\overline{W}'(\overline{\phi})|^2 \Big). \end{aligned}$$

Note that the last term $|F+\overline{W}'(\overline{\phi})|^2$ can be eliminated by solving the equation of motion as

(12.40)
$$F = -\overline{W}'(\overline{\phi}).$$

Setting F to this value can also be viewed as a result of integrating out F in the path-integral. To summarize, we have obtained the action for the scalar ϕ and the Dirac fermion $\psi_\pm$, $\overline{\psi}_\pm$ with a potential $|W'(\phi)|^2$ for ϕ and the fermion mass term (or Yukawa interaction) $W''(\phi)\psi_+\psi_-$.

By construction, the action is invariant under the variation δ from Eq. (12.21). This variation on the superfield Φ can actually be identified as a certain variation of the ordinary fields ϕ, $\psi_\pm$, $\overline{\psi}_\pm$ and F — the component fields of Φ. This is obvious if the superfield $\mathcal{F}$ is unconstrained. Simply *define* each coefficient field of the θ-expansion of $\delta\mathcal{F}$ as the variation of the corresponding coefficient field of the θ-expansion of $\mathcal{F}$. For example, for the general superfield given in Eq. (12.5), the δ-variation yields

(12.41)
$$\delta\mathcal{F} = \epsilon_+ f_- - \epsilon_- f_+ + \overline{\epsilon}_- f' + \overline{\epsilon}_+ f'_- + \theta^+(\cdots) + \cdots,$$

Then we *define* $\delta f_0 = \epsilon_+ f_- - \epsilon_- f_+ + \overline{\epsilon}_- f' + \overline{\epsilon}_+ f'_-$, $\delta f_+ = (\cdots)$, etc. A chiral superfield is not an arbitrary superfield but rather satisfies $\overline{D}_\pm\Phi = 0$. The last condition means that there are relations between the coefficient fields, as can be explicitly seen in Eq. (12.30). Thus, it is not obvious whether the variation δ of Φ can be represented by a variation of the component fields of Φ. However, this is actually the case. The key point is that the differential

operators $\mathcal{Q}_\pm$, $\overline{\mathcal{Q}}_\pm$ anti-commute with $\overline{D}_\pm$ (and also with $D_\pm$) and hence the variation $\delta\Phi$ is also a chiral superfield

(12.42)
$$\overline{D}_\pm\delta\Phi = \delta\overline{D}_\pm\Phi = 0.$$

Indeed, one can explicitly show that the variation in question is given by

(12.43)
$$\begin{aligned} \delta\phi &= \epsilon_+\psi_- - \epsilon_-\psi_+, \\ \delta\psi_\pm &= \pm 2i\overline{\epsilon}_\mp\partial_\pm\phi + \epsilon_\pm F, \\ \delta F &= -2i\overline{\epsilon}_+\partial_-\psi_+ - 2i\overline{\epsilon}_-\partial_+\psi_-. \end{aligned}$$

One can replace F by its equation of motion and write a supersymmetry variation of the ϕ and ψ fields alone (true after imposing the equations of motion). One can explicitly check (though it is not necessary) that the action S (or S_{kin} and S_W) is invariant under this variation of the component fields. By the anti-commutation relations from Eq. (12.9), the variations for different parameters ϵ_1 and ϵ_2 satisfy the commutation relation

(12.44)
$$[\delta_1, \delta_2] = 2i(\epsilon_{1-}\overline{\epsilon}_{2-} - \epsilon_{2-}\overline{\epsilon}_{1-})\partial_+ + 2i(\epsilon_{1+}\overline{\epsilon}_{2+} - \epsilon_{2+}\overline{\epsilon}_{1+})\partial_-.$$

This is a relation in quantum mechanics that generalizes the supersymmetry relation given by Eq. (10.77). We refer to this situation by saying *the classical field theory with the action given by Eq. (12.39) has $\mathcal{N}=(2,2)$ supersymmetry generated by Eq. (12.43).*

Since the classical system has a symmetry, one can find via the Noether procedure the conserved currents and conserved charges. The conserved currents are

(12.45)
$$G^0_\pm = 2\partial_\pm\overline{\phi}\,\psi_\pm \mp i\overline{\psi}_\mp\overline{W}'(\overline{\phi}),$$

(12.46)
$$G^1_\pm = \mp 2\partial_\pm\overline{\phi}\psi_\pm - i\overline{\psi}_\mp\overline{W}'(\overline{\phi}),$$

(12.47)
$$\overline{G}^0_\pm = 2\overline{\psi}_\pm\partial_\pm\phi \pm i\psi_\mp W'(\phi),$$

(12.48)
$$\overline{G}^1_\pm = \mp 2\overline{\psi}_\pm\partial_\pm\phi \pm i\psi_\mp W'(\phi),$$

and the conserved charges (supercharges) are

(12.49)
$$Q_\pm = \int dx^1\, G^0_\pm,\ \ \overline{Q}_\pm = \int dx^1\, \overline{G}^0_\pm.$$

These charges transform as spinors

(12.50)
$$Q_\pm \mapsto e^{\mp\gamma/2}Q_\pm,\ \ \overline{Q}_\pm \mapsto e^{\mp\gamma/2}\overline{Q}_\pm,$$

under Lorentz transformation as shown by Eq. (12.3).

EXERCISE 12.2.1. *Verify the expressions from Eq. (12.49) for the supercharges.*

This system has more global symmetries. First, by assigning axial R-charge 0 for Φ, the action is invariant under axial R-rotation;

$$\Phi(x^{\pm},\theta^{\pm},\overline{\theta}^{\pm}) \mapsto \Phi(x^{\pm},\, e^{\mp i\alpha}\theta^{\pm},\, e^{\pm i\alpha}\overline{\theta}^{\pm}). \tag{12.51}$$

This is obvious in the superspace expressions given by Eq. (12.32) and Eq. (12.35): the products θ^4 and θ^2 are both invariant under the axial rotation. Thus, the system has an axial R-symmetry. The axial rotation of the superfields can be realized as a transformation of the component fields (this can also be understood by looking at the commutation relation of $\overline{D}_{\pm}$ and the axial rotation). The transformation is given by

$$\phi \mapsto \phi,\ \psi_{\pm} \mapsto e^{\mp i\alpha}\psi_{\pm}. \tag{12.52}$$

The corresponding current is given by

$$J_A^0 = \overline{\psi}_+\psi_+ - \overline{\psi}_-\psi_-, \tag{12.53}$$

$$J_A^1 = -\overline{\psi}_+\psi_+ - \overline{\psi}_-\psi_-, \tag{12.54}$$

and the conserved charge is

$$F_A = \int J_A^0\, dx^1. \tag{12.55}$$

We note that the axial R-rotation rotates the supercharges as

$$Q_{\pm} \mapsto e^{\mp i\alpha}Q_{\pm},\ \overline{Q}_{\pm} \mapsto e^{\pm i\alpha}\overline{Q}_{\pm}. \tag{12.56}$$

Second, depending on the form of the superpotential $W(\Phi)$, the system is also invariant under the vector R-rotation. Since θ^4 is invariant under the vector R-rotation and $\overline{\Phi}\Phi$ is invariant under the phase rotation of Φ, the D-term is invariant under an arbitrary choice of vector R-charge. However, θ^2 has vector R-charge -2 (namely it transforms as $\theta^2 \mapsto \theta^2 e^{-2i\alpha}$). Thus, the F-term is invariant under vector R-rotation if and only if one can assign the vector R-charge of Φ so that $W(\Phi)$ has vector R-charge 2. This is the case when $W(\Phi)$ is a monomial. If

$$W(\Phi) = c\Phi^k, \tag{12.57}$$

then, by assigning vector R-charge $2/k$ to Φ, the F-term is made invariant under vector R-rotation. Namely, the system has a vector R-symmetry. In such a case, the vector R-rotation of the superfield is realized as a transformation of the component field as

$$\phi \mapsto e^{(2/k)i\alpha}\phi,\ \psi_{\pm} \mapsto e^{((2/k)-1)i\alpha}\psi_{\pm}. \tag{12.58}$$

The conserved current is

$$J_V^0 = (2i/k)(\partial_0\overline{\phi}\phi - \overline{\phi}\partial_0\phi) - (2/k-1)(\overline{\psi}_+\psi_+ + \overline{\psi}_-\psi_-), \tag{12.59}$$

$$J_V^1 = (2i/k)(-\partial_1\overline{\phi}\phi + \overline{\phi}\partial_1\phi) + (2/k-1)(\overline{\psi}_+\psi_+ - \overline{\psi}_-\psi_-), \tag{12.60}$$

and the conserved charge is

$$F_V = \int J_V^0\, dx^0. \tag{12.61}$$

The vector R-rotation transforms the supercharges as

$$Q_{\pm} \mapsto e^{-i\alpha}Q_{\pm},\ \overline{Q}_{\pm} \mapsto e^{i\alpha}\overline{Q}_{\pm}. \tag{12.62}$$

Also, the axial and vector R-rotations commute with each other.

12.2.2. Theory of a Twisted Chiral Superfield. One can also find a similar supersymmetric action for a twisted chiral superfield U. This time the action is expressed in the superspace as

$$S = -\int d^2x\, d\theta^4\, \overline{U}U + \left(\int d^2x d^2\widetilde{\theta}\, \widetilde{W}(U) + c.c.\right). \tag{12.63}$$

Note the minus sign in front of the D-term. This is required for the component fields to have the standard sign for the kinetic term. Chiral and twisted chiral superfields are related by the exchange of θ^- and $-\overline{\theta}^-$ which flips the sign for the D-term: $d\theta^- d\overline{\theta}^- = -d\overline{\theta}^- d\theta^-$ (the minus sign in $\theta^- \leftrightarrow -\overline{\theta}^-$ is for $Q_- \leftrightarrow \overline{Q}_-$). This last point enables us to borrow the formulae for a chiral superfield in finding the expression for the supersymmetry transformations, supercurrents, and R-symmetry generators in terms of the component fields. All we need to do is to make the replacements $\phi \to v$, $\psi_+ \to \overline{\chi}_+$, $\psi_- \to -\chi_-$, $F \to -E$, $\epsilon_+ \to -\overline{\epsilon}_+$, $Q_- \to \overline{Q}_-$ (or $G_-^{\mu} \to \overline{G}_-^{\mu}$), $F_V \to F_A$ and $F_A \to F_V$ with the others kept intact. For completeness we record here the relevant expressions. The supersymmetry transformation is

$$\begin{aligned}
\delta v &= \overline{\epsilon}_+\chi_- - \epsilon_-\overline{\chi}_+,\\
\delta\overline{\chi}_+ &= 2i\overline{\epsilon}_-\partial_+ v + \overline{\epsilon}_+ E,\\
\delta\chi_- &= -2i\epsilon_+\partial_- v + \epsilon_- E,\\
\delta E &= -2i\epsilon_+\partial_-\overline{\chi}_+ - 2i\overline{\epsilon}_-\partial_+\chi_-.
\end{aligned}$$

The supercharges are

$$\begin{aligned}
Q_+ &= \int dx^1 \left\{ 2\partial_+\overline{v}\,\overline{\chi}_+ + i\overline{\chi}_-\overline{\widetilde{W}}'(\overline{v}) \right\},\\
\overline{Q}_+ &= \int dx^1 \left\{ 2\chi_+\partial_+ v - i\chi_-\widetilde{W}'(v) \right\},\\
Q_- &= \int dx^1 \left\{ -2\overline{\chi}_-\partial_- v - i\overline{\chi}_+\widetilde{W}'(v) \right\},\\
\overline{Q}_- &= \int dx^1 \left\{ -2\partial_-\overline{v}\,\chi_- + i\chi_+\overline{\widetilde{W}}'(\overline{v}) \right\}.
\end{aligned}$$

The action is always invariant under the $U(1)$ vector R-rotation by assigning the vector R-charge of U to be zero, but it is not always invariant under the $U(1)$ axial R-rotation. It has an axial $U(1)$ R-symmetry only if the twisted superpotential $\widetilde{W}(U)$ is a monomial, say, U^k. The vector and axial R-symmetry generators are then expressed as

(12.64)
$$\begin{aligned}
F_V &= \int dx^1 \left\{ -\overline{\chi}_+\chi_+ - \overline{\chi}_-\chi_- \right\},\\
F_A &= \int dx^1 \left\{ (2i/k)(\partial_0\overline{v}v - \overline{v}\partial_0 v) - (2/k-1)(-\overline{\chi}_+\chi_+ + \overline{\chi}_-\chi_-) \right\}.
\end{aligned}$$

12.3. $\mathcal{N}=(2,2)$ Supersymmetric Quantum Field Theories

Suppose we have a classical supersymmetric field theory — an $\mathcal{N}=(2,2)$ supersymmetric action for a number of fields. Then we obtain four supercharges

(12.65) $$Q_+,\ Q_-,\ \overline{Q}_+,\ \overline{Q}_-.$$

As in any Poincaré invariant quantum field theory, we will also have Hamiltonian, momentum, and angular momentum

(12.66) $$H,\ P,\ M,$$

which are the Noether charges for the time translations $\partial/\partial x^0$, spatial translations $\partial/\partial x^1$, and Lorentz rotations $x^0\partial/\partial x^1 + x^1\partial/\partial x^0$. If the action is invariant under both vector and axial R-rotations, there are also corresponding Noether charges

(12.67) $$F_V,\ F_A.$$

If these symmetries in the classical system are not lost in the quantum theory,[1] then the conserved charges correspond, in the quantum theory, to the generators of the corresponding symmetry transformations. In particular, the conserved charges $Q_\pm, \overline{Q}_\pm$ generate the supersymmetry transformation δ by

(12.68) $$\delta\mathcal{O} = [\widehat{\delta}, \mathcal{O}],$$

where

(12.69) $$\widehat{\delta} := i\epsilon_+ Q_- - i\epsilon_- Q_+ - i\overline{\epsilon}_+\overline{Q}_- + i\overline{\epsilon}_-\overline{Q}_+.$$

Note that $\widehat{\delta}^\dagger = -\widehat{\delta}$ as a consequence of $\overline{Q}_\pm = Q_\pm^\dagger$, which is consistent with $(\delta\mathcal{O})^\dagger = \delta\mathcal{O}^\dagger$. The (anti-)commutation relations of the symmetry transformations imply the following (anti-)commutation relation of the generators;

(12.70) $$Q_+^2 = Q_-^2 = \overline{Q}_+^2 = \overline{Q}_-^2 = 0,$$

(12.71) $$\{Q_\pm, \overline{Q}_\pm\} = H \pm P,$$

(12.72) $$\{\overline{Q}_+, \overline{Q}_-\} = \{Q_+, Q_-\} = 0,$$

(12.73) $$\{Q_-, \overline{Q}_+\} = \{Q_+, \overline{Q}_-\} = 0,$$

(12.74) $$[iM, Q_\pm] = \mp Q_\pm,\ \ [iM, \overline{Q}_\pm] = \mp\overline{Q}_\pm,$$

(12.75) $$[iF_V, Q_\pm] = -iQ_\pm,\ \ [iF_V, \overline{Q}_\pm] = i\overline{Q}_\pm,$$

(12.76) $$[iF_A, Q_\pm] = \mp iQ_\pm,\ \ [iF_A, \overline{Q}_\pm] = \pm i\overline{Q}_\pm.$$

The Hermiticity property of the generators follows that of the classical one. In particular, we have

(12.77) $$Q_\pm^\dagger = \overline{Q}_\pm,$$

and other generators are Hermitian. The relations (12.72) and those in Eq. (12.73) can actually be relaxed to

(12.78) $$\{\overline{Q}_+, \overline{Q}_-\} = Z,\ \ \{Q_+, Q_-\} = Z^*,$$

(12.79) $$\{Q_-, \overline{Q}_+\} = \widetilde{Z},\ \ \{Q_+, \overline{Q}_-\} = \widetilde{Z}^*,$$

as long as Z and $\widetilde{Z}$ commute with all operators in the theory. In particular, Z and $\widetilde{Z}$ must commute with other symmetry generators and are called central charges. Thus, Z must be zero if F_V is conserved while $\widetilde{Z}$ is zero if

[1]We will see later some examples in which that is not the case due to the fact that the measure of the path-integral does not respect that symmetry; such a loss of symmetry in the quantum theory is called an anomaly.

F_A is conserved. The central charge Z will appear later in our discussion of soliton sectors of Landau–Ginzburg models. The (graded) algebra defined by the above (anti-)commutation relations of symmetry generators is called an $\mathcal{N}=(2,2)$ supersymmetry algebra.

The component fields of a superfield constitute a representation of the $\mathcal{N}=2$ supersymmetry algebra. For example, the component fields $\phi,\psi_\pm,F$ of a chiral superfield determines a representation called a *chiral multiplet* via Eq. (12.43), where we replace the transformation δ by commutation with $\widehat{\delta}$ in Eq. (12.69). Similarly, the component fields $v,\overline{\chi}_+,\chi_-,\widetilde{F}$ of a twisted chiral superfield determine a representation called a *twisted chiral multiplet.*

The lowest component ϕ of a chiral multiplet satisfies

$$[\overline{Q}_\pm,\phi]=0. \tag{12.80}$$

This can be seen as follows:

$$[\overline{Q}_\pm,\phi]=\overline{\mathcal{Q}}_+\mathcal{F}\Big|_{\theta^\pm=\overline{\theta}^\pm=0}=(\overline{D}_\pm+2i\theta^\pm\partial_\pm)\mathcal{F}\Big|_{\theta^\pm=\overline{\theta}^\pm=0}=0. \tag{12.81}$$

Conversely, if we have an operator ϕ such that $[\overline{Q}_\pm,\phi]=0$, we can construct a chiral multiplet (ϕ,ψ_+,ψ_-,F) by

$$\begin{aligned}\psi_\pm &:= [iQ_\pm,\phi],\\ F &:= \{Q_+,[Q_-,\phi]\}.\end{aligned} \tag{12.82}$$

Similarly, the lowest component v of a twisted chiral multiplet obeys

$$[\overline{Q}_+,v]=[Q_-,v]=0. \tag{12.83}$$

Conversely, if we have such a field, we can construct a twisted chiral multiplet $(v,\overline{\chi}_+,\chi_-,E)$ by

$$\begin{aligned}\overline{\chi}_+ &:= [iQ_+,v],\ \chi_- := -[i\overline{Q}_-,v],\\ E &:= -\{Q_+,[\overline{Q}_-,v]\}.\end{aligned} \tag{12.84}$$

12.4. The Statement of Mirror Symmetry

We note here an unusual symmetry of the $\mathcal{N}=(2,2)$ supersymmetry algebra. The algebra is invariant under a $\mathbb{Z}_2$ outer automorphism given by the exchange of the generators

$$\begin{aligned}Q_- &\longleftrightarrow \overline{Q}_-,\\ F_V &\longleftrightarrow F_A,\\ Z &\longleftrightarrow \widetilde{Z},\end{aligned} \tag{12.85}$$

with all other generators kept intact. Two $\mathcal{N}=(2,2)$ supersymmetic quantum field theories are said to be *mirror* to each other if they are equivalent as quantum field theories where the isomorphism of the Hilbert spaces transforms the generators of the $\mathcal{N}=(2,2)$ supersymmetry algebra according to Eq. (12.85).

Thus, if there is a pair of mirror symmetric theories, a chiral multiplet of one theory is mapped to a twisted chiral multiplet of the mirror. If the axial R-symmetry is unbroken (broken) in one theory, the vector R-symmetry is unbroken (broken) in the mirror.

It is actually a matter of convention which to call Q_- or $\overline{Q}_-$. Here we are assuming a certain convention that applies to a class of theories, called non-linear sigma models and Landau–Ginzburg models, that generalizes the basic examples considered in this section and will be studied in the following sections in more detail. The convention is that holomorphic coordinates of the non-linear sigma models or holomorphic variables of the Landau–Ginzburg models are represented by the lowest components of chiral superfields (as in the first of the basic examples). One could switch the convention so that the holomorphic coordinates/variables are represented by the lowest components of twisted chiral superfields (as in the second of the basic examples). Therefore, *if* we flip the convention of one of a mirror symmetric pair, then the two theories are equivalent without the exchange as shown by Eq. (12.85). We will sometimes encounter mirror symmetric pairs realized in this way.

12.5. Appendix

We obtain supersymmetries with half as many supercharges — $(1,1)$ and $(0,2)$ supersymmetries — by restriction of $(2,2)$ superspace to its subspaces.

12.5.1. $(1,1)$ **Supersymmetry.** We can obtain supersymmetries with fewer supercharges by restriction to a subspace of the $\mathcal{N}=(2,2)$ superspace. Here we consider $(1,1)$ supersymmetries which has two real supercharges, one with positive chirality and one with negative chirality. The relevant sub-superspace is the one where θ^+ and θ^- are real up to phases. Namely, the subspace such that

$$\theta^+=i\,e^{i\nu_+}\theta_1^+,\ \theta_1^+\ \text{real}, \tag{12.86}$$

$$\theta^-=i\,e^{i\nu_-}\theta_1^-,\ \theta_1^-\ \text{real}, \tag{12.87}$$

for arbitrary (but fixed) phases $e^{i\nu_\pm}$, where "θ_1^α real" means $(\theta_1^\alpha)^\dagger = \theta_1^\alpha$. The subspace can also be defined by the equations

$$e^{-i\nu_+}\theta^+ + e^{i\nu_+}\overline{\theta}^+ = 0, \tag{12.88}$$

$$e^{-i\nu_-}\theta^- + e^{i\nu_-}\overline{\theta}^- = 0. \tag{12.89}$$

$\theta_1^\pm$ are the fermionic coordinates of this subspace, which we call *(1,1) superspace.* The following combinations of differential operators preserves Eqs. (12.88)–(12.89), and can be written as differential operators on the $(1,1)$ superspace:

$$\mathcal{Q}_\pm^1 := e^{i\nu_\pm}\mathcal{Q}_\pm + e^{-i\nu_\pm}\overline{\mathcal{Q}}_\pm = -i\frac{\partial}{\partial\theta_1^\pm} + 2\theta_1^\pm\partial_\pm, \tag{12.90}$$

$$D_\pm^1 := e^{i\nu_\pm}D_\pm + e^{-i\nu_\pm}\overline{D}_\pm = -i\frac{\partial}{\partial\theta_1^\pm} - 2\theta_1^\pm\partial_\pm. \tag{12.91}$$

These operators obey the anti-commutation relations

$$\{\mathcal{Q}_\pm^1, \mathcal{Q}_\pm^1\} = -4i\partial_\pm,\ \{\mathcal{Q}_+^1, \mathcal{Q}_-^1\} = 0, \tag{12.92}$$

$$\{D_\pm^1, D_\pm^1\} = 4i\partial_\pm,\ \{D_+^1, D_-^1\} = 0, \tag{12.93}$$

$$\{\mathcal{Q}_\alpha^1, D_\beta^1\} = 0. \tag{12.94}$$

A superfield on the $(1,1)$ superspace (or a $(1,1)$ superfield) can be expanded as

$$\mathbf{\Phi} = \phi + i\theta_1^+\psi_+ + i\theta_1^-\psi_- + i\theta_1^+\theta_1^- f. \tag{12.95}$$

It can be complex or real, bosonic or fermionic. It is bosonic and real if $[\theta_1^\pm, \mathbf{\Phi}] = 0$ and all the component fields $(\phi, \psi_\pm, f)$ are real. Let us define the integral on the $(1,1)$ superspace as

$$\int d^2x\, d^2\theta_1\, \mathbf{F} := \int d^2x\, d\theta_1^+ d\theta_1^-\, \mathbf{F}, \tag{12.96}$$

for any function $\mathbf{F} = \mathbf{F}(\mathbf{\Phi}_i, D_\pm^1\mathbf{\Phi}_i, \ldots)$ of superfields $\mathbf{\Phi}_i$ and their $D_\pm^1$ derivatives. Then the integral is invariant under the $(1,1)$ supersymmetry transformations $\delta^1 = i\epsilon_-^1\mathcal{Q}_+^1 - i\epsilon_+^1\mathcal{Q}_-^1$. For instance, the following functional of a real bosonic superfield $\mathbf{\Phi}$ is invariant under the $(1,1)$ supersymmetry,

$$S = \int d^2x\, d^2\theta_1 \left\{ \frac{1}{2}D_-^1\mathbf{\Phi}D_+^1\mathbf{\Phi} + ih(\mathbf{\Phi}) \right\} \tag{12.97}$$

where $h(\mathbf{\Phi})$ is an arbitrary differentiable function of $\mathbf{\Phi}$. This functional can be written in terms of the component fields as

$$\begin{aligned} S = \int d^2x \Big\{ & \frac{1}{2}(\partial_0\phi)^2 - \frac{1}{2}(\partial_1\phi)^2 + \frac{1}{2}f^2 + h'(\phi)f \\ & + \frac{i}{2}\psi_-(\partial_0+\partial_1)\psi_- + \frac{i}{2}\psi_+(\partial_0-\partial_1)\psi_+ - ih''(\phi)\psi_+\psi_- \Big\}. \end{aligned} \tag{12.98}$$

By eliminating the auxiliary field f (or completing the square), we obtain the term $-\frac{1}{2}(h'(\phi))^2$. Thus, this is the action for a supersymmetric potential theory with the potential

$$U(\phi) = \frac{1}{2}\Big(h'(\phi)\Big)^2. \tag{12.99}$$

When a $(1,1)$ supersymmetric field theory is quantized appropriately, we obtain Noether charges $Q_\pm^1$ that generate the supersymmetry transformations. These will obey the anti-commutation relations

$$\{Q_\pm^1, Q_\pm^1\} = 2(H \pm P),\ \{Q_+^1, Q_-^1\} = 0. \tag{12.100}$$

A $(2,2)$ supersymmetric field theory can be regarded as a $(1,1)$ supersymmetric field theory. In particular, an invariant action on the $(2,2)$ superspace can be written as an expression on the $(1,1)$ subspace as shown by Eqs. (12.88)–(12.89). For D-terms, where one integrates over all four fermionic coordinates, one simply integrates over the two coordinates orthogonal to the subspace from Eqs. (12.88)–(12.89). This leads to the identity

$$\begin{aligned} \int d^4\theta\, \mathcal{F} =& \frac{1}{4}\int d^2\theta_1 \left[\left(e^{i\nu_+}\frac{\partial}{\partial\theta^+} + e^{-i\nu_+}\frac{\partial}{\partial\overline{\theta}^+} \right) \right. \\ & \left. \times \left(e^{i\nu_-}\frac{\partial}{\partial\theta^-} + e^{-i\nu_-}\frac{\partial}{\partial\overline{\theta}^-} \right) \mathcal{F} \right]_{(1,1)} \\ =& \frac{1}{4}\int d^2\theta_1 \left[(e^{i\nu_+}D_+ - e^{-i\nu_+}\overline{D}_+) \right. \\ & \left. \times (e^{i\nu_-}D_- - e^{-i\nu_-}\overline{D}_-)\mathcal{F} \right]_{(1,1)} + \cdots, \end{aligned}$$

where $[\cdots]_{(1,1)}$ stands for restriction to the $(1,1)$ subspace Eqs. (12.88)–(12.89), and $+\cdots$ are total derivatives in the bosonic coordinates. As for F-terms, we have the identity

$$\int d^2\theta\, W(\mathbf{\Phi}) = e^{-i(\nu_+ + \nu_-)} \int d^2\theta_1 [W(\mathbf{\Phi})]_{(1,1)}. \tag{12.101}$$

Using these identities, it is easy to see that

$$\int d^2x\, d^4\theta\, \overline{\Phi}\Phi + \frac{1}{2}\left(\int d^2x\, d^2\theta\, W(\Phi) + c.c.\right) \tag{12.102}$$
$$= \int d^2x\, d^2\theta_1 \left\{ \frac{1}{2}\sum_{I=1,2} D^1_-\Phi^I D^1_+\Phi^I + i\,\mathrm{Im}\left[\mathrm{e}^{-i(\nu_+ + \nu_-)}W(\Phi)\right]_{(1,1)} \right\},$$

where $\mathbf{\Phi}^I$ are defined by $[\Phi]_{(1,1)} = (\mathbf{\Phi}^1 + i\mathbf{\Phi}^2)/\sqrt{2}$.

12.5.2. $(0,2)$ **Supersymmetry.** We next consider $(0,2)$ supersymmetry, which has two supercharges of positive chirality. The relevant subspace of the $(2,2)$ superspace is the $(0,2)$ *superspace* defined by

$$\theta^- = \overline{\theta}^- = 0. \tag{12.103}$$

This subspace is preserved by the differential operators

$$\mathcal{Q}_+ = \frac{\partial}{\partial\theta^+} + i\overline{\theta}^+\partial_+,\ \overline{\mathcal{Q}}_+ = -\frac{\partial}{\partial\overline{\theta}^+} - i\theta^+\partial_+, \tag{12.104}$$

$$D_+ = \frac{\partial}{\partial\theta^+} - i\overline{\theta}^+\partial_+,\ \overline{D}_+ = -\frac{\partial}{\partial\overline{\theta}^+} + i\theta^+\partial_+. \tag{12.105}$$

R-rotation of the superfield is defined by

$$\mathcal{F}(x^\mu, \theta^+, \overline{\theta}^+) \mapsto \mathcal{F}(x^\mu, \mathrm{e}^{-i\alpha}\theta^+, \mathrm{e}^{i\alpha}\overline{\theta}^+). \tag{12.106}$$

A $(0,2)$ superfield $\mathit{\Phi}$ is called *chiral* when it satisfies

$$\overline{D}_+\mathit{\Phi} = 0. \tag{12.107}$$

A bosonic scalar chiral superfield $\mathit{\Phi}$ has an expansion

$$\mathit{\Phi} = \phi + \theta^+\psi_+ - i\theta^+\overline{\theta}^+\partial_+\phi. \tag{12.108}$$

We often call a fermionic chiral superfield a *Fermi superfield*. A negative chirality Fermi superfield $\mathit{\Psi}_-$ has an expansion

$$\mathit{\Psi}_- = \psi_- + \theta^+ G - i\theta^+\overline{\theta}^+\partial_+\psi_-. \tag{12.109}$$

One can find functionals of the superfields that are invariant under the $(0,2)$ supersymmetry transformations $\delta = \overline{\epsilon}_-\mathcal{Q}_- - \epsilon_-\overline{\mathcal{Q}}_+$. One is the $(0,2)$ D-term

$$\int d\theta^+ d\overline{\theta}^+\,\mathcal{F} \tag{12.110}$$

for any $(0,2)$ superfield $\mathcal{F}$ and the other is the $(0,2)$ F-term

$$\int d\theta^+ \mathcal{G}\Big|_{\overline{\theta}^+=0} \tag{12.111}$$

for any $(0,2)$ Fermi superfield $\mathcal{G}$. Examples are the following actions for a chiral superfield $\mathit{\Phi}$ and a Fermi superfield $\mathit{\Psi}_-$;

$$\begin{aligned} S_{\mathit{\Phi}} &= \int d^2x d\theta^+ d\overline{\theta}^+\, i\overline{\mathit{\Phi}}(\partial_0 - \partial_1)\mathit{\Phi} \\ &= \int d^2x\left(|\partial_0\phi|^2 - |\partial_1\phi|^2 + i\overline{\psi}_+(\partial_0-\partial_1)\psi_+\right), \end{aligned} \tag{12.112}$$

$$S_{\mathit{\Psi}_-} = \int d^2x d\theta^+ d\overline{\theta}^+\,\overline{\mathit{\Psi}}_-\mathit{\Psi}_- = \int d^2x\left(i\overline{\psi}_-(\partial_0+\partial_1)\psi_- + |G|^2\right). \tag{12.113}$$

Also, for a holomorphic function $\mathcal{V}(\mathit{\Phi})$ we have

(12.114)
$$S_{\mathcal{V}} = \int d^2x d\theta^+\,\mathit{\Psi}_-\mathcal{V}(\mathit{\Phi})\Big|_{\overline{\theta}^+=0} + c.c. = \int d^2x\Big(\mathcal{V}(\phi)G + \mathcal{V}'(\phi)\psi_+\psi_-\Big) + c.c.$$

where $\mathcal{V}(\mathit{\Phi})$ is an arbitrary holomorphic function of $\mathit{\Phi}$. When a $(0,2)$ supersymmetric field theory is quantized appropriately, we will obtain the supercharges Q_+ and $\overline{Q}_+$ that obey the anti-commutation relation

$$\{Q_+, \overline{Q}_+\} = H + P,\ Q_+^2 = \overline{Q}_+^2 = 0. \tag{12.115}$$

A $(2,2)$ supersymmetric theory can be considered as a $(0,2)$ supersymmetric theory. To obtain the $(0,2)$ expression of a $(2,2)$ invariant action, it is useful to note that

(12.116)
$$\int d^4\theta\,\mathcal{F} = \int d\theta^+ d\overline{\theta}^+ \frac{\partial}{\partial\theta^-}\frac{\partial}{\partial\overline{\theta}^-}\mathcal{F} = -\int d\theta^+ d\overline{\theta}^+ D_-\overline{D}_-\mathcal{F}\Big|_{\theta^-=\overline{\theta}^-=0}.$$

Let us consider the $(2,2)$ supersymmetric field theory of a single chiral superfield Φ considered in Sec. 12.2. The $(2,2)$ chiral multiplet splits into $(0,2)$ chiral and Fermi multiplets $(\mathit{\Phi}, \mathit{\Psi}_-)$ as follows;

$$\mathit{\Phi} = \Phi\Big|_{\theta^-=\overline{\theta}^-=0},\ \mathit{\Psi}_- = D_-\Phi\Big|_{\theta^-=\overline{\theta}^-=0}. \tag{12.117}$$

Then the $(2,2)$ invariant action $S = S_{\mathrm{kin}} + S_W$ can be written as $S_{\mathit{\Phi}} + S_{\mathit{\Psi}_-} + S_{\mathcal{V}}$ where the holomorphic function $\mathcal{V}(\mathit{\Phi})$ is given by

$$\mathcal{V}(\mathit{\Phi}) = -W'(\mathit{\Phi}). \tag{12.118}$$

CHAPTER 13

Non-linear Sigma Models and Landau–Ginzburg Models

13.1. The Models

Let us generalize our basic example of a single chiral multiplet Φ to the case with many chiral multiplets $\Phi^1, \ldots, \Phi^n$ and replace $\overline{\Phi}\Phi$ by a general real function $K(\Phi^i, \overline{\Phi^{\bar\imath}})$ of the Φ_i's and $\overline{\Phi}_{\bar\imath}$'s. For the kinetic term of the component fields to be non-degenerate with a correct sign, we assume that the matrix

$$g_{i\bar\jmath} := \partial_i \partial_{\bar\jmath} K(\Phi^i, \overline{\Phi^{\bar\imath}}) \tag{13.1}$$

is positive definite. Then one can consider this matrix as determining a Kähler metric on $\mathbb{C}^n = \{(z^1, \ldots, z^n)\}$

$$ds^2 = g_{i\bar\jmath} dz^i \, d\overline{z}^{\bar\jmath}, \tag{13.2}$$

which further defines the Levi–Civita connection $\Gamma^i_{jk} = g^{i\bar\jmath}\partial_j g_{k\bar\jmath}$ on the tangent bundle $T\mathbb{C}^n$. Under this assumption, we consider the Lagrangian density

$$\mathcal{L}_{\text{kin}} = \int d^4\theta \, K(\Phi^i, \overline{\Phi^{\bar\imath}}). \tag{13.3}$$

In terms of component fields $\phi^i, \psi^i_\pm, F^i$ of Φ^i, $\mathcal{L}_{\text{kin}}$ can be expressed as

$$\begin{aligned} \mathcal{L}_{\text{kin}} = &- g_{i\bar\jmath}\partial^\mu \phi^i \partial_\mu \overline{\phi}^{\bar\jmath} + i g_{i\bar\jmath} \overline{\psi}^{\bar\jmath}_- (D_0 + D_1)\psi^i_- \\ &+ i g_{i\bar\jmath}\overline{\psi}^{\bar\jmath}_+ (D_0 - D_1)\psi^i_+ + R_{i\bar\jmath k\bar l}\psi^i_+ \psi^k_- \overline{\psi}^{\bar\jmath}_- \overline{\psi}^{\bar l}_+ \\ &+ g_{i\bar\jmath}(F^i - \Gamma^i_{jk}\psi^j_+\psi^k_-)(\overline{F}^{\bar\jmath} - \Gamma^{\bar\jmath}_{\bar k \bar l}\overline{\psi}^{\bar k}_- \overline{\psi}^{\bar\jmath}_+), \end{aligned} \tag{13.4}$$

up to total derivatives in x^μ. The kinetic terms are non-singular under the assumption that $g_{i\bar\jmath}$ is positive definite. In the above expression, $R_{i\bar\jmath k\bar l}$ is the Riemannian curvature of the metric in Eq. (13.2) and D_μ is defined by

$$D_\mu \psi^i_\pm := \partial_\mu \psi^i_\pm + \partial_\mu \phi^j \Gamma^i_{jk}\psi^k_\pm. \tag{13.5}$$

We note here that the expression shown in Eq. (13.4) is covariant under holomorphic coordinate changes of $z^1, \ldots, z^n$ except for the last term, which can be eliminated by the equation of motion. If we change the coordinates, the action is invariant under an appropriate change of variables. Also, the action is invariant under the "Kähler transformation"

$$K(\Phi^i, \overline{\Phi^i}) \to K(\Phi^i, \overline{\Phi^i}) + f(\Phi^i) + \overline{f}(\overline{\Phi^i});\ f(\Phi^i) \text{ holomorphic,} \tag{13.6}$$

which leaves the metric from Eq. (13.2) invariant. This is manifest in the component expression as shown by Eq. (13.4) but can also be understood by the fact that $\int d^4\theta f(\Phi)$ is a total derivative if $f(\Phi^i)$ is holomorphic. Thus, we can apply this construction for each coordinate patch of a Kähler manifold M (possibly with more complicated topology than $\mathbb{C}^n$), and then glue the patches together by the invariance of the action under coordinate change and Kähler transformation. This will lead us to define an action for a map of the worldsheet to any Kähler manifold:

$$\phi : \Sigma \to M. \tag{13.7}$$

Then the fermions are the spinors with values in the pull-back of the tangent bundle, ϕ^*TM;

$$\psi_\pm \in \Gamma(\Sigma, \phi^*TM^{(1,0)} \otimes S_\pm), \tag{13.8}$$

$$\overline{\psi}_\pm \in \Gamma(\Sigma, \phi^*TM^{(0,1)} \otimes S_\pm). \tag{13.9}$$

The derivative in Eq. (13.5) is the covariant derivative with respect to the Levi–Civita connection pulled back to the worldsheet Σ by the map ϕ. This system is called a supersymmetric non-linear sigma model on a Kähler manifold M with metric g. Note that this formulation is not global, and the supersymmetry must be checked patch-by-patch. This is a limitation of this formulation and it indeed has some drawbacks (e.g., one cannot see the separation of parameters into the *cc* and *ac* parts that we will later introduce). Later in this section, we will find a global formulation of another model that falls into the same "universality class" as the non-linear sigma model.

Let us next consider an F-term

$$\mathcal{L}_W = \frac{1}{2}\left(\int d^2\theta\, W(\Phi^i) \ +\ c.c.\right). \tag{13.10}$$

Here $W(\Phi^i)$ is the superpotential, which is a holomorphic function of $\Phi^1, \ldots, \Phi^n$, or in the case of a sigma model on M, W is a holomorphic function on M (which is non-trivial only when M is non-compact). In terms of the component fields, the F-term is expressed as

$$\mathcal{L}_W = \frac{1}{2}F^i\partial_i W - \frac{1}{2}\partial_i\partial_j\psi_+^i\psi_-^j + \frac{1}{2}\overline{F}^{\bar{i}}\partial_{\bar{i}}\overline{W} - \frac{1}{2}\partial_{\bar{i}}\partial_{\bar{j}}\overline{W}\overline{\psi}_-^{\bar{i}}\overline{\psi}_+^{\bar{j}}. \tag{13.11}$$

The total Lagrangian is the sum of $\mathcal{L}_{\text{kin}}$ and $\mathcal{L}_W$

$$\mathcal{L} = \int d^4\theta\, K(\Phi^i, \overline{\Phi^i}) + \frac{1}{2}\left(\int d^2\theta\, W(\Phi^i) \ +\ c.c.\right). \tag{13.12}$$

The fields F^i and $\overline{F}^{\bar{i}}$ are again auxiliary fields and can be eliminated by their equations of motion,

$$F^i = \Gamma^i_{jk}\psi_+^j\psi_-^k - \frac{1}{2}g^{i\bar{l}}\partial_{\bar{l}}\overline{W}, \tag{13.13}$$

$$\overline{F}^{\bar{j}} = \Gamma^{\bar{j}}_{\bar{i}\bar{k}}\overline{\psi}_-^{\bar{i}}\overline{\psi}_+^{\bar{k}} - \frac{1}{2}g^{\bar{j}l}\partial_l W. \tag{13.14}$$

Then the total Lagrangian can be expressed in terms of the component fields as

$$\begin{aligned}\mathcal{L} = &- g_{i\bar{j}}\partial^\mu\phi^i\partial_\mu\overline{\phi}^{\bar{j}} + ig_{i\bar{j}}\overline{\psi}_-^{\bar{j}}(D_0 + D_1)\psi_-^i \\ &+ ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}(D_0 - D_1)\psi_+^i + R_{i\bar{j}k\bar{l}}\psi_+^i\psi_-^k\overline{\psi}_-^{\bar{j}}\overline{\psi}_+^{\bar{l}} \\ &- \frac{1}{4}g^{\bar{i}j}\partial_{\bar{i}}\overline{W}\partial_j W - \frac{1}{2}D_i\partial_j W\psi_+^i\psi_-^j - \frac{1}{2}D_{\bar{i}}\partial_{\bar{j}}\overline{W}\overline{\psi}_-^{\bar{i}}\overline{\psi}_+^{\bar{j}}\end{aligned} \tag{13.15}$$

By construction, the above Lagrangian is invariant under $\mathcal{N} = (2,2)$ supersymmetry. The supersymmetry variations of the component fields are expressed as

$$\begin{aligned}\delta\phi^i &= \epsilon_+\psi_-^i - \epsilon_-\psi_+^i, & \delta\overline{\phi}^{\bar{i}} &= -\overline{\epsilon}_+\overline{\psi}_-^{\bar{i}} + \overline{\epsilon}_-\overline{\psi}_+^{\bar{i}}, \\ \delta\psi_+^i &= 2i\overline{\epsilon}_-\partial_+\phi^i + \epsilon_+F^i, & \delta\overline{\psi}_+^{\bar{i}} &= -2i\epsilon_-\partial_+\overline{\phi}^{\bar{i}} + \overline{\epsilon}_+\overline{F}^{\bar{i}}, \\ \delta\psi_-^i &= -2i\overline{\epsilon}_+\partial_-\phi^i + \epsilon_-F^i, & \delta\overline{\psi}_-^{\bar{i}} &= 2i\epsilon_+\partial_-\overline{\phi}^{\bar{i}} + \overline{\epsilon}_-\overline{F}^{\bar{i}},\end{aligned} \tag{13.16}$$

where F^i and $\overline{F}^{\bar{i}}$ are as given in Eqs. (13.13)–(13.14). Following the Noether procedure, we find the four conserved currents $G^\mu_\pm$ and $\overline{G}^\mu_\pm$, which are defined by

$$\delta\int \mathcal{L}d^2x = \int d^2x\{\partial_\mu\epsilon_+ G^\mu_- - \partial_\mu\epsilon_- G^\mu_+ + \partial_\mu\overline{\epsilon}_-\overline{G}^\mu_+ - \partial_\mu\overline{\epsilon}_+\overline{G}^\mu_-\}. \tag{13.17}$$

These currents — *supercurrents* — are expressed as

$$G^0_\pm = 2g_{i\bar{j}}\partial_\pm\overline{\phi}^{j}\psi^i_\pm \mp \frac{i}{2}\overline{\psi}^{\bar{i}}_\mp\partial_{\bar{i}}\overline{W}, \tag{13.18}$$

$$G^1_\pm = \mp 2g_{i\bar{j}}\partial_\pm\overline{\phi}^{j}\psi^i_\pm - \frac{i}{2}\overline{\psi}^{\bar{i}}_\mp\partial_{\bar{i}}\overline{W}, \tag{13.19}$$

$$\overline{G}^0_\pm = 2g_{\bar{i}j}\overline{\psi}^{\bar{i}}_\pm\partial_\pm\phi^j \pm \frac{i}{2}\psi^i_\mp\partial_i W, \tag{13.20}$$

$$\overline{G}^1_\pm = \mp 2g_{\bar{i}j}\overline{\psi}^{\bar{i}}_\pm\partial_\pm\phi^j \pm \frac{i}{2}\psi^i_\mp\partial_i W. \tag{13.21}$$

The conserved charges — *supercharges* — are given by

$$Q_\pm = \int dx^1 G^0_\pm, \tag{13.22}$$

$$\overline{Q}_\pm = \int dx^1 \overline{G}^0_\pm. \tag{13.23}$$

Inclusion of B-field. As in the bosonic sigma model with the target space T^2, if there is a non-trivial cohomology class $B \in H^2(M, \mathbb{R})$ one can modify the theory by putting the phase factor

$$\exp\left(i\int \phi^* B\right) \tag{13.24}$$

in the path-integral. This factor is invariant under a continuous deformation of the map ϕ. In particular, it is invariant under the supersymmetry variation and this modification does not break the supersymmetry. Also, the forms of the supercurrent and the supercharges remain the same as above.

13.2. R-Symmetries

We recall that the vector and axial R-rotations act on the superfields as

$$V: \Phi^i(x,\theta^\pm,\overline{\theta}^\pm) \mapsto e^{i\alpha q^i_V}\Phi^i(x,\, e^{-i\alpha}\theta^\pm,\, e^{i\alpha}\overline{\theta}^\pm), \tag{13.25}$$

$$A: \Phi^i(x,\theta^\pm,\overline{\theta}^\pm) \mapsto e^{i\alpha q^i_V}\Phi^i(x,\, e^{\mp i\beta}\theta^\pm,\, e^{\pm i\beta}\overline{\theta}^\pm). \tag{13.26}$$

These can be considered as the action of group $U(1)_V \times U(1)_A$ of R-rotations. We would like to ask under what conditions these R-rotations are symmetries of the system.

13.2.1. Classical Level.

At the classical level, these are symmetries under which the action is invariant. Since the D-term $S_{\rm kin} = \int d^2x\mathcal{L}_{\rm kin}$ and the F-term $S_W = \int d^2x\mathcal{L}_W$ are not mixed under R-rotations, these must be independently invariant. Let us first consider the D-term $S_{\rm kin}$. As remarked in the single-variable case, θ^4 is invariant under both R-rotations. Thus,

$S_{\rm kin}$ is invariant under $U(1)_V$ ($U(1)_A$) if one can assign vector (axial) R-charges for Φ^i such that $K(\Phi^i, \overline{\Phi}^{j})$ has vector (axial) charge zero. This is usually possible by assigning trivial R-charges to the fields Φ^i. However, if $K(\Phi^i, \overline{\Phi}^i)$ is a function of $|\Phi^i|^2 = \Phi^i\overline{\Phi}^i$, the D-term is R-invariant under *any* assignment of R-charges to Φ^i's. Next let us consider the F-term S_W. Since θ^2 has vector R-charge -2 and axial R-charge 0, the F-term is invariant under $U(1)_V$ ($U(1)_A$) if one can assign R-charges to the Φ^i's so that $W(\Phi^i)$ has vector R-charge 2 (axial R-charge 0). For $U(1)_A$, this can be done by assigning trivial R-charges to Φ^i. For $U(1)_V$, this depends on the form of the superpotential. We call a holomorphic function W such that this is possible a *quasi-homogeneous function*. Namely, it is quasi-homogeneous when

$$W(\lambda^{q^i}\Phi^i) = \lambda^2 W(\Phi^i), \tag{13.27}$$

for some q^i which is identified as the right vector R-charge to make the F-term vector R-invariant. Let us summarize what we have seen *at the classical level: $U(1)_A$ is always a symmetry by assigning axial R-charge zero to all fields. However, $U(1)_V$ is a symmetry only if the superpotential is quasi-homogeneous.* The Kähler potential must also be invariant (up to Kähler transformations) by the assignment of the vector R-charge determined by the quasi-homogeneity. The non-linear sigma model without superpotential has both $U(1)_V$ and $U(1)_A$ symmetries.

What we have said above concerns the full $U(1)$ groups of R-rotations. However, even if the full $U(1)$ is not a symmetry it is possible that some subgroup is still a symmetry. For example, such is the case if Eq. (13.27) holds under some non-trivial phase λ. For instance it always holds for $\lambda = -1$ by assigning $q_i = 0$. Thus, the $\mathbb{Z}_2$ subgroup of the vector R-rotation group $U(1)_V$ is always a symmetry. Actually this has to be the case since this $\mathbb{Z}_2$ action is the same as the action of the $\mathbb{Z}_2$ subgroup of $U(1)_A$. The generator of this $\mathbb{Z}_2$ group is denoted by $(-1)^F$ and is an important operator in a supersymmetric theory, as noted before. In some cases, a $\mathbb{Z}_{2p}$ subgroup can be a symmetry. (An example is the theory with superpotential $W = \Phi^{p+1} + \Phi$, with suitable D-term.)

13.2.2. Anomaly.

The invariance of the action does not necessarily mean the symmetry of the quantum theory. It is symmetric if the correlation functions are invariant:

$$\langle \delta \mathcal{O} \rangle = \int \mathcal{D}X \, \mathrm{e}^{iS} \mathcal{O} = 0. \tag{13.28}$$

This is the case when the path-integral measure is also invariant, $\delta(\mathcal{D}X) = 0$, or more generally

$$\delta(\mathcal{D}X \, \mathrm{e}^{iS}) = 0. \tag{13.29}$$

When the classical symmetry $\delta S = 0$ is lost in the quantum theory by $\delta \mathcal{D}X \neq 0$, we say that the symmetry is *anomalous*. Now, let us examine whether the $U(1)_V$ and $U(1)_A$ R-symmetries of the non-linear sigma model without superpotential, $W = 0$, are really symmetries of the quantum theory. We recall that these R-rotations act only on the fermions:

$$V : \psi^i_\pm \to \mathrm{e}^{-i\alpha} \psi^i_\pm, \tag{13.30}$$

$$A : \psi^i_\pm \to \mathrm{e}^{\mp i\beta} \psi^i_\pm. \tag{13.31}$$

Thus, the question is whether the path-integral measure for fermions is invariant under these phase rotations.

A Toy Model. To see this, let us consider the simpler system of a Dirac fermion coupled to a background (Hermitian) gauge field A on the worldsheet Σ. We take Σ to be a Euclidean torus $\Sigma = T^2$ with a flat coordinate $z \equiv z+1 \equiv z+\tau$. The gauge field A is considered as a Hermitian connection of a complex vector bundle E with a Hermitian metric, and the fermions are spinors with values in E:

$$\psi_\pm \in \Gamma(T^2, E \otimes S_\pm), \tag{13.32}$$

$$\overline{\psi}_\pm \in \Gamma(T^2, E^* \otimes S_\pm). \tag{13.33}$$

Here $S_\pm$ are the positive and negative spinor bundles and E^* is the dual bundle of E. The action is given by

$$S = \int_{T^2} d^2 z (i\overline{\psi}_+ D_z \psi_+ + i\overline{\psi}_- D_{\overline{z}} \psi_-) \tag{13.34}$$

where

$$D_z = \partial_z + A_z, \; D_{\overline{z}} = \partial_{\overline{z}} + A_{\overline{z}}. \tag{13.35}$$

This action is invariant under the phase rotations of the fermions — the vector and axial rotations as in Eqs. (13.30)–(13.31). We denote the corresponding groups by $U(1)_V$ and $U(1)_A$.

Suppose the first Chern class of E is nonzero, say positive:

$$k := \int_{T^2} c_1(E) = \frac{i}{2\pi} \int_{T^2} \mathrm{Tr}\, F_A > 0. \tag{13.36}$$

Then by the index theorem

$$\dim \mathrm{Ker} D_{\overline{z}} - \dim \mathrm{Ker} D_z = \int_{T^2} c_1(E) = k, \tag{13.37}$$

the number of ψ_- zero modes (= the number of $\overline{\psi}_+$ zero modes) is larger by k than the number of $\overline{\psi}_-$ zero modes (= the number of ψ_+ zero modes). Thus, the partition function vanishes due to integration over the zero modes.

$$\int \mathcal{D}\psi \mathcal{D}\overline{\psi} \, \mathrm{e}^{-S[\psi, \overline{\psi}]} = 0. \tag{13.38}$$

To obtain a nonzero correlation function we need a certain kind of operator to absorb the zero modes. Let us consider the generic case where there are exactly k $D_{\overline{z}}$-zero modes and no D_z zero modes. Then the following correlator is non-vanishing:

$$\langle \psi_-(z_1) \cdots \psi_-(z_k) \overline{\psi}_+(w_1) \cdots \overline{\psi}_+(w_k) \rangle \neq 0. \tag{13.39}$$

Under the vector and axial rotations of the inserted operators, this correlation function transforms as

$$\begin{aligned} &\langle \psi_-(z_1) \cdots \psi_-(z_k) \overline{\psi}_+(w_1) \cdots \overline{\psi}_+(w_k) \rangle \\ &\xrightarrow{V} \langle \mathrm{e}^{-i\alpha} \psi_-(z_1) \cdots \mathrm{e}^{-i\alpha} \psi_-(z_k) \, \mathrm{e}^{i\alpha} \overline{\psi}_+(w_1) \cdots \mathrm{e}^{i\alpha} \overline{\psi}_+(w_k) \rangle \\ &= \langle \psi_-(z_1) \cdots \psi_-(z_k) \overline{\psi}_+(w_1) \cdots \overline{\psi}_+(w_k) \rangle \end{aligned} \tag{13.40}$$

$$\begin{aligned} &\langle \psi_-(z_1) \cdots \psi_-(z_k) \overline{\psi}_+(w_1) \cdots \overline{\psi}_+(w_k) \rangle \\ &\xrightarrow{A} \langle \mathrm{e}^{i\beta} \psi_-(z_1) \cdots \mathrm{e}^{i\beta} \psi_-(z_k) \, \mathrm{e}^{i\beta} \overline{\psi}_+(w_1) \cdots \mathrm{e}^{i\beta} \overline{\psi}_+(w_k) \rangle \\ &= \mathrm{e}^{2ik\beta} \langle \psi_-(z_1) \cdots \psi_-(z_k) \overline{\psi}_+(w_1) \cdots \overline{\psi}_+(w_k) \rangle. \end{aligned} \tag{13.41}$$

Thus, we see an anomaly of the $U(1)_A$ symmetry since a $U(1)_A$ non-invariant field can acquire an expectation value, while *the $U(1)_V$ symmetry is never anomalous*. One can also see explicitly that the measure is not $U(1)_A$ invariant (but is $U(1)_V$ invariant). Let us expand the fermions $\psi_\pm$, $\overline{\psi}_\pm$ in the eigenfunctions of the operators $D_{\bar{z}}^{\dagger}D_{\bar{z}}$ and $D_z^{\dagger}D_z$ etc:

$$\overline{\psi}_- = \sum_{n=1}^{\infty} b_n \overline{\varphi}_-^n, \ \psi_- = \sum_{\alpha=1}^{k} c_{0\alpha}\varphi_-^{0\alpha} + \sum_{n=1}^{\infty} c_n \varphi_-^n, \tag{13.42}$$

$$\psi_+ = \sum_{n=1}^{\infty} \widetilde{b}_n \varphi_+^n, \ \overline{\psi}_+ = \sum_{\alpha=1}^{k} \widetilde{c}_{0\alpha}\overline{\varphi}_+^{0\alpha} + \sum_{n=1}^{\infty} \widetilde{c}_n \overline{\varphi}_+^n, \tag{13.43}$$

where $\varphi_\pm^n$ and $\overline{\varphi}_\pm^n$ are the nonzero modes with eigenvalues λ_n while $\varphi_-^{0\alpha}$ and $\overline{\varphi}_+^{0\alpha}$ are the zero modes. The path-integral measure is given by
(13.44)

$$\mathcal{D}\psi\,\mathcal{D}\overline{\psi}\ \mathrm{e}^{-S} = \prod_{\alpha=1}^{k} dc_{0\alpha}d\widetilde{c}_{0\alpha} \prod_{n=1}^{\infty} db_n dc_n d\widetilde{b}_n d\widetilde{c}_n\ \mathrm{e}^{-\sum_{n=1}^{n}\lambda_n(b_n c_n + \widetilde{c}_n\widetilde{b}_n)}.$$

The measure $db_n dc_n d\widetilde{b}_n d\widetilde{c}_n$ is invariant under both $U(1)_V$ and $U(1)_A$ but $dc_{0\alpha}d\widetilde{c}_{0\alpha}$ has vector charge zero but axial charge 2. This is a direct way to see that the measure is $U(1)_V$-invariant but not $U(1)_A$-invariant, showing that $U(1)_V$ symmetry is not anomalous but $U(1)_A$ symmetry is anomalous. This argument also applies to non-generic cases where D_z has some zero modes and $D_{\bar{z}}$ has k more zero modes.

Although the full $U(1)_A$ symmetry is broken, its $\mathbb{Z}_{2k}$ subgroup $\{\mathrm{e}^{2\pi i l/2k}\}$, $0 \leq l \leq 2k-1$, remains a symmetry of the quantum theory, as can be seen from Eq. (13.41) or Eq. (13.44). If, over the different components of the space of maps, k assumes every integer value, then only a $\mathbb{Z}_2$ subgroup is anomaly-free. If k is allowed to take only integer multiples of some integer p, then a larger subgroup $\mathbb{Z}_{2p}$ is anomaly-free.

Back to the Sigma Model. Now let us come back to the R-symmetry of the non-linear sigma models. On the Euclidean torus T^2 the fermionic kinetic terms are expressed as

$$-2ig_{i\bar{j}}\overline{\psi}_-^{\bar{j}}D_{\bar{z}}\psi_-^i + 2ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}D_z\psi_+^i, \tag{13.45}$$

which is of the form shown in Eq. (13.34) with $E = \phi^*TM^{(1,0)}$. The action also includes the four-fermi terms $R_{i\bar{j}k\bar{l}}\psi_+^i\psi_-^k\overline{\psi}_-^{\bar{j}}\overline{\psi}_+^{\bar{l}}$. In the large radius expansion of the sigma model (which will be explained systematically in later chapters), the four-fermi terms are treated as a perturbation and the path-integral measure is constructed using the spectral decomposition of the Dirac operator that appears in the kinetic term from Eq. (13.45). Thus, as far as the R-symmetry is concerned, the situation for a fixed $\phi : \Sigma \to M$ is identical to the one in the above toy model with $E = \phi^*TM^{(1,0)}$.

One consequence is that the vector R-symmetry $U(1)_V$ is not anomalous and is a symmetry of the quantum theory. Also, for a given map ϕ, the $U(1)_A$ R-symmetry is broken to $\mathbb{Z}_{2k}$ where k is

$$k = \int_\Sigma c_1(\phi^*TM^{(1,0)}) = \int_\Sigma \phi^* c_1(T^{(1,0)}) = \langle c_1(M), \phi_*[\Sigma]\rangle. \tag{13.46}$$

This depends only on the homology class $\phi_*[\Sigma]$. If k can take all integer values by varying the homology class $\phi_*[\Sigma]$, then $U(1)_A$ is broken to $\mathbb{Z}_2$. If k is divisible by p for any map $\phi : \Sigma \to M$, then $U(1)_A$ is broken to $\mathbb{Z}_{2p}$. Such is the case when $c_1(M)$ is p times some integral cohomology class (e.g., for $M = \mathbb{CP}^{N-1}$ $c_1(M)$ is N times the generator of $H^2(M,\mathbb{Z}) \cong \mathbb{Z}$; thus $U(1)_A$ is broken to $\mathbb{Z}_{2N}$ in the $\mathbb{CP}^{N-1}$ sigma model). Finally, if $k=0$ for any map ϕ, $U(1)_A$ is not anomalous and is a symmetry of the quantum theory. Such is the case when $c_1(M) = 0$, namely when M is a Calabi–Yau manifold. Another way to state the axial anomaly is in terms of the B-field. Since in the path-integral h has the phase factor

$$\exp\left(i\int_\Sigma \phi^*B\right), \tag{13.47}$$

the phase rotation of the measure by $\mathrm{e}^{2ik\beta}$ with k given by Eq. (13.46) is equivalent to the shift in the cohomology class of the B-field

$$[B] \to [B] - 2\beta c_1(M). \tag{13.48}$$

Summary:

	$U(1)_V$	$U(1)_A$
CY sigma model	○	○
sigma model on M with $c_1(M) \neq 0$	○	×
LG model on CY with generic W	×	○
LG model on CY with quasi-homogeneous W	○	○

The ×'s in the table denote lack of the corresponding $U(1)$ R-symmetries. Depending on the manifold or superpotential, some discrete subgroup of even order is unbroken.

13.3. Supersymmetric Ground States

Let us study the supersymmetric ground states and Witten index of the system. We first compactify the spatial direction on the circle S^1 and put periodic boundary conditions on S^1 for all the fields. Let Q and $Q^\dagger$ be either

$$\left\{\begin{array}{l} Q_A = \overline{Q}_+ + Q_-, \\ Q_A^\dagger = Q_+ + \overline{Q}_-, \end{array}\right. \quad \text{or} \quad \left\{\begin{array}{l} Q_B = \overline{Q}_+ + \overline{Q}_-, \\ Q_B^\dagger = Q_+ + Q_-. \end{array}\right. \tag{13.49}$$

Then by the supersymmetry algebra (with $Z = \widetilde{Z} = 0$) we see

$$\{Q, Q^\dagger\} = 2H, \tag{13.50}$$

$$Q^2 = Q^{\dagger 2} = 0. \tag{13.51}$$

We notice that this is the relation defining a supersymmetric quantum mechanics (SQM). In fact, we can consider the system as a quantum mechanics with infinitely many degrees of freedom. The supersymmetric ground state we are after is the supersymmetric ground states of this SQM. As explained in the lectures on SQM, we can characterize the supersymmetric ground states as the cohomology classes of the Q-complex, and the Witten index is the Euler characteristic of the Q-complex. We also note that if F_A and F_V are conserved, we have

$$[F_A, Q_A] = Q_A \text{ and } [F_V, Q_B] = Q_B. \tag{13.52}$$

Thus, the Q-complex and cohomology groups are graded by the axial charge for $Q = Q_A$ and by the vector charge for $Q = Q_B$. Even if F_A or F_V is not conserved, if some subgroup $\mathbb{Z}_{2p}$ of $U(1)_A$ or $U(1)_V$ is a symmetry of the theory, the Q-complex/cohomology is graded by the $\mathbb{Z}_{2p}$ charges.

Let us take a closer look at the operator $Q = Q_A = \overline{Q}_+ + Q_-$. Using Eqs. (13.18)–(13.21) we find the expression

$$Q = -i \int_{S^1} \Big\{ ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}\partial_0\phi^i + ig_{i\bar{j}}\psi_-^i\partial_0\overline{\phi}^{\bar{j}} - ig_{i\bar{j}}\psi_-^i\partial_1\overline{\phi}^{\bar{j}} + ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}\partial_1\phi^i - \frac{1}{2}\psi_-^i\partial_i W - \frac{1}{2}\overline{\psi}_+^{\bar{i}}\partial_{\bar{i}}\overline{W} \Big\} dx^1. \tag{13.53}$$

If there is a functional h of $\phi(x^1)$ such that

$$\frac{\delta h}{\delta \phi^i} = -ig_{i\bar{j}}\partial_1\overline{\phi}^{\bar{j}} - \frac{1}{2}\partial_i W, \tag{13.54}$$

$$\frac{\delta h}{\delta \overline{\phi}^{\bar{j}}} = ig_{i\bar{j}}\partial_1\phi^i - \frac{1}{2}\partial_{\bar{j}}\overline{W}, \tag{13.55}$$

then the operator Q can be written in the form

$$Q = \int_{S^1} \overline{\psi}^I(x^1)\left(ig_{IJ}(x^1)\partial_0\phi^J(x^1) + \frac{\delta h}{\delta\phi^I(x^1)}\right) dx^1, \tag{13.56}$$

where we have set $\overline{\psi}^i = -i\psi_-^i$ and $\overline{\psi}^{\bar{j}} = -i\overline{\psi}_+^{\bar{j}}$. This is exactly the same form as the supercharge as shown by Eq. (10.241) for the SQM deformed by a function h. In the present case, the target space is an infinite-dimensional space of $\phi(x^1)$, namely the space of loops in M,

$$LM = \left\{ \phi : S^1 \to M \right\}. \tag{13.57}$$

Now, the question is whether there is a function h on LM such that the infinitesimal variations are given by Eqs. (13.54)–(13.55).

The function h_2 that yields the second terms is easy to find; it is simply

$$h_2 = -\int_{S^1} \mathrm{Re}\left[W(\phi^i)\right] dx^1. \tag{13.58}$$

The function h_1 that yields the first terms can be constructed as follows. The connected components of the loop space LM are classified by the fundamental group $\pi_1(M)$. We choose and fix a loop, a base loop, in each connected component. Let us pick a component and denote the base loop there by ϕ_0. For a loop ϕ in that component we choose a homotopy $\widehat{\phi}$ that connects ϕ_0 to ϕ. Namely, $\widehat{\phi} = \widehat{\phi}(x^1, \tau)$ is a map from $S^1 \times [0,1]$ to M such that $\widehat{\phi}(x^1, 0) = \phi_0(x^1)$ and $\widehat{\phi}(x^1, 1) = \phi(x^1)$. Now, let us consider the area

$$h_1 = \int_{S^1\times[0,1]} \widehat{\phi}^*\omega \tag{13.59}$$

where ω is the Kähler form of M;

$$\omega = ig_{i\bar{j}}dz^i \wedge d\bar{z}^{\bar{j}}. \tag{13.60}$$

For a variation of $\widehat{\phi}$, the pull-back $\widehat{\phi}^*\omega$ changes by a total derivative

$$\delta\widehat{\phi}^*\omega = d\left(ig_{i\bar{j}}\delta\widehat{\phi}^i d\overline{\widehat{\phi}}^{\bar{j}} - ig_{i\bar{j}}d\widehat{\phi}^i\delta\overline{\widehat{\phi}}^{\bar{j}}\right). \tag{13.61}$$

and therefore the area changes by the boundary terms

$$\begin{aligned} \delta\int_{S^1\times[0,1]} \widehat{\phi}^*\omega &= \int_{S^1} \left\{ -ig_{i\bar{j}}\delta\widehat{\phi}^i d\overline{\widehat{\phi}}^{\bar{j}} + ig_{i\bar{j}}d\widehat{\phi}^i\delta\overline{\widehat{\phi}}^{\bar{j}} \right\}\Big|_{\tau=0}^{\tau=1} \\ &= \int_{S^1} \left\{ -ig_{i\bar{j}}\delta\phi^i d\overline{\phi}^{\bar{j}} + ig_{i\bar{j}}d\phi^i\delta\overline{\phi}^{\bar{j}} \right\}, \end{aligned} \tag{13.62}$$

where we have used the constraint that $\widehat{\phi}|_{\tau=0}$ is fixed to be ϕ_0 and thus $\delta\widehat{\phi}|_{\tau=0} = 0$. In particular, for a fixed loop ϕ the functional $h_1 = \int \widehat{\phi}^*\omega$ does periodic boundary conditions on S^1 for all the fields. Let Q and $Q^\dagger$ be either

$$\left\{\begin{array}{l} Q_A = \overline{Q}_+ + Q_-, \\ Q_A^\dagger = Q_+ + \overline{Q}_-, \end{array}\right. \quad \text{or} \quad \left\{\begin{array}{l} Q_B = \overline{Q}_+ + \overline{Q}_-, \\ Q_B^\dagger = Q_+ + Q_-. \end{array}\right. \tag{13.49}$$

Then by the supersymmetry algebra (with $Z = \widetilde{Z} = 0$) we see

$$\{Q, Q^\dagger\} = 2H, \tag{13.50}$$

$$Q^2 = Q^{\dagger 2} = 0. \tag{13.51}$$

We notice that this is the relation defining a supersymmetric quantum mechanics (SQM). In fact, we can consider the system as a quantum mechanics with infinitely many degrees of freedom. The supersymmetric ground state we are after is the supersymmetric ground states of this SQM. As explained in the lectures on SQM, we can characterize the supersymmetric ground states as the cohomology classes of the Q-complex, and the Witten index is the Euler characteristic of the Q-complex. We also note that if F_A and F_V are conserved, we have

$$[F_A, Q_A] = Q_A \text{ and } [F_V, Q_B] = Q_B. \tag{13.52}$$

Thus, the Q-complex and cohomology groups are graded by the axial charge for $Q = Q_A$ and by the vector charge for $Q = Q_B$. Even if F_A or F_V is not conserved, if some subgroup $\mathbb{Z}_{2p}$ of $U(1)_A$ or $U(1)_V$ is a symmetry of the theory, the Q-complex/cohomology is graded by the $\mathbb{Z}_{2p}$ charges.

Let us take a closer look at the operator $Q = Q_A = \overline{Q}_+ + Q_-$. Using Eqs. (13.18)–(13.21) we find the expression

$$Q = -i \int_{S^1} \left\{ ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}\partial_0\phi^i + ig_{i\bar{j}}\psi_-^i\partial_0\overline{\phi}^{\bar{j}} - ig_{i\bar{j}}\psi_-^i\partial_1\overline{\phi}^{\bar{j}} + ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}\partial_1\phi^i - \frac{1}{2}\psi_-^i\partial_iW - \frac{1}{2}\overline{\psi}_+^{\bar{\imath}}\partial_{\bar{\imath}}\overline{W} \right\} dx^1. \tag{13.53}$$

If there is a functional h of $\phi(x^1)$ such that

$$\frac{\delta h}{\delta\phi^i} = -ig_{i\bar{j}}\partial_1\overline{\phi}^{\bar{j}} - \frac{1}{2}\partial_iW, \tag{13.54}$$

$$\frac{\delta h}{\delta\overline{\phi}^{\bar{j}}} = ig_{i\bar{j}}\partial_1\phi^i - \frac{1}{2}\partial_{\bar{j}}\overline{W}, \tag{13.55}$$

then the operator Q can be written in the form

$$Q = \int_{S^1} \overline{\psi}^I(x^1)\left(ig_{IJ}(x^1)\partial_0\phi^J(x^1) + \frac{\delta h}{\delta\phi^I(x^1)}\right)dx^1, \tag{13.56}$$

where we have set $\overline{\psi}^i = -i\psi_-^i$ and $\overline{\psi}^{\bar{j}} = -i\overline{\psi}_+^{\bar{j}}$. This is exactly the same form as the supercharge as shown by Eq. (10.241) for the SQM deformed by a function h. In the present case, the target space is an infinite-dimensional space of $\phi(x^1)$, namely the space of loops in M,

$$LM = \left\{ \phi : S^1 \to M \right\}. \tag{13.57}$$

Now, the question is whether there is a function h on LM such that the infinitesimal variations are given by Eqs. (13.54)–(13.55).

The function h_2 that yields the second terms is easy to find; it is simply

$$h_2 = -\int_{S^1} \mathrm{Re}\left[W(\phi^i)\right] dx^1. \tag{13.58}$$

The function h_1 that yields the first terms can be constructed as follows. The connected components of the loop space LM are classified by the fundamental group $\pi_1(M)$. We choose and fix a loop, a base loop, in each connected component. Let us pick a component and denote the base loop there by ϕ_0. For a loop ϕ in that component we choose a homotopy $\widehat{\phi}$ that connects ϕ_0 to ϕ. Namely, $\widehat{\phi} = \widehat{\phi}(x^1, \tau)$ is a map from $S^1 \times [0,1]$ to M such that $\widehat{\phi}(x^1, 0) = \phi_0(x^1)$ and $\widehat{\phi}(x^1, 1) = \phi(x^1)$. Now, let us consider the area

$$h_1 = \int_{S^1\times[0,1]} \widehat{\phi}^*\omega \tag{13.59}$$

where ω is the Kähler form of M;

$$\omega = ig_{i\bar{j}}dz^i \wedge d\bar{z}^{\bar{j}}. \tag{13.60}$$

For a variation of $\widehat{\phi}$, the pull-back $\widehat{\phi}^*\omega$ changes by a total derivative

$$\delta\widehat{\phi}^*\omega = d\left(ig_{i\bar{j}}\delta\widehat{\phi}^i d\overline{\widehat{\phi}}^{\bar{j}} - ig_{i\bar{j}}d\widehat{\phi}^i\delta\overline{\widehat{\phi}}^{\bar{j}}\right). \tag{13.61}$$

and therefore the area changes by the boundary terms

$$\begin{aligned} \delta\int_{S^1\times[0,1]} \widehat{\phi}^*\omega &= \int_{S^1} \left\{-ig_{i\bar{j}}\delta\widehat{\phi}^i d\overline{\widehat{\phi}}^{\bar{j}} + ig_{i\bar{j}}d\widehat{\phi}^i\delta\overline{\widehat{\phi}}^{\bar{j}}\right\}\Big|_{\tau=0}^{\tau=1} \\ &= \int_{S^1} \left\{-ig_{i\bar{j}}\delta\phi^i d\overline{\phi}^{\bar{j}} + ig_{i\bar{j}}d\phi^i\delta\overline{\phi}^{\bar{j}}\right\}, \end{aligned} \tag{13.62}$$

where we have used the constraint that $\widehat{\phi}|_{\tau=0}$ is fixed to be ϕ_0 and thus $\delta\widehat{\phi}|_{\tau=0} = 0$. In particular, for a fixed loop ϕ the functional $h_1 = \int \widehat{\phi}^*\omega$ does not change for a deformation of the homotopy $\widehat{\phi}$. Namely, h_1 is a locally well-defined function on the loop space LM. Now, if we look at Eq. (13.62), we see that h_1 yields exactly the first terms of the required variation in Eq. (13.54) and Eq. (13.55). Thus, we can take as the function h the sum of h_1 and h_2;

$$h = \int_{S^1\times[0,1]} \widehat{\phi}^*\omega - \int_{S^1} \mathrm{Re}\left[W(\phi^i)\right] dx^1. \tag{13.63}$$

To be precise, this function can change if we change the homotopy class of $\widehat{\phi}$. Let us see how it changes by taking another homotopy $\widehat{\phi}' : S^1 \times [0,1] \to M$. The difference in h is

$$\begin{aligned}\Delta h &= \int_{S^1\times[0,1]} \widehat{\phi}'^*\omega - \int_{S^1\times[0,1]} \widehat{\phi}^*\omega \\ &= \int_{S^1\times S^1} \widetilde{\phi}^*\omega\end{aligned} \tag{13.64}$$

where $\widetilde{\phi}$ is a map $S^1 \times S^1 \to M$ obtained by gluing $\widehat{\phi}$ to $\widehat{\phi}'$ with the orientation of $\widehat{\phi}$ being reversed.[1] Thus, the function h is not a single-valued function on LM if there is a 2-cycle in M on which the Kähler class $[\omega]$ has a nonzero period. One can, however, make it single-valued on a certain covering space of LM. The relevant covering space can be identified with the set of maps $\widehat{\phi} : S^1 \times [0,1] \to M$ with $\widehat{\phi}(x^1, 0) = \phi_0(x^1)$ modulo the following equivalence relation: $\widehat{\phi} \equiv \widehat{\phi}'$ if and only if $\widehat{\phi} = \widehat{\phi}'$ at $\tau = 1$ and $\widehat{\phi}$ can be continuously deformed to $\widehat{\phi}'$. We denote this covering space by $\widetilde{LM}$. In such a situation, we first quantize the covering space $\widetilde{LM}$ and then project to the wave-functions invariant under the action of the covering group. Over the component of contractible loops where the base loop ϕ_0 is chosen to be a constant map to a point, $\phi_0(x^1) = * \in M$, the covering group is canonically isomorphic to the second homotopy group $\pi_2(M, *)$.

13.3.1. Non-Linear Sigma Models. Let us first consider the non-linear sigma model on a compact connected Kähler manifold M, with the superpotential set equal to zero, $W = 0$. We wish to find the number of supersymmetric ground states of this theory.

Due to the cohomological characterization, the spectrum of ground states does not change under the deformation of the function $h = \int_{S^1\times[0,1]} \widehat{\phi}^*\omega$. As

[1] If we parametrize the second S^1 by $\tau \in [0,2]$ with an identification $\tau = 0 \equiv 2$, the glued map is given by $\widetilde{\phi}(x^1, \tau) = \widehat{\phi}'(x^1, \tau)$ for $\tau \in [0,1]$ and $\widehat{\phi}(x^1, 2-\tau)$ for $\tau \in [1,2]$.

we have done in SQM, we rescale h by a large number, or equivalently, we consider the case where the Kähler form ω is taken to be large — the large volume limit. Then the ground state wave-functions are localized near the critical points of h.

The critical points of h are found by solving the equations

$$\frac{\delta h}{\delta \phi^i} = -ig_{i\bar{j}}\partial_1\overline{\phi^j} = 0, \quad \frac{\delta h}{\delta \overline{\phi^{\bar{\imath}}}} = ig_{j\bar{\imath}}\partial_1\phi^j = 0. \tag{13.65}$$

The solutions are obviously the constant maps $\phi : S^1 \to$ a point $\in M$ (which belong to the trivial component of LM). Thus, the critical point set is the space of constant maps, which is isomorphic to the target space manifold M itself. In the covering space $\widetilde{LM}$ the critical point set is the union of copies of M that are permuted by the covering group $\pi_2(M, *)$. The function h is not non-degenerate. To see if it is non-degenerate in the normal directions (i.e., Bott–Morse in the sense of Sec. 10.5.5), let us examine the Hessian of h. The Hessian at a constant loop is given by the second derivative

$$\delta_1\delta_2 h = i\int_{S^1} \left(g_{i\bar{j}}\delta_1\overline{\phi^j}d\delta_2\phi^i - g_{i\bar{j}}\delta_1\phi^i d\delta_2\overline{\phi^j}\right). \tag{13.66}$$

Thus, it is zero only if $d\delta\phi^I = 0$, namely only if the variation is tangent to the constant map locus. The function h is indeed Bott–Morse and the argument of Sec. 10.5.5 applies.

If we coordinatize the loop as

$$\phi^i = \sum_{n\in\mathbb{Z}} z_n^i\, e^{inx^1}, \quad \overline{\phi^{\bar{\imath}}} = \sum_{n\in\mathbb{Z}} \overline{z}_n^{\bar{\imath}}\, e^{-inx^1}, \tag{13.67}$$

the directions where the Hessian is negative definite are spanned by $(z_m^i, \overline{z}_m^{\bar{\imath}})$ with $m > 0$. Thus, the perturbative ground state at a constant loop is given by Eq. (10.304)

$$|\omega\rangle = e^{-\sum_{m\neq 0} m\|z_m\|^2}\omega \wedge d^{2n}z_1 \wedge d^{2n}z_2 \wedge \cdots \tag{13.68}$$

where ω is a harmonic form of $(z_0^i, \overline{z}_0^{\bar{\imath}})$ and $d^{2n}z_m$ is $dz_m^1 \wedge d\overline{z}_m^{\bar{1}} \wedge \cdots \wedge dz_m^n \wedge d\overline{z}_m^{\bar{n}}$. The question is whether this differential form glues together to define a differential form on LM around M. For this we need the negative normal bundle (the bundle of tangent vectors on which the Hessian is negative definite) to be orientable. In the present case it is indeed orientable since multiplication by i on the holomorphic coordinate induces a canonical orientation. Thus, we expect that we can find $|\omega\rangle$ as a well-defined differential form on LM

around M, which is a perturbative supersymmetric ground state if ω is a harmonic form of M.

We recall that the critical point set of h in $\widetilde{LM}$ consists of many connected components, each of which is a copy of M. If we require the invariance under the covering group action, we may focus on only one copy of M and we expect that the ground states can be identified as the harmonic forms on M. However, as we saw in SQM, it is in general possible that instanton effects lift the ground state degeneracy. To see whether there is such an instanton effect, let us compute the relative Morse index between different copies of M in $\widetilde{LM}$. Thus, we choose a path in the loop space LM that

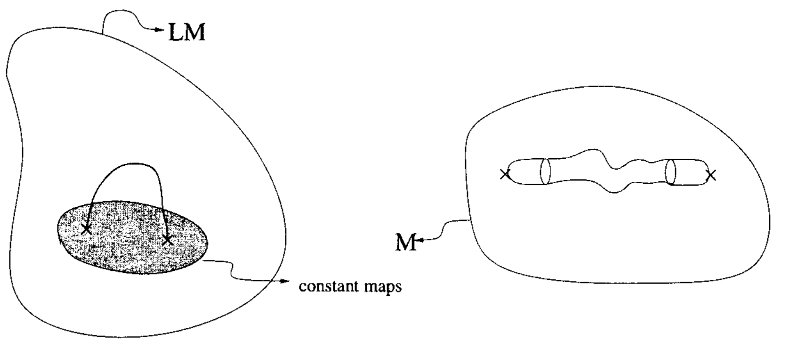


FIGURE 1. The path in the loop space LM connecting two trivial loops. It corresponds to a two-sphere mapped to M. If the map is homotopically non-trivial the lift of the path in the covering space $\widetilde{LM}$ connects different copies of M

starts at a constant loop $x \in M$ and end on another constant loop $y \in M$. (See Fig. 1.) This yields a trajectory of S^1's that shrinks at the two ends: namely, a map $\widetilde{\phi}$ of the two -sphere S^2 to M which maps the two tips (say the north and south poles) to x and y. If the map $\widetilde{\phi} : S^2 \to M$ defines a non-trivial homotopy class in $\pi_2(M)$, this path lifts to a path in $\widetilde{LM}$ that connects different copies of M. Now what is the relative Morse index? We can use here the relation of the relative Morse index and the index of the Dirac operator for fermions, which was noted in Sec. 10.5.2. The relevant Dirac operator here is the one acting on ψ_-^i and $\overline{\psi_+^i}$. The index of this operator is given by

$$\text{index } = 2\int_{S^2} \mathrm{ch}(\widetilde{\phi}^* TM^{(1,0)})\widehat{A}(S^2) = 2\int_{S^2} \widetilde{\phi}^* c_1(M). \tag{13.69}$$

If M is a Calabi–Yau manifold, $c_1(M) = 0$, then the relative Morse index vanishes. This means that all the critical submanifolds have the same Morse index. Then applying the procedure described in Sec. 10.5.5, we find that there is no non-trivial instanton effect. Thus, the perturbative ground states remain as the exact ground states. In fact, this is true even if M is not Calabi–Yau.[2] The reason is that our function $h = h_1$ is the moment map associated with a $U(1)$ action on the loop space LM. Note that the loop space is an infinite dimensional Kähler manifold whose Kähler form is given by

$$\widehat{\omega}(\delta_1\phi, \delta_2\phi) = \int_{S^1} i g_{i\bar{j}}(\delta_1\phi^i\delta_2\overline{\phi^j} - \delta_2\phi^i\delta_1\overline{\phi^j})dx^1. \tag{13.70}$$

The shift of the domain parameter x^1, $\phi(x^1) \to \phi(x^1 + \Delta)$, preserves the metric of LM as well as the above Kähler form. The tangent vector field generating this action is $v^I = \partial_1\phi^I$, and we find

$$i_v\widehat{\omega}(\delta\phi) = \int_{S^1} i g_{i\bar{j}}(\partial_1\phi^i\delta\overline{\phi^j} - \delta\phi^i\partial_1\overline{\phi^j})dx^1 = \delta h, \tag{13.71}$$

where (13.54) and (13.54) with $W = 0$ are used in the last step. Thus, h is indeed the moment map associated with the $U(1)$ action. Applying the result of Ch 10.5.6, we find that there is no non-trivial instanton effect that lifts the perturbative ground states.

Thus, we conclude that *the supersymmetric ground states are in one-to-one correspondence with the harmonic forms on* M. What are the quantum numbers (i.e., charges) of a ground state? The Q-complex is graded by the Morse index. However, as we have seen above, the relative Morse index can be nonzero (if $c_1(M) \neq 0$) even between the same point of M. This shows that the Morse index is well defined only modulo some integer. In the case where $\int_{S^2} \widetilde{\phi}^* c_1(M)$ can take arbitrary integer values, the Morse index is well-defined mod 2; if it can take only integer multiples of $p \in \mathbb{Z}$, then the Morse index is well-defined mod $2p$. Since the Q-complex is graded by the axial R-symmetry, this of course reflects the axial R-anomaly. On the other hand, the vector R-symmetry is not anomalous and the corresponding quantum number must be well defined. For the ground state corresponding

[2]We will find an alternative derivation of this fact in Ch. 16.

to a harmonic (p,q)-form, the vector R-charge is given by Eqs. (13.30)–(13.31).

$$q_V = -p + q. \tag{13.72}$$

In the case $c_1(M) = 0$, where the target space M is a Calabi–Yau manifold, the relative Morse index is well defined (in $\mathbb{Z}$). This corresponds to the existence of $U(1)_A$ axial R-symmetry or the conservation of the axial R-charge F_A. We fix the zero of the Morse index by requiring the invariance of the spectrum under the "CPT conjugation" (which requires that for every state in the Hilbert space there should be a conjugate state with opposite charge) that acts on F_A as $F_A \to -F_A$. Then the axial R-charge of the ground state corresponding to the harmonic (p,q)-form is

$$q_A = p + q - \dim_{\mathbb{C}} M. \tag{13.73}$$

13.3.2. Ground States of the LG Model. Let us consider an LG model with a non-trivial superpotential $W(\Phi^i)$. We assume that $W(\Phi^i)$ has isolated and non-degenerate critical points only. Here we will show that the number of ground states is in one-to-one correspondence with the number of critical points, just as we found for the corresponding one-dimensional QFT.

The equation for a critical point of h is

$$\frac{d\phi^i}{dx^1} = -\frac{i}{2} g^{i\bar{j}} \partial_{\bar{j}} \overline{W}. \tag{13.74}$$

The above equations imply

$$\frac{dW}{dx^1} = \partial_i W \frac{d\phi^i}{dx^1} = -\frac{i}{2} g^{i\bar{j}} \partial_i W \partial_{\bar{j}} \overline{W} = -\frac{i}{2} |\partial W|^2. \tag{13.75}$$

Integrating over the circle S^1, we obtain

$$-\frac{i}{2} \int_{S^1} dx^1 |\partial W|^2 = \int_{S^1} \frac{dW}{dx^1} dx^1 = 0, \tag{13.76}$$

where we have used the periodic boundary condition along S^1. This shows that $\partial_i W = 0$ everywhere on the circle S^1, which implies that ϕ is the constant map to a critical point of W. Since the (mod 2) Morse index is constant as in SQM, there is no room for instanton effects that lift the ground state degeneracy. Therefore, *the ground states are in one-to-one correspondence with the critical points of the superpotential.*

13.4. Supersymmetric Sigma Model on T^2 and Mirror Symmetry

In this section we show how the T-duality discussed in the context of bosonic sigma models, can be extended to the supersymmetric case. This leads to the first (and most basic) example of mirror symmetry.

13.4.1. The Spectrum and Supersymmetric Ground States. Let us consider the supersymmetric sigma model on T^2. For simplicity, we consider the rectangular metric on T^2 with radius R_1 and R_2 and we set $B = 0$, but this assumption is not essential for what we will show here. The model is described by a chiral superfield Φ representing flat coordinates of T^2. In particular the lowest component ϕ has periodicity

$$\phi \equiv \phi + 2\pi R_1 \equiv \phi + 2\pi R_2 i. \tag{13.77}$$

The action is given by

$$S = \frac{1}{4\pi} \int d^2x \int d^4\theta \, \overline{\Phi}\Phi. \tag{13.78}$$

In terms of the component fields ϕ, $\psi_\pm$ and $\overline{\psi}_\pm$ the action is expressed as

$$S = \frac{1}{4\pi} \int \left\{ |\partial_0 \phi|^2 - |\partial_1 \phi|^2 + i\overline{\psi}_-(\partial_0 + \partial_1)\psi_- + i\overline{\psi}_+(\partial_0 - \partial_1)\psi_+ \right\} d^2x. \tag{13.79}$$

Now we see that the system consists of the free bosonic sigma model on T^2 plus the free theory of a Dirac fermion — which are decoupled from each other. The bosonic sigma model is identical to the one considered in Sec. 11.1. The fermion does not know about the periodicity of the coordinates ϕ and is nothing but the free system analyzed in detail in Sec. 11.3 (up to a field normalization $\psi_\pm \to \sqrt{2}\psi_\pm$ that has no effect). Accordingly, the Hilbert space is the tensor product of the Hilbert spaces of the bosonic and fermionic systems

$$\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_F. \tag{13.80}$$

The Hamiltonian and momentum are the sums of those for the corresponding systems

$$H = H_B + H_F, \tag{13.81}$$

$$P = P_B + P_F. \tag{13.82}$$

Since $c_1(T^2) = 0$, the $U(1)_A$ R-symmetry is preserved, as well as the $U(1)_V$ R-symmetry. They act trivially on the bosonic component ϕ, and therefore the R-charges F_V and F_A are the same as those for the fermionic system:

$$F_V = \frac{1}{2\pi}\int_{S^1} (\overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+)\, dx^1, \tag{13.83}$$

$$F_A = \frac{1}{2\pi}\int_{S^1} (-\overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+)\, dx^1. \tag{13.84}$$

These can be expressed in terms of the oscillator modes as in Eqs. (11.136)–(11.137). The states in $\mathcal{H}$ are constructed by acting with the oscillator modes α^i_n, $\widetilde{\alpha}^i_n$ $(i = 1, 2)$ and ψ_n, $\overline{\psi}_n$, $\overline{\widetilde{\psi}}_n$ and $\widetilde{\psi}_n$ on the states

$$|\vec{l}, \vec{m}\rangle := |l_1, l_2, m^1, m^2\rangle_B \otimes |0\rangle_F. \tag{13.85}$$

Here $|l_1, l_2, m^1, m^2\rangle_B$ is the state with momentum $\vec{l} = (l_1, l_2)$ and winding number $\vec{m} = (m^1, m^2)$ which is annihilated by the positive frequency modes (α^i_n, $\widetilde{\alpha}^i_n$ with $n \geq 1$), while $|0\rangle_F$ is the state given by Eq. (11.128) annihilated by the positive frequency modes of $\psi_\pm$ and $\overline{\psi}_\pm$, and by half of the zero modes $\overline{\psi}_0$ and $\widetilde{\psi}_0$. There are four lowest-energy states

$$\begin{array}{ccc} & \psi_0\overline{\widetilde{\psi}}_0|\vec{0},\vec{0}\rangle & \\ \psi_0|\vec{0},\vec{0}\rangle & & \overline{\widetilde{\psi}}_0|\vec{0},\vec{0}\rangle \\ & |\vec{0},\vec{0}\rangle, & \end{array} \tag{13.86}$$

with R-charges

$$q_V = \begin{array}{ccc} & 0 & \\ -1 & & 1 \\ & 0 & \end{array} \quad q_A = \begin{array}{ccc} & 1 & \\ 0 & & 0 \\ & -1 & \end{array} \tag{13.87}$$

and energy

$$E_0 = \left(-\frac{1}{12}\right) \times 2 + \frac{1}{6} = 0. \tag{13.88}$$

Since these are the zero energy states, they are the supersymmetric ground states. We note here that these supersymmetric states take the form shown in Eq. (13.68) that is obtained by the semi-classical method. Indeed, $\vec{m} = \vec{0}$ shows that the states are in the component of the contractible loops. The bosonic piece $|0, 0, 0, 0\rangle_B$ is identified as the wave-function

$$|0,0,0,0\rangle_B \leftrightarrow \Psi(z_n, \overline{z}_n) = \exp\left(-\sum_{m\neq 0} |z_m|^2\right). \tag{13.89}$$

The fermionic piece $|0\rangle_F$ interpreted as Dirac's sea in Eq. (11.143) can be identified as the $\frac{\infty}{2}$-form

$$|0\rangle_F \leftrightarrow d^2z_1 \wedge d^2z_2 \wedge d^2z_3 \wedge \cdots \tag{13.90}$$

under the identification $\psi_n \leftrightarrow dz_n$, $\overline{\widetilde{\psi}}_n \leftrightarrow d\overline{z}_n$. The four states from Eq. (13.86) are then identified as the state $|\omega\rangle$ in Eq. (13.68) with

$$\omega = \begin{array}{ccc} & dz \wedge d\overline{z} & \\ dz & & d\overline{z} \\ & 1. & \end{array} \tag{13.91}$$

Notice that the R-charges in Eq. (13.87) obtained by the exact quantization agree with the result in Eqs. (13.72)–(13.73) obtained by the semi-classical method plus CPT invariance.

13.4.2. T-duality. Let us perform T-duality on the second circle of T^2. This inverts the radius R_2 to $R_2' = 1/R_2$ and therefore the dual field ϕ' has periodicity

$$\phi \equiv \phi' + 2\pi R_1 \equiv \phi' + (2\pi/R_2)i. \tag{13.92}$$

It is related to the original field ϕ by $\mathrm{Re}\,\phi = \mathrm{Re}\,\phi'$ and

$$\partial_+\mathrm{Im}\,\phi = \partial_+\mathrm{Im}\,\phi', \tag{13.93}$$

$$\partial_-\mathrm{Im}\,\phi = -\partial_-\mathrm{Im}\,\phi'. \tag{13.94}$$

In terms of the complex variables, the relation is

$$\partial_+\phi = \partial_+\phi', \tag{13.95}$$

$$\partial_-\phi = \partial_-\overline{\phi}'. \tag{13.96}$$

On the other hand, we do not touch the fermions. Since T-duality is an equivalence of theories, the dual theory also has $(2, 2)$ supersymmetry. The supercharges are expressed as

(13.97)
$$\begin{aligned} Q_+ &= \tfrac{1}{2\pi}\textstyle\int \psi_+\partial_+\overline{\phi} = \tfrac{1}{2\pi}\int \psi_+\partial_+\overline{\phi}', & \overline{Q}_+ &= \tfrac{1}{2\pi}\textstyle\int \overline{\psi}_+\partial_+\phi = \tfrac{1}{2\pi}\int \overline{\psi}_+\partial_+\phi', \\ Q_- &= \tfrac{1}{2\pi}\textstyle\int \psi_-\partial_-\overline{\phi} = \tfrac{1}{2\pi}\int \psi_-\partial_-\phi', & \overline{Q}_- &= \tfrac{1}{2\pi}\textstyle\int \overline{\psi}_-\partial_-\phi = \tfrac{1}{2\pi}\int \overline{\psi}_-\partial_-\overline{\phi}'. \end{aligned}$$

We notice that they take the standard form of the supercharges if we denote $\psi_+ = \psi'_+$, $\psi_- = \overline{\psi}'_-$, $\overline{\psi}_+ = \overline{\psi}'_+$ and $\overline{\psi}_- = \psi'_-$ and also

$$Q_+ = Q'_+, \quad \overline{Q}_+ = \overline{Q}'_+, \tag{13.98}$$

$$Q_- = \overline{Q}'_-, \quad \overline{Q}_- = Q'_-. \tag{13.99}$$

Also, the R-symmetry generators are

$$F_V = \frac{1}{2\pi}\int_{S^1} \left(-\overline{\psi}'_-\psi'_- + \overline{\psi}'_+\psi'_+\right) dx^1 = F'_A, \tag{13.100}$$

$$F_A = \frac{1}{2\pi}\int_{S^1} \left(\overline{\psi}'_-\psi'_- + \overline{\psi}'_+\psi'_+\right) dx^1 = F'_V. \tag{13.101}$$

These mean that under T-duality, the supercharges Q_- and $\overline{Q}_-$ as well as $U(1)_V$ and $U(1)_A$ R-symmetries are exchanged with each other. Thus, we have shown that *T-duality is a mirror symmetry*.

The above change of notation yields the change of notation $\psi_n = \overline{\psi}'_n$, $\overline{\psi}_n = \psi'_n$, $\widetilde{\psi}_n = \widetilde{\psi}'_n$ and $\overline{\widetilde{\psi}}_n = \overline{\widetilde{\psi}}'_n$ for the oscillator modes. In particular, the state $|\vec{l},\vec{m}\rangle'$ is annihilated by all the positive frequency modes and two zero modes $\overline{\psi}_0 = \psi'_0$, $\widetilde{\psi}_0 = \widetilde{\psi}'_0$. Thus, it is appropriate to write it in the dual theory as

$$|\vec{l},\vec{m}\rangle = \psi'_0|\vec{l}',\vec{m}'\rangle', \tag{13.102}$$

where the momentum and winding number for the second circle are exchanged

$$\vec{l}' = (l_1, m^2),\ \ \vec{m}' = (m^1, l_2). \tag{13.103}$$

The four ground states shown in Eq. (13.86) are then written as

$$\begin{array}{ccc} & -\overline{\widetilde{\psi}}'_0|\vec{0},\vec{0}\rangle' & \\ |\vec{0},\vec{0}\rangle' & & -\psi'_0\overline{\widetilde{\psi}}'_0|\vec{0},\vec{0}\rangle' \\ & \psi'_0|\vec{0},\vec{0}\rangle'. & \end{array} \tag{13.104}$$

The R-charges of these states are

$$q'_V = \begin{array}{ccc} & 1 & \\ 0 & & 0 \\ & -1 & \end{array} \quad q'_A = \begin{array}{ccc} & 0 & \\ -1 & & 1 \\ & 0. & \end{array} \tag{13.105}$$

Indeed, the axial and vector R-charges are exchanged, $q_V = q'_A$ and $q_A = q'_V$.

Path-integral Derivation. One can repeat the path-integral derivation of T-duality shown in Sec. 11.2 for the superfields. For the superspace calculus used here, see Sec. 12.1.3. We start with the following Lagrangian for a real superfield B and the chiral superfield Φ,

$$L' = \int d^4\theta \left(-\frac{1}{4}B^2 + \frac{1}{2}(\Phi + \overline{\Phi})B\right). \tag{13.106}$$

We first integrate out the real superfield B. Then B is solved by

$$B = \Phi + \overline{\Phi}. \tag{13.107}$$

Inserting this into L' we obtain

$$L = \int d^4\theta\, \frac{1}{2}\overline{\Phi}\Phi, \tag{13.108}$$

which is the Lagrangian for the supersymmetric sigma model on T^2 with radius (R_1, R_2). Now, reversing the order of integration, we consider integrating out Φ and $\overline{\Phi}$ first. This yields the following constraint on B,

$$\overline{D}_+\overline{D}_-B = D_+D_-B = 0, \tag{13.109}$$

which is solved by

$$B = \Theta + \overline{\Theta}, \tag{13.110}$$

where Θ is a twisted chiral superfield of periodicity

$$\Theta \equiv \Theta + 2\pi R_1 \equiv \Theta + (2\pi/R_2)i. \tag{13.111}$$

Inserting this into the original Lagrangian we obtain

$$\widetilde{L} = \int d^4\theta \left(-\frac{1}{2}\overline{\Theta}\Theta\right), \tag{13.112}$$

which is the Lagrangian for the supersymmetric sigma model on a torus of radius $(R_1, 1/R_2)$. This time, however, the complex coordinate is described by the twisted chiral superfield Θ. This is another manifestation of mirror symmetry. The two theories are equivalent without the exchange of the supercharges Q_- and $\overline{Q}_-$ (see the remark at the end of Sec. 12.4). The supercharges given by Eq. (13.97) have the right expression in terms of the twisted chiral superfield $\Theta = \phi' + \theta^+\overline{\psi}'_+ + \overline{\theta}^-\psi'_- + \cdots$, where the fermions $\psi'_\pm$, $\overline{\psi}'_\pm$ are related to $\psi_\pm$, $\overline{\psi}_\pm$ simply by the renaming $\psi'_\pm = \pm\overline{\psi}_\pm$ and $\overline{\psi}'_\pm = \pm\psi_\pm$. This renaming is dictated by the relation

$$\Phi + \overline{\Phi} = \Theta + \overline{\Theta}, \tag{13.113}$$

which follows from Eqs. (13.107)–(13.110).

CHAPTER 14

Renormalization Group Flow

We are now in a position to study one of the most important aspects of quantum field theory. This is the fact that the behavior of a theory depends on the scale. What one means by this is how the expectation values of fields vary as a function of the distance between fields, or equivalently under rescaling of the metric on the manifold over which the quantum field theory is defined. Quite often, their behavior at long distances is very different from their behavior at short distances and often one introduces a new set of fields at long distances which give a more useful description of the theory. In this section, we will see such a change of behavior and description in the non-linear sigma models and the Landau–Ginzburg models. In particular, we will see that the target space metric changes as a function of the scale. In supersymmetric field theories, however, there are certain quantities that do not depend on the scale. The superpotential in a Landau–Ginzburg model is one such object. This is the famous non-renormalization theorem of the superpotential. We present the proof of this theorem and its generalizations.

14.1. Scales

Let us consider the correlation function of operators

$$\langle \mathcal{O}_1(x_1)\cdots\mathcal{O}_s(x_s)\rangle \tag{14.1}$$

of a quantum field theory formulated on a Euclidean plane. We are interested in how this function behaves at various scales, or how the behavior changes as we change the scale. Here what we mean by "scale" is the average distance between the insertion points, $|x_i - x_j|$.

A change of scale can be implemented by a scale transformation

$$x_i^\mu \longrightarrow \lambda x_i^\mu, \tag{14.2}$$

where λ is a nonzero constant. If we take $\lambda > 1$, we change the scale to longer distances while $\lambda < 1$ corresponds to shorter distances. There is an equivalent way to perform a scale transformation that is applicable to a more general setting. Let us consider a worldsheet Σ with a metric $h_{\mu\nu}$. Then the notion of distance is defined with respect to $h_{\mu\nu}$. The correlation function depends on the metric and we denote the dependence as a subscript $\langle \mathcal{O}_1(x_1)\cdots\mathcal{O}_s(x_s)\rangle_h$. Then the scale transformation is implemented by

$$h_{\mu\nu} \longrightarrow \lambda^2 h_{\mu\nu}. \tag{14.3}$$

On the Euclidean plane, it is easy to see that the two transformations, Eqs. (14.2)-(14.3), are equivalent: $\langle \mathcal{O}_1(\lambda x_1)\cdots\mathcal{O}_s(\lambda x_s)\rangle_h = \langle \mathcal{O}_1(x_1)\cdots\mathcal{O}_s(x_s)\rangle_{\lambda^2 h}$

As a convention, we will refer to extremely short distances as *ultraviolet* while extremely long distances will be called *infrared*. This terminology has its origin in the electromagnetic waves which behave as $\mathrm{Re}\, e^{ik(t-x)}$ where t is the time coordinate and x is a spatial coordinate. The phase $e^{ik(t-x)}$ rotates once in the distance

$$\lambda_k = 2\pi/k \tag{14.4}$$

in the x or t direction. This length is called the wavelength of the wave $\mathrm{Re}\, e^{ik(t-x)}$. The electromagnetic wave with its wavelength in a certain range is a visible light. It is violet near the shorter and red near the longer wavelengths of the range. This is the origin of the terminology. k is called *frequency* since it counts how frequently the phase rotates over a given distance or time. Thus, a long wavelength corresponds to low frequency (red) and a short wavelength corresponds to high frequency (violet).

In quantum field theory, scattering amplitudes of particles are interesting objects to study (although we do not treat them here). They are obtained from the correlation functions, such as Eq. (14.1), essentially by performing the Fourier transform of the coordinates $x_1,\ldots,x_s$:

$$S(p_1,\ldots,p_s) = \int [\cdots]\langle \mathcal{O}_1(x_1)\cdots\mathcal{O}_s(x_s)\rangle \prod_{i=1}^{s} e^{ip_i x_i} d^2x_i, \tag{14.5}$$

where $[\cdots]$ may contain differential operators in the x_i. This represents the scattering amplitude of s particles, and the frequencies $p_1,\ldots,p_s$ represent the energy-momenta of the particles. As usual in a Fourier transform or as Eq. (14.4) suggests, high (resp. low) energy behavior of the scattering amplitude corresponds to short (resp. long) distance behavior of the correlation functions. In the terminology introduced above, very high energy corresponds to ultraviolet and very low energy corresponds to infrared.

14.2. Renormalization of the Kähler Metric

Let us consider the supersymmetric non-linear sigma model on a Kähler manifold M with metric g. In the previous sections, we have been working on the worldsheet with a flat normalized metric, say $\delta_{\mu\nu}$ in the case of Euclidean signature. Now consider a general worldsheet metric, $h_{\mu\nu}$. The classical action can be written as

(14.6)
$$S = \int_\Sigma \left\{ g_{i\bar{j}} h^{\mu\nu}\partial_\mu\phi^i\partial_\nu\overline{\phi^j} + i g_{i\bar{j}}\overline{\psi^j}\gamma^\mu D_\mu\psi^i + R_{i\bar{j}k\bar{l}}\psi_+^i\psi_-^k\overline{\psi_-^j}\,\overline{\psi_+^l} \right\} \sqrt{h}\, d^2x.$$

Consider rescaling the worldsheet metric

$$h_{\mu\nu} \to \lambda^2 h_{\mu\nu}. \tag{14.7}$$

The gamma matrices transform as

$$\gamma^\mu \to \lambda^{-1}\gamma^\mu \tag{14.8}$$

since they obey the relations $\{\gamma^\mu,\gamma^\nu\} = -2h^{\mu\nu}$. Then the action is invariant under this rescaling provided the fermionic fields are transformed as

$$\psi_\pm \to \sqrt{\lambda}^{-1}\psi_\pm,\ \overline{\psi}_\pm \to \sqrt{\lambda}^{-1}\overline{\psi}_\pm, \tag{14.9}$$

while the bosons ϕ^i are kept intact. Thus the scale transformation from Eq. (14.7), or "dilatation," is a classical symmetry of the theory. The question is: Is it a symmetry of the quantum theory? In other words, is it a symmetry of the correlation functions of quantum field theory?

14.2.1. The Kähler Class.

To examine this question let us see whether the correlation functions on a torus T^2 are scale invariant. Consider the correlation function of some combination of ψ_-^i's and $\overline{\psi_+^j}$'s;

$$f(h,g) := \langle (\psi_-)^k(\overline{\psi}_+)^k\rangle_h. \tag{14.10}$$

Here h and g stand for the metrics of the worldsheet torus T^2 and the target space M respectively. This correlation function may also depend on the insertion points of ψ_-^i's and $\overline{\psi_+^j}$'s, as in Sec. 13.2.2, but we omit dependence in the notation $f(h,g)$ as it is irrelevant in our discussion. We saw in Eqs. (13.39) and (13.41), in the context of the axial anomaly, that this correlation function is generically non-vanishing when

$$k = \int_{T^2} \phi^* c_1(M), \tag{14.11}$$

for some $\phi : T^2 \to M$. Now assume that $h_{\mu\nu}$ is a flat metric and the inserted operator commutes with the supercharge $Q = \overline{Q}_+ + Q_-$,

$$[Q, (\psi_-)^k(\overline{\psi}_+)^k] = 0. \tag{14.12}$$

The correlation function has some very special properties that we will systematically explore when we discuss topological field theory in Ch. 16 (for the moment we take them as facts). One special property is dependence on the worldsheet metric: the correlation function $\langle(\psi_-)^k(\overline{\psi}_+)^k\rangle_h$ is invariant under the scaling $h \to \lambda^2 h$. Since the scale transformation acts on the fermionic fields as $\psi_\pm \to \sqrt{\lambda}^{-1}\psi_\pm$ and $\overline{\psi}_\pm \to \sqrt{\lambda}^{-1}\overline{\psi}_\pm$ as shown by Eq. (14.9), this means that

$$f(h, g) = f(\lambda^2 h, g) \cdot \lambda^k. \tag{14.13}$$

Another property is that it receives contributions only from holomorphic maps $\phi : T^2 \to M$, and the correlation function can be written as

$$f(h, g) = n_h \, e^{-A_g}. \tag{14.14}$$

Here n_h is a number depending only on h, and A_g is the area of the image $\phi(T^2)$ measured by the metric g. Combining the two properties Eqs. (14.13)–(14.14), we find the relation

$$f(h, g) = f(\lambda^2 h, g)\lambda^k = n_{\lambda^2 h}\, e^{-(A_g - k\log\lambda)}.$$

This means that

$$f(h, g) = f(\lambda^2 h, g'), \tag{14.15}$$

for a metric g' such that $A_{g'} = A_g - k\log\lambda$. Thus, under the scaling $h_{\mu\nu} \to \lambda^2 h_{\mu\nu}$, the metric must be changed as $g \to g'$ in order for the correlation function to remain the same. Namely, the scale transformation effectively changes the metric on M so that the area changes as

$$A_g \to A_g - k\log\lambda. \tag{14.16}$$

The area is expressed as

$$\begin{aligned} A_g &= \int_{T^2} g_{i\bar{j}} h^{\mu\nu}\partial_\mu\phi^i\partial_\nu\overline{\phi}^{\bar{j}}\sqrt{h}d^2x \\ &= \int_{T^2} g_{i\bar{j}}(\partial_z\phi^i\partial_{\bar{z}}\overline{\phi}^{\bar{j}} + \partial_{\bar{z}}\phi^i\partial_z\overline{\phi}^{\bar{j}})idz\wedge d\bar{z} \\ &= \int_{T^2} 2g_{i\bar{j}}\partial_{\bar{z}}\phi^i\partial_z\overline{\phi}^{\bar{j}}idz\wedge d\bar{z} + \int_{T^2}\phi^*\omega \end{aligned} \tag{14.17}$$

where ω is the Kähler form of M,

$$\omega = ig_{i\bar{j}}dz^i \wedge d\bar{z}^{\bar{j}}. \tag{14.18}$$

For a holomorphic map $\partial_{\bar{z}}\phi^i = 0$, this area is given by

$$A_g = \int_{T^2}\phi^*\omega. \tag{14.19}$$

We note that the integer k also has a similar integral expression shown in Eq. (14.11). Therefore the effect shown in Eq. (14.16) of the scale transformation $h_{\mu\nu} \to \lambda^2 h_{\mu\nu}$ is nothing but to change the Kähler class, $[\omega]$:

$$[\omega] \to [\omega] - (\log\lambda)c_1(M). \tag{14.20}$$

From these considerations, we see that the scale invariance of the classical system is broken in the quantum theory if the first Chern class $c_1(M)$ of M is non-vanishing. If the first Chern class is positive definite, the above result shows that the Kähler class becomes large as $h \to \lambda^2 h$ with $\lambda \ll 1$, namely at short distances on the worldsheet. In other words, the Kähler class becomes smaller at longer distances of the worldsheet. If the first Chern class vanishes $c_1(M) = 0$ (i.e., for Calabi–Yau manifolds), the Kähler class is not modified according to the change in the scale. Thus, the classical scale invariance is not broken only for Calabi–Yau sigma models.

Since the first Chern class $c_1(M)$ is represented by the Ricci form

$$c_1(M) = \frac{i}{2\pi}R_{i\bar{j}}dz^i \wedge d\bar{z}^{\bar{j}}, \tag{14.21}$$

Eq. (14.20) may suggest that the metric effectively changes under the worldsheet rescaling $h \to \lambda^2 h$ as

$$g_{i\bar{j}} \to g_{i\bar{j}} - \frac{\log\lambda}{2\pi}R_{i\bar{j}}. \tag{14.22}$$

One can see how the metric changes under the change of scale in an approximation scheme called *sigma model perturbation theory*. This is the topic of the next discussion. We will indeed see that the metric changes as Eq. (14.22) to first non-trivial order in this approximation.

14.2.2. Sigma Model Perturbation Theory. Let us consider the bosonic non-linear sigma model on a Riemannian manifold M with metric g. The model is described by bosonic scalar fields ϕ^I ($I = 1, \ldots, n = \dim M$)

that represent a map of the worldsheet to M. The classical action is given by

$$S = \frac{1}{2}\int g_{IJ}(\phi)\partial_\mu\phi^I\partial^\mu\phi^J d^2x, \tag{14.23}$$

where we have chosen a (conformally) Euclidean metric on the worldsheet. We expand the fields around a point $\phi_0^I \in M$,

$$\phi^I = \phi_0^I + \xi^I. \tag{14.24}$$

If the coordinate is chosen appropriately, the metric is expanded as

$$g_{IJ}(\phi) = \delta_{IJ} - \frac{1}{3}R_{IKJL}(\phi_0)\xi^K\xi^L + O(\xi^3). \tag{14.25}$$

The ξ-linear term is eliminated here by our choice of good coordinates (Riemann normal coordinates), but the ξ-bilinear term is proportional to the curvature at ϕ_0 and cannot be eliminated by a further change of coordinates. Thus, if M is not flat, the action

$$S = \frac{1}{2}\int\left(\partial^\mu\xi^I\partial_\mu\xi^I - \frac{1}{3}R_{IKJL}(\phi_0)\xi^K\xi^L\partial^\mu\xi^I\partial_\mu\xi^J + O(\xi^5)\right)d^2x \tag{14.26}$$

is not purely quadratic in any choice of variables. Namely, the system is interacting, where the non-quadratic terms are regarded as providing the interactions between ξ^I for different I's. To organize the interaction terms, let us consider rescaling the target space metric as

$$g_{IJ} \to t^2 g_{IJ}. \tag{14.27}$$

If we change the variables ξ^I as $\widetilde{\xi}^I = t\xi^I$, the metric is expressed by

$$\widetilde{g}_{IJ}(\phi) = \delta_{IJ} - \frac{1}{3t^2}R_{IKJL}(\phi_0)\widetilde{\xi}^K\widetilde{\xi}^L + O(\xi^3/t^3). \tag{14.28}$$

The interaction terms are small for large t and higher-order terms are smaller by powers of $1/t$. Thus, we can consider a systematic perturbation theory in powers of $1/t$. This is the large volume expansion of the non-linear sigma model.

In Ch. 9, we studied zero-dimensional QFTs as toy models, where we encountered integrals such as

$$Z(M,C) = \int\prod_{i=1}^n dX_i \, \exp\left(-\frac{1}{2}X_iM_{ij}X_j + C_{ijkl}X_iX_jX_kX_l\right) \tag{14.29}$$

and also the correlation function

$$\langle\mathcal{O}\rangle = \frac{1}{Z(M,C)} \times\int\prod_{i=1}^n dX_i \, \exp\left(-\frac{1}{2}X_iM_{ij}X_j + C_{ijkl}X_iX_jX_kX_l\right)\mathcal{O}(X_i,X_j,\ldots), \tag{14.30}$$

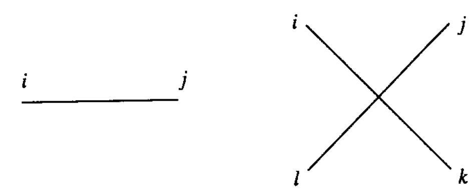
where $\mathcal{O}(X_i,X_j,\ldots)$ is some expression of $X_i, X_j, \ldots$. The perturbative expansion of the partition function $Z(M,C)$ and the correlation function $\langle\mathcal{O}\rangle$ is obtained by first expanding $e^{C_{ijkl}X_iX_jX_kX_l}$ as $\sum_{r=0}^\infty \frac{1}{r!}(C_{ijkl}X_iX_jX_kX_l)^r$ and computing the integral for each term. This leads to a diagrammatic evaluation of the integral based on the propagator (two-point functions at $C_{ijkl}=0$)

$$\langle X_iX_j\rangle_{(0)} = \frac{1}{Z(M,0)}\int\prod_{i=1}^n dX_i \, e^{-\frac{1}{2}X_kM_{kl}X_l}\,X_iX_j = (M^{-1})^{ij}, \tag{14.31}$$

which solves the equation

$$M_{ij}\langle X_jX_k\rangle_{(0)} = \delta_{ik}. \tag{14.32}$$

The diagrammatic computation is carried out by using this propagator and the interaction vertex $C_{ijkl}X_iX_jX_kX_l$. These are represented by the diagrams (1) and (2) in Fig. 1 respectively. The two point-function $\langle X_iX_j\rangle$



(1) (2)

FIGURE 1. (1) Propagator and (2) Vertex

and the four-point function $\langle X_iX_jX_kX_l\rangle$ can be computed by the diagrams of the form given in Fig. 2. The holes in each diagram are called the *loops* of the diagram. A diagram is called an L-loop diagram if it has L loops. For example, the first one of each series in Fig. 2 is the zero-loop diagram. The second of (A) and the second and third of (B) are the one-loop diagrams. (The third of (A) is one of the two-loop diagrams.) One can organize the sum over the diagrams by the number of loops. We denote the sum over s-loop

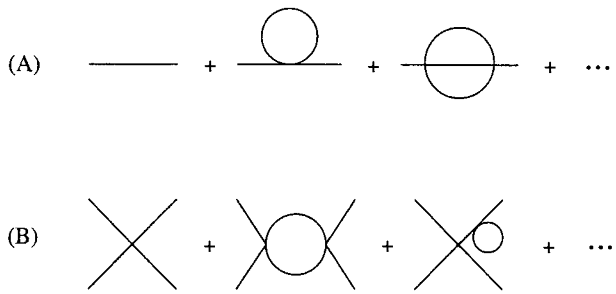


FIGURE 2. Feynman diagrams: (A) for two-point function $\langle X_i X_j \rangle$, (B) for four-point function $\langle X_i X_j X_k X_l \rangle$

diagrams with $s = 0, 1, \ldots, L$ in Fig. 2 by $\langle X_i X_j \rangle_{(L)}$ and $\langle X_i X_j X_k X_l \rangle_{(L)}$ respectively and call them the two- and four-point functions at the L-loop level. In the present example, the number of loops is the same as the number of C_{ijkl}'s up to a constant.

EXERCISE 14.2.1. *Compute $\langle X_i X_j \rangle_{(1)}$ and $\langle X_i X_j X_k X_l \rangle_{(1)}$, the two- and the four-point functions at the one-loop level.*

As in the above example, we can also consider the diagrammatic evaluation of path-integrals based on the propagator and the interaction vertex. The analogue of the matrix M in the present case is the Laplace operator $M = -\partial^\mu \partial_\mu$. Thus the propagator obeys the analogue of Eq. (14.32), namely

$$-\partial^\mu \partial_\mu \langle \xi^I(x) \xi^J(y) \rangle_{(0)} = \delta(x-y) \delta^{IJ} \tag{14.33}$$

which is solved by

$$\langle \xi^I(x) \xi^J(y) \rangle_{(0)} = \int \frac{d^2k}{(2\pi)^2} \frac{e^{ik(x-y)}}{k^2} \delta^{IJ}. \tag{14.34}$$

(We notice that the integral is logarithmically divergent at $k = 0$. This is the long-distance singularity which is special two dimensions. Here we leave it as it is. We will shortly make it finite by introducing a cut-off and later interpret what the manipulation means.) The leading term (in powers of the curvature) in the interaction vertex is given by

$$\frac{1}{6} \int d^2x \, R_{MKNL} \xi^K \xi^L \partial^\mu \xi^M \partial_\mu \xi^N. \tag{14.35}$$

Thus, this system including only this interaction vertex is almost identical to the toy model considered above. (The only difference is that the indices $i, j, k, \ldots$ in the present case run over infinitely many values.) Therefore, we can try to repeat what we have done there to obtain the correlation functions as power series in the Riemannian curvature, R_{IJKL}.

The two-point function at the one-loop level is obtained by summing the first and the second diagrams in Fig. 2 (A). It is straightforward to find

$$\langle \xi^I(x) \xi^J(y) \rangle_{(1)} = \int \frac{d^2p}{(2\pi)^2} \frac{e^{ip(x-y)}}{p^2} \left\{ \delta^{IJ} + \frac{1}{3} \int \frac{d^2k}{(2\pi)^2} \frac{1}{k^2} R_{IJ} \right\}. \tag{14.36}$$

The momentum integral in the second term is logarithmically divergent at large k, in addition to the divergence at small k.[1] The origin of the divergence at large k is clear if we look at the second diagram of Fig. 2 (A); it comes from setting x equal to y in the propagator given by Eq. (14.34). It is a short-distance divergence coming from the singularity of the propagator $\langle \xi^I(x) \xi^J(y) \rangle_{(0)}$ at $x = y$.

For now, we avoid the divergences at short-distance as well as long-distance by simply cutting off the high and low momenta. In other words,

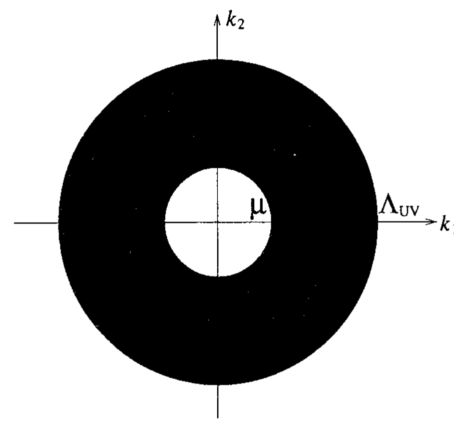


FIGURE 3. Cut-off

we perform the momentum integral in the region

$$\mu^2 \le k^2 \le \Lambda_{\mathrm{UV}}^2 \tag{14.37}$$

[1] There are actually quadratically divergent terms as well. In the present discussion we simply omit them, in order to avoid too many complications in our presentation.

where μ and Λ_{UV} are the lower and higher momentum cut-off (see Fig. 3). This manipulation is called *regularization*, and we will later interpret what it means. The momentum integral restricted to this region is given by

$$\int_{\mu\le|k|\le\Lambda_{\mathrm{UV}}} \frac{d^2k}{(2\pi)^2}\frac{1}{k^2} = \frac{1}{2\pi}\log\left(\frac{\Lambda_{\mathrm{UV}}}{\mu}\right). \tag{14.38}$$

We find similar divergences in the four-point functions of the ξ^I's as well. The four-point function at the one-loop level is obtained by summing the first three diagrams in Fig. 2 (B). It is given by

$$\begin{aligned}(14.39)\quad &\langle\xi^{I_1}(x_1)\xi^{I_2}(x_2)\xi^{I_3}(x_3)\xi^{I_4}(x_4)\rangle_{(1)}\\ &= -\frac{1}{3}\int\prod_{i=1}^{4}\left(\frac{d^2p_i}{(2\pi)^2}\frac{e^{ip_ix_i}}{p_i^2}\right)(2\pi)^2\delta(p_1+p_2+p_3+p_4)\\ &\times\left[(p_3\cdot p_4)\left\{\left(R_{(4)}+\frac{1}{6\pi}\log\left(\frac{\Lambda_{\mathrm{UV}}}{\mu}\right)R_{(4)}\cdot R_{(2)}\right)_{I_1I_2I_3I_4}\right.\right.\\ &\left.\left.+(I_1\leftrightarrow I_2)\right\}+\cdots\right]\\ &+\text{third diagram},\end{aligned}$$

where $R_{(4)}$ is the Riemannian curvature and $R_{(4)}\cdot R_{(2)}$ is defined by

$$\begin{aligned}(14.40)\quad &(R_{(4)}\cdot R_{(2)})_{IJKL}\\ &:= R_{NJKL}R^N{}_I + R_{INKL}R^N{}_J + R_{IJNL}R^N{}_K + R_{IJKN}R^N{}_L,\end{aligned}$$

and $+\cdots$ are permutations in (1234). Here again, there is a logarithmic divergence that is regularized by restricting the momentum integral to the region shown in Eq. (14.37). The last line of Eq. (14.39) is the term coming from the third diagram of Fig. 2 (B) and also has a divergence of the same order; it is simply obtained by replacing one of the four propagators by $\langle\xi^I(x)\xi^J(y)\rangle_{(1)}$ in Eq. (14.36) and summing over permutations.

These regularized correlation functions are divergent if we remove the cut-off as $\Lambda_{\mathrm{UV}}/\mu\to\infty$. These divergences can actually be tamed by a manipulation called *renormalization*. Let us modify the fields ξ^I and the target space metric g_{IJ} at ϕ_0 as

$$\begin{aligned} g_{IJ}=\delta_{IJ} &\;\to\; g_{0IJ}=\delta_{IJ}+a_{IJ},\\ \xi^I &\;\to\; \xi_0^I=\xi^I+b^I{}_J\xi^J.\end{aligned}$$

Namely, we replace ξ^I and g_{IJ} (and all quantities that depend on g_{IJ}, e.g., R_{IJKL}) in the classical action by ξ_0^I and g_{0IJ} and consider it as the action for ξ^I. The action is then expressed as

$$\begin{aligned} S_0=\frac{1}{2}\int\Big(&(1+a+2b)_{IJ}\partial_\mu\xi^I\partial^\mu\xi^J\\ &-\frac{1}{3}(R_{(4)}+R_{(4)}\cdot b)_{IKJL}\xi^K\xi^L\partial_\mu\xi^I\partial^\mu\xi^J+\cdots\Big)d^2x,\end{aligned}$$

where we set $b_{IJ}:=b^I{}_J$. Let us consider the two- and four-point functions at the one-loop level, $\langle\xi^I(x)\xi^J(y)\rangle_{(1)}$ and $\langle\xi^{I_1}(x_1)\xi^{I_2}(x_2)\xi^{I_3}(x_3)\xi^{I_4}(x_4)\rangle_{(1)}$. We regard the a_{IJ} and $b^I{}_J$ to be already of one-loop order in the loop expansion. Now, let us choose a_{IJ} and $b^I{}_J$ to be proportional to $\log(\Lambda_{\mathrm{UV}}/\mu)$ and try to find the coefficients so that the divergences they produce cancel the divergences in Eqs. (14.36) and (14.39) which are regularized by the cut-off in Eq. (14.37). We can actually find such a and b. The solutions are

$$g_{0IJ}=g_{IJ}+\frac{1}{2\pi}\log\left(\frac{\Lambda_{\mathrm{UV}}}{\mu}\right)R_{IJ}, \tag{14.41}$$

$$\xi_0^I=\left(\delta^I{}_J-\frac{1}{6\pi}\log\left(\frac{\Lambda_{\mathrm{UV}}}{\mu}\right)R^I{}_J\right)\xi^J. \tag{14.42}$$

Then the two- and four-point functions are finite at the one-loop level even as we remove the cut-off $\Lambda_{\mathrm{UV}}/\mu\to\infty$. Namely, if we change the target space metric and coordinate variables in a way depending on the cut-off, the correlation functions become finite when the cut-off is removed. This change of variables and the metric is what we call *renormalization/*.

14.2.3. Renormalization Group. What we have done above – regularization of the divergences and renormalization – has important physical significance beyond being a technical manipulation to make the correlation functions finite. It makes manifest an important aspect of quantum field theory, i.e., how its description changes as we change the energy scale. We give a short account of this important idea, called the *renormalization group*, which was introduced by Ken Wilson. We consider a theory of scalar fields with several coupling constants. The collection of fields and the coupling constants are denoted by $\phi(x)$ and g respectively. We denote the action by $S(\phi,g)$. In the non-linear sigma model under consideration, $\phi(x)$ corresponds to the fields $\xi^I(x)$ and the metric g_{IJ} is considered as a collection of infinitely many coupling constants.

Let us consider the Fourier mode expansion of $\phi(x)$;

$$\phi(x) = \int \frac{d^2k}{(2\pi)^2} \, \mathrm{e}^{ikx} \widehat{\phi}(k). \tag{14.43}$$

Usually the integral is over all frequencies $0 \leq |k| < \infty$. Setting an ultraviolet cut-off Λ_{UV} means that we restrict the integral to the disc

$$|k| \leq \Lambda_{\mathrm{UV}}, \tag{14.44}$$

and remove the higher frequency modes from $\phi(x)$. We denote such a field by $\phi_0(x)$ and call it a field at the cut-off scale Λ_{UV}. Thus,

$$\phi_0(x) = \int_{0 \leq |k| \leq \Lambda_{\mathrm{UV}}} \frac{d^2k}{(2\pi)^2} \, \mathrm{e}^{ikx} \widehat{\phi}(k). \tag{14.45}$$

We also denote the coupling constant by g_0. The path-integral is over this field ϕ_0;

$$Z = \int \mathcal{D}\phi_0 \, \mathrm{e}^{-S(\phi_0, g_0)}. \tag{14.46}$$

Then the momentum integral is cut off at Λ_{UV} and the ultraviolet divergences as in Eq. (14.36) are avoided. Since some of the Fourier modes are missing, the field $\phi_0(x)$ is not the most general one. In particular it is almost a constant within a distance $\Delta x \sim 1/\Lambda_{\mathrm{UV}}$. Thus, setting a UV cut-off is essentially the same as setting a short-distance cut-off. Introduction of a cut-off breaks the Poincaré invariance of the theory. Eventually, we would like to take the continuum limit, $1/\Lambda_{\mathrm{UV}} \to 0$, where Poincaré invariance is recovered. The question is whether one can achieve this by making the physics at a finite energy M regular.

Let us decompose the integration region of Eq. (14.45) into two parts:

$$0 \leq |k| \leq \mu \text{ and } \mu \leq |k| \leq \Lambda_{\mathrm{UV}}. \tag{14.47}$$

We denote the corresponding mode expansions as

$$\phi_L(x) = \int_{0 \leq |k| \leq \mu} \frac{d^2k}{(2\pi)^2} \, \mathrm{e}^{ikx} \widehat{\phi}(k), \tag{14.48}$$

$$\phi_H(x) = \int_{\mu \leq |k| \leq \Lambda_{\mathrm{UV}}} \frac{d^2k}{(2\pi)^2} \, \mathrm{e}^{ikx} \widehat{\phi}(k), \tag{14.49}$$

where "L" and "H" stand for Low and High energies. We would like to study the behavior of the system at energies of order μ or less, e.g., scattering

amplitudes of particles of momentum $\lesssim \mu$. Then it is convenient if there is an action in terms of $\phi_L(x)$ only that reproduces the low energy behavior. This can be obtained by integrating over ϕ_H in the path-integral but keeping $\phi_L(x)$ as a variable:

$$\mathrm{e}^{-S_{\mathrm{eff}}(\phi_L, g_0)} = \int \mathcal{D}\phi_H \, \mathrm{e}^{-S(\phi_L + \phi_H, g_0)}. \tag{14.50}$$

This is called the *effective action* at energy μ. The regularization we have done in the non-linear sigma model — keeping only momenta in the range $\mu \leq |k| \leq \Lambda_{\mathrm{UV}}$ — is precisely this integration over the "high energy field" $\phi_H(x)$. In that example, we have also observed that some correlation functions diverge when we take the limit $\Lambda_{\mathrm{UV}}/\mu \to \infty$. Such a divergence means that the resulting effective action $S_{\mathrm{eff}}(\phi_L, g)$ is ill defined or irregular as we take the limit $\Lambda_{\mathrm{UV}}/\mu \to \infty$. Such an irregularity can be regarded as a mandate to change the description of the theory at the low energy scale $\mu \ll \Lambda_{\mathrm{UV}}$. If one can find another set of variables and parameters such that the effective action is regular, that is a good description of the theory at the scale μ. In many cases, the change of variables and parameters takes the form

$$g_0 = g_0(g, \tfrac{\Lambda_{\mathrm{UV}}}{\mu}), \tag{14.51}$$

$$\phi_0(x) = Z(g, \tfrac{\Lambda_{\mathrm{UV}}}{\mu})\phi(x) + \phi_H(x). \tag{14.52}$$

Here $\phi(x)$ and g are new fields and the coupling constants in terms of which the effective action

$$\mathrm{e}^{-S_{\mathrm{eff}}(\phi, g; \mu)} = \int \mathcal{D}\phi_H \, \mathrm{e}^{-S(\phi_0, g_0)} \tag{14.53}$$

is regular in the continuum limit $\Lambda_{\mathrm{UV}}/\mu \to \infty$. The fields $\phi_0(x)$ and the couplings g_0 at the cut-off scale Λ_{UV} are called the *bare fields* and the *bare couplings*.

One can look at this change of fields and couplings in two ways. One viewpoint is to fix μ and move Λ_{UV}. We fix the fields $\phi(x)$ and couplings g at the scale μ but change the bare fields $\phi_0(x)$ and the bare couplings g_0 according to Eqs. (14.52)–(14.51). If we can move Λ_{UV} to infinity without changing the behavior of the system at a finite energy μ (described in terms of $\phi(x)$ and g), the continuum limit is well defined and we obtain a continuous field theory with Poincaré invariance. Another viewpoint is to move

μ, fixing the cut-off scale Λ_{UV} along with the bare fields and couplings.[2] Then the renormalized fields $\phi(x)$ and couplings g change according to Eqs. (14.52)–(14.51). In particular, if we change the scale from μ_1 to μ_2, the couplings change from $g_1 = g(g_0, \frac{\mu_1}{\Lambda_{\text{UV}}})$ to $g_2 = g(g_0, \frac{\mu_2}{\Lambda_{\text{UV}}})$, where $g(g_0, \mu/\Lambda_{\text{UV}})$ is the inverse function of Eq. (14.51). Alternatively, one can also obtain the effective action at the scale μ_2 from the one at a higher energy scale μ_1 by performing the integration over the modes of $\phi(x)$ with frequencies in the range $\mu_2 \leq |k| \leq \mu_1$. Then a similar action that occured when integrating over modes in $\mu_1 \leq |k| \leq \Lambda_{\text{UV}}$ will occur again. In particular, the couplings g_2 can also be written in terms of the coupling g_1 as $g_2 = g(g_1, \frac{\mu_2}{\mu_1})$. Thus, as we change the energy scale, the coupling flows along the vector field

$$\beta(g) = \mu \frac{d}{d\mu} g(g_1, \tfrac{\mu}{\mu_1})\Big|_{g_1 = g, \mu_1 = \mu} \tag{14.54}$$

in the space of coupling constants. The vector field $\beta(g)$ is called the *beta function* for the coupling constants g.

The Massive Fields. In the above discussion, we kept all the fields $\phi(x)$ when we described the low energy effective action. However, there are instances where it is more appropriate to integrate out all the modes of some field so that it does not appear in the effective theory. This is the case where there are *massive fields* with mass larger than the scale we are interested in. Such massive fields do not appear in non-linear sigma models but can appear in Landau–Ginzburg models.

The simplest example of massive fields is the free scalar field $\varphi(x)$, which has the action

$$S = \int_\Sigma \left(\partial^\mu \varphi \partial_\mu \varphi + m^2 \varphi^2 \right) d^2x. \tag{14.55}$$

The parameter m is the mass of the field φ. The second term, the mass term, explicitly breaks the classical scale invariance. In Minkowski space, the equation of motion is given by $(\partial_0^2 - \partial_1^2 + m^2)\varphi = 0$. The solution has a Fourier expansion where the frequencies are restricted to those which satisfy

$$(k_0)^2 = (k_1)^2 + m^2. \tag{14.56}$$

This is indeed the relation of the energy and momentum of a particle of mass m. We note from this that the energy is bounded from below by m. Thus,

[2] Or more precisely we fix the family of $\phi_0(x)$ and g_0 parametrized by Λ_{UV} that defines a single continuum theory.

any mode is highly fluctuating or rapidly varying in a distance much larger than the length $1/m$. (This length is called the *Compton wavelength.*) In Euclidean space, its two-point function behaves for $|x| \gg 1/m$ as

$$\langle \varphi(x)\varphi(0)\rangle = \int \frac{d^2k}{(2\pi)^2} \frac{e^{ikx}}{k^2 + m^2} \simeq \frac{e^{-m|x|}}{\sqrt{8\pi m|x|}}. \tag{14.57}$$

Thus it rapidly decays at distances larger than the Compton wavelength $1/m$.

Suppose a theory contains a field φ with a mass term of mass m. At energy $\mu \gg m$, the Compton wavelength is much larger than the scale $1/\mu$. Therefore we should keep this field in the effective theory at that energy. The mass parameter m can be renormalized as any other parameter of the theory. Suppose we take μ very small, much smaller than the renormalized mass. In the effective theory at energy μ we should not see fields fluctuating rapidly compared to the distance $1/\mu$. In particular, the massive fields are rapidly fluctuating within the distance $1/\mu$, which is much larger than the Compton wavelength. Thus, it is appropriate to integrate out all modes of φ, the modes of the frequencies in the whole range $0 \leq |k| \leq \Lambda_{\text{UV}}$.

14.2.4. Back to the Sigma Model.

Now let us come back to the bosonic non-linear sigma model. The relations Eqs. (14.52)–(14.51) between high- and low-energy couplings/fields are given in this case by Eqs. (14.42)–(14.41). Thus, the beta function (or beta functional as there are infinitely many couplings) can be found by $0 = \mu \frac{d}{d\mu} g_{0IJ} = \beta_{IJ} - \frac{1}{2\pi} R_{IJ}$. Namely,

$$\beta_{IJ} = \frac{1}{2\pi} R_{IJ}. \tag{14.58}$$

This is the beta function determined at the one-loop level.

The behavior of the theory thus depends crucially on the Ricci tensor R_{IJ}. We separate the discussion into three different cases; the cases where R_{IJ} is positive definite, $R_{IJ} = 0$, and R_{IJ} is negative definite.

- $R_{IJ} > 0$ — Asymptotic Freedom.

When the Ricci tensor is positive definite. $R_{IJ} > 0$, the bare metric

$$g_{0IJ} = g_{IJ} + \frac{1}{2\pi} \log\left(\frac{\Lambda_{\text{UV}}}{\mu}\right) R_{IJ} \tag{14.59}$$

grows large as we increase the UV cut-off Λ_{UV} (fixing the scale μ and the renormalized metric g_{IJ}). This means that the sigma model is weakly coupled at higher and higher energies, since the sigma model coupling is inversely proportional to the size of the target space. Thus, the perturbation theory becomes better and better as we take $\Lambda_{\text{UV}} \to \infty$. This is a good sign for the existence of the continuum limit. This property is called *asymptotic freedom*, and the sigma model on a Ricci positive Riemannian manifold is said to be *asymptotically free*. This is a property shared with four-dimensional Yang-Mills theory or quantum chromodynamics with a small number of flavors. On the other hand, the above equation also shows that g_{IJ} becomes smaller as we lower the energy M (fixing the UV cut-off Λ_{UV} and the bare metric). This means that the sigma model is strongly coupled at lower energies or at longer distances. Thus the sigma model perturbation theory becomes worse at lower energies and will break down at some point. The description in terms of the coordinate variables $\xi^I(x)$ will no longer be valid at low enough energies. Finding the low-energy description and behavior of the sigma model on a manifold with $R_{IJ} > 0$ is thus a difficult problem. This is one point where an analogue of $R \to 1/R$ duality (if it exists) is possibly useful; that may make it easier to study long distance behavior.

- $R_{IJ} = 0$ — Scale Invariance.

When the Ricci tensor is vanishing, $R_{IJ} = 0$, the one-loop beta function vanishes. Thus the theory is scale invariant at the one-loop level. Of course the beta function may receive nonzero contributions from higher loops, and the behavior of the theory depends on them. The sigma model on the torus we considered earlier is an example where the scale invariance holds exactly.

- $R_{IJ} < 0$ — Ultraviolet Singularity.

When the Ricci tensor is negative definite, $R_{IJ} < 0$, the bare metric decreases as we increase the cut-off Λ_{UV}. Thus the sigma model perturbation theory becomes worse at higher energies. In particular, there is a problem in taking the continuum limit $\Lambda_{\text{UV}} \to \infty$. Thus the sigma model on a Riemannian manifold with negative Ricci tensor is not a well-defined theory by itself. However, it may happen that such a theory appears as a low energy effective theory of some other (possibly well-defined) theory. In such a case, the low energy behavior is easy to study; the metric increases as we lower the

scale M and thus the sigma model is weakly coupled at lower energies. We can use sigma model perturbation theory to study the low energy physics.

RG Flow for Supersymmetric Sigma Models. So far, we have been considering the bosonic non-linear sigma models. What about supersymmetric sigma models? One can carry out a similar computation of the two- and four-point functions taking into account the fermion loops. It turns out that having fermions does not modify the one-loop beta function of the metric as shown in Eqs. (14.58) or (14.59). Thus, what we have said above for the three cases applies equally well to the supersymmetric sigma models as well. The sigma model perturbation theory is well defined only for $R_{IJ} \geq 0$. The sigma model on a Ricci-flat Kähler manifold is scale invariant at the one-loop level. The sigma model on a Ricci-positive Kähler manifold is asymptotically free. It is known, however, that the beta functions at higher loops are modified. For example, the two-loop beta function can be written in terms of the covariant derivatives of Ricci tensor. Thus, the two-loop beta function vanishes again for Ricci-flat manifolds. It also vanishes for symmetric spaces. Thus, it had originally been expected that the beta function vanishes to all orders in perturbation theory for Calabi–Yau manifolds – for which $R_{IJ} = 0$ – and receives contribution only at one loop for symmetric spaces. Further study showed, however, that the beta function is actually non-vanishing at the four-loop level for the Calabi–Yau sigma model. For Hermitian symmetric spaces such as $\mathbb{CP}^{N-1}$ and Grassmannians, there is an argument that the beta function receives contributions only at one loop.

One important remark is now in order. We have actually seen, in the supersymmetric sigma model, how the Kähler class changes according to the change of the scale; see Eq. (14.20). This suggested a change of the metric Eq. (14.22) under the scale transformation. This is actually nothing but what we have found in Eq. (14.58) at the one-loop level in the sigma model perturbation theory. However, as we have stated, this one-loop answer is not always the exact result for the renormalization of the metric; there can be higher-loop corrections. Is it consistent with the result from Eq. (14.20)? In the argument to derive Eq. (14.20) we made no approximation, and Eq. (14.20) is indeed an exact result. The solution of this apparent puzzle is that the possible higher-loop corrections to the Kähler metric or the Kähler form ω are of a form such that $\Delta\omega = d\alpha$ for some one-form α. Then the *Kähler* *class* $[\omega]$ receives no higher-loop correction. Without computing the higher-loop amplitudes to determine the exact renormalization group flow, we know at least some information exactly by a very elementary consideration. This is the power of supersymmetry. The essential point in the argument for Eq. (14.20) was that the correlation function scales simply as Eq. (14.13), as long as the inserted operator is invariant under some supercharge. We also notice the similarity of the argument to the one for the axial anomaly: both reduce to counting the index of the fermion Dirac operator. This is actually not a coincidence. The axial R-rotation and the scale transformation are in fact related by supersymmetry. Likewise, the Kähler class and the class of the B-field are superpartners of each other. (This last point will be made more explicit and precise in the next chapter where we provide a global definition of the supersymmetric non-linear sigma models for a certain class of target spaces.)

Another lesson we learn from these considerations is that the Kähler metric itself is not necessarily a good quantity to parametrize the theory; it can be corrected by infinitely many loops, which are practically impossible to compute (usually). Rather, the Kähler class is the one whose renormalization property is controlled as in Eq. (14.20), and can be a good parameter of the theory. The coordinates of $H^2(M,\mathbb{R})$ are the natural parameters for the Kähler class and are called the Kähler parameters. Thus, if $\dim H^2(M,\mathbb{R}) = k$ there are k Kähler parameters. There is actually one other real (periodic) parameter corresponding to each Kähler parameter. This is the parameter for the class $[B]$ of the B-field and takes values in the torus $H^2(M,\mathbb{R})/H^2(M,\mathbb{Z})$.[3] As we will see in the next section, the Kähler parameter and the corresponding parameter for the class $[B]$ naturally combine into one complex parameter. In total there are k complex parameters. To be more precise, for the case $c_1(M) \neq 0$, where the Kähler class is indeed renormalized, it is more appropriate to introduce a scale parameter Λ so that the Kähler class at the energy μ is given by

$$[\omega](\mu) = [\widetilde{\omega}] + \log(\mu/\Lambda)c_1(M). \tag{14.60}$$

Here $[\widetilde{\omega}]$ is a class in $H^2(M,\mathbb{R})$ transverse to the line spanned by $c_1(M)$. The scale parameter Λ replaces one of the Kähler parameters. This is a phenomenon called *dimensional transmutation*. In such a case, there is also an axial anomaly. This means that the shift of the class of the B field in the direction of $c_1(M)$ can be undone by a field redefinition (axial rotation). See Eq. (13.48). Thus one of the B-class parameters is unphysical and can be removed. Then if $c_1(M) \neq 0$, there are $k-1$ complex parameters and one scale parameter Λ.

[3]There are more sophisticated proposals for where the cohomology class of the B-field lies; we content ourselves here with the simplest interpretation.

14.3. Superspace Decouplings and Non-Renormalization of Superpotential

In the context of $(2,2)$ supersymmetric quantum field theories in two dimensions, we have seen that we can vary the action in five different ways: by deforming the chiral or twisted chiral superpotential and their conjugates, and also by deforming the D-terms. Here we wish to prove certain decoupling and non-renormalization theorems involving these terms. In particular we will show that varying the D-terms does not induce any corrections to the superpotential terms. Secondly we will show that the superpotential terms (chiral and twisted anti-chiral) are decoupled from each other, and neither gets renormalized. However, the D-terms do get renormalized.

14.3.1. Decoupling of D-term, F-term and Twisted F-term. The basic idea to prove decoupling is to consider an enlarged QFT where certain parameters in the action are promoted to fields. Morever, one considers a one-parameter family of such theories given by an action S_ϵ, where in the limit as $\epsilon \to 0$ one recovers the original theory. For the theory with action S_ϵ one proves a certain decoupling theorem which therefore leads to the decoupling result also in the limit $\epsilon \to 0$. In particular we will see that in the effective action, F-terms and twisted F-terms cannot mix. Moreover the D-terms cannot enter into the effective action for the F-terms or twisted F-terms. But the reverse can happen: the effective theory of D-terms does in general include F-term and twisted F-term couplings.

Let us consider a theory of chiral superfields Φ_i and twisted chiral superfields $\widetilde{\Phi}_{\tilde{i}}$ with the Lagrangian,

$$\int d^4\theta\, K(\Phi_i, \overline{\Phi_i}, \widetilde{\Phi}_{\tilde{i}}, \overline{\widetilde{\Phi}}_{\tilde{i}}, \gamma_b) \tag{14.61}$$

$$+\left(\int d^2\theta\, W(\Phi_i, \lambda_a) + c.c.\right) \tag{14.62}$$

$$+\left(\int d^2\widetilde{\theta}\, \widetilde{W}(\widetilde{\Phi}_{\tilde{i}}, \widetilde{\lambda}_{\tilde{a}}) + c.c.\right). \tag{14.63}$$

Here λ_a and $\widetilde{\lambda}_{\tilde{a}}$ are parameters in the superpotential W and the twisted superpotential $\widetilde{W}$ respectively and γ_b are parameters in the D-term. We want to see whether the parameters γ_b, λ_a can enter into the effective twisted superpotential $\widetilde{W}_{\text{eff}}$ at a lower energy and whether $\gamma_b, \widetilde{\lambda}_{\tilde{a}}$ can enter into W_{eff}. Let us now promote the parameters λ_a and $\widetilde{\lambda}_{\tilde{a}}$ to chiral superfields Λ_a and twisted chiral superfields $\widetilde{\Lambda}_{\tilde{a}}$. For the γ_b we consider two cases. We promote γ_b to a field Γ_b which is chiral for the proof of the first decoupling and twisted chiral in the second case.

We introduce the kinetic terms

$$\frac{1}{\epsilon}\int d^4\theta\left(\sum_b \pm|\Gamma_b|^2 + \sum_a |\Lambda_a|^2 - \sum_{\tilde{a}} |\widetilde{\Lambda}_{\tilde{a}}|^2\right) \tag{14.64}$$

where the $\pm$ sign in front of the Γ_b term depends on whether we are considering it to be a chiral or a twisted chiral field. We thus have an enlarged theory with an action we denote as S_ϵ. Since Λ_a is a chiral superfield it cannot enter into $\widetilde{W}_{\text{eff}}$. Also, $\widetilde{\Lambda}_{\tilde{a}}$ cannot enter into W_{eff}. Similarly, if we choose Γ_b to be a chiral superfield it cannot enter into $\widetilde{W}_{\text{eff}}$, and if we choose it to be a twisted chiral superfield it cannot enter into W_{eff}. Otherwise, supersymmetry would be violated. This statement is valid for any ϵ. Now let us consider the limit $\epsilon \to 0$. In this limit the kinetic term of the fields $\Lambda_a, \widetilde{\Lambda}_{\tilde{a}}, \Gamma_b$ becomes very large. Thus any variation of the corresponding fields over the two-dimensional space-time manifold gives a very large action. Thus in this limit the fields are frozen to constant values. In other words the scalar components of these new superfields become constants, and all other components "vanish." We have thus recovered the effective action for the original system in this limit. We thus see that there is no mixing of the parameters between the superpotential and twisted superpotential. Nor do parameters in the D-term enter the superpotential terms. However, this argument does not preclude the possibility that in the effective D-term

the couplings in the superpotential terms appear. And in fact the effective D-terms do receive corrections involving the superpotential couplings.

14.3.2. The Non-renormalization Theorem. Here we will argue that the terms in the superpotential do not change in the effective theory. This was not precluded by the decoupling argument above, as in principle the superpotential terms may have changed depending only on the superpotential coupling constants.

The argument is rather simple: We can *demote* fields to parameters by changing D-terms. In other words, if we change the D-terms for the chiral and twisted chiral fields by the ϵ deformation,

$$\Delta_\epsilon S = \frac{1}{\epsilon}\int d^4\theta\left(\sum_i |\Phi_i|^2 - \sum_{\tilde{i}} |\widetilde{\Phi}_{\tilde{i}}|^2\right), \tag{14.65}$$

so that the D-terms will give rise to large kinetic terms, we see that in this limit Φ_i and $\widetilde{\Phi}_{\tilde{i}}$ become parameters and all the quantum fluctuations are suppressed by the action. Thus in the limit as $\epsilon \to 0$ there cannot be any renormalization of the superpotential. However in the previous section we had shown that the D-term parameters do not affect the chiral and twisted chiral superpotentials. Thus the statement is that for *any* ϵ the superpotential does not get renormalized, including the $\epsilon \to 0$ limit. This proves the important result that all the chiral and twisted-chiral superpotential terms are not renormalized.

14.3.3. Another Derivation of the F-term Non-renormalization Theorem. The non-renormalization theorem for chiral and twisted chiral superpotential terms is so important that we will present another proof for it here, based on symmetry arguments. As a simplest example, let us consider a single-variable Landau–Ginzburg theory with the superpotential

$$W(\Phi, m, \lambda) = m\Phi^2 + \lambda\Phi^3. \tag{14.66}$$

We would like to study the low energy effective action of this system at some scale μ, integrating out modes with frequencies in the range $\mu \le |k| \le \Lambda_{\text{UV}}$. This leads to an effective superpotential $W_{\text{eff}}(\Phi)$. In this model, since the superpotential is not homogeneous, the vector R-symmetry is explicitly broken. Also, there is no other global symmetry except the axial R-symmetry that acts on the lowest scalar component ϕ of Φ trivially. Thus, it appears that we cannot constrain the form of $W_{\text{eff}}(\Phi)$ using the symmetry. Is it possible that all kinds of new terms are generated in $W_{\text{eff}}(\Phi)$?

The answer is no. The effective superpotential is exactly the same as the superpotential in Eq. (14.66) at the cut-off scale. One way to see this is to explicitly compute the effective action using Feynman diagrams. There is a supergraph formalism developed by Grisaru, Siegel and Roček which makes it easier to see. Another is the argument by Seiberg which makes use of holomorphy and other physical conditions as the basic constraints. Having in mind applications in other contexts, we describe the latter argument here.

The first step again is to promote the parameters that enter into the superpotential to chiral superfields. In the above example we promote the parameters m and λ to chiral superfields M and Λ. Take the Kähler potential for these new variables as

$$K_M + K_\Lambda = \frac{1}{\epsilon}\overline{M}M + \frac{1}{\epsilon}\overline{\Lambda}\Lambda, \tag{14.67}$$

and consider the limit $\epsilon \to 0$. This will freeze the fluctuations of M and Λ around some background value and give us a starting system where m and λ are simply parameters.

Before the limit $\epsilon \to 0$, the superpotential is

$$W(\Phi, M, \Lambda) = M\Phi^2 + \Lambda\Phi^3. \tag{14.68}$$

Now this system has a larger symmetry. The superpotential is quasi-homogeneous; it has vector R-charge 2 if we assign vector R-charge $(1,0,-1)$ for (Φ, M, Λ). Thus the vector R-rotation is a symmetry of the system. Also, there is another anomaly-free global $U(1)$ symmetry where the superfields (Φ, M, Λ) have charge $(1,-2,-3)$ so that the superpotential W is invariant.

There are three following basic constraints on the effective superpotential $W_{\text{eff}}(\Phi, M, \Lambda)$.

Symmetry: W_{eff} must have charge 2 under $U(1)_V$ and must be invariant under the global $U(1)$ symmetry.

Holomorphy: W_{eff} must be a holomorphic function of Φ, M, Λ.

Asymptotic Behavior: W_{eff} must approach the classical value $M\Phi^2 + \Lambda\Phi^3$ for an arbitrary limit in which $M, \Lambda \to 0$.

The first two conditions constrain the form of W_{eff} as

$$W_{\text{eff}}(\Phi, M, \Lambda) = M\Phi^2 f(t);\ t := \Lambda\Phi/M, \tag{14.69}$$

where $f(t)$ is a holomorphic function of t. In the limit where $M, \Lambda \to 0$ as $M = \alpha M_*$, $\Lambda = \alpha\Lambda_*$ with $\alpha \to 0$, the parameter t is $t_* := \Lambda_*\Phi/M_*$ and W_{eff} approaches $M\Phi^2 f(t_*)$. This is equal to the classical expression only if $f(t_*) = 1 + t_*$. Since t_* is arbitrary, we conclude $f(t) = 1 + t$. Thus we have shown

$$W_{\text{eff}} = M\Phi^2(1+t) = M\Phi^2 + \Lambda\Phi^3. \tag{14.70}$$

Now let us take the limit where $\epsilon \to 0$. This cannot change the superpotential and thus we have shown the non-renormalization of the superpotential.

EXERCISE 14.3.1. *Generalize the above argument to show the non-renormalization of F-terms for multi-variable LG models.*

14.3.4. Integrating Out Fields.

The non-renormalization theorem above applies to the case where we write an effective theory involving all the fields in the theory. However, when the masses (that appear in the superpotential) of some of the fields are larger than the scale we are interested in, it is appropriate to integrate out these heavy fields, and we obtain an effective action in terms of the light fields. The non-renormalization theorems above do not apply to the effective superpotential in terms of the fields we keep. In fact, as we will now see, the effective superpotential will look as though it had received "quantum corrections" in terms of the fields we retain.

For instance, let us consider a theory of two chiral superfields Φ and Φ_1 with the superpotential

$$W = \lambda\Phi^3 + \kappa\Phi_1\Phi^2 + m\Phi_1^2. \tag{14.71}$$

Suppose we are interested in the effective action at scale μ which is much smaller than m. Then it is appropriate to integrate out the field Φ_1. This is carried out by eliminating Φ_1 by using the equation of motion $\Phi_1 = -\frac{\kappa}{2m}\Phi^2$, which comes from setting $\partial_{\Phi_1} W = 0$:

$$W_{\text{eff}} = W\Big|_{\Phi_1 = -\frac{\kappa}{2m}\Phi^2} = \lambda\Phi^3 - \frac{\kappa^2}{4m}\Phi^4. \tag{14.72}$$

Thus, not only is the field Φ_1 gone, but a new term $-\frac{\kappa}{4m}\Phi^4$ is generated.

14.4. Infrared Fixed Points and Conformal Field Theories

It is natural to look for fixed points of RG flows. QFTs corresponding to such fixed points are called conformal field theories. This implies that under the rescaling of the metric the theory preserves its form. In particular all the correlations have simple scaling properties based on their "scaling dimensions". For (1+1)-dimensional QFTs the existence of scale invariance gives rise to an infinite-dimensional group of symmetries whose generators satisfy the Virasoro algebra:

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{c}{12}(n^3 - n)\delta_{n+m,0}. \tag{14.73}$$

Here c is a central element and is realized as a c-number. The generator L_n acts on the coordinates of the Euclidean worldsheet by

$$z \to z + \epsilon z^{n+1}. \tag{14.74}$$

In other words the action of the generators on the coordinates is given by

$$L_n \leftrightarrow z^{n+1} d/dz. \tag{14.75}$$

One also has the anti-holomorphic version of these generators acting on the anti-holomorphic coordinates. For example, for the free field theory, which is conformal, the Virasoro algebra is realized as

$$L_n = \frac{1}{2} : \sum_m \alpha_m \alpha_{n-m} :, \tag{14.76}$$

where : : is the normal ordering defined in Sec. 11.1.

In the context of Landau–Ginzburg theories we have a non-renormalization theorem for the superpotential which means that W does not renormalize. However the superspace integral measure $d^2z d^2\theta$ rescales by a factor of λ as we rescale $z \to \lambda z$ and $d\theta \to \lambda^{-1/2} d\theta$. Thus for an LG theory to correspond to a conformal theory we must be able to redefine fields by some scaling factor such that

$$\lambda W(\Phi_i) = W(\lambda_i \Phi_i) \tag{14.77}$$

where $\lambda_i = \lambda^{q_i}$. In other words W is a quasi-homogeneous function. Moreover the field Φ_i has scaling dimension q_i which is also its axial $U(1)$ charge. Thus a necessary condition for an LG theory to correspond to a conformal theory is having a quasi-homogeneous superpotential. Note that if we do not have a quasi-homogeneous superpotential, W *effectively* flows. What we mean by this is that by a redefinition of the fields the form of W changes, maintaining the form of higher dimension operators. For example, consider

$$W(\Phi) = \Phi^n + \Phi^k \tag{14.78}$$

with $n > k$. Then under $W \to \lambda W$ we can maintain the higher dimension operator, namely Φ^n of the same form, which means that we rescale $\Phi \to \lambda^{-1/n}\Phi$. Then we see that

$$W(\Phi) \to \Phi^n + \lambda^{\frac{n-k}{n}} \Phi^k. \tag{14.79}$$

This implies that the theory in the UV, corresponding to $\lambda \to 0$, has a superpotential which is effectively Φ^n and in the IR, corresponding to $\lambda \to \infty$, has a superpotential which is effectively Φ^k.

It is believed that the LG theories with quasi-homogeneous superpotential flow, with suitable D-terms, to unique conformal field theories. In this way we are led to attribute to each quasi-homogeneous W a $(2,2)$ superconformal field theory (which in addition to the Virasoro symmetry given above has a supercurrent and a $U(1)$ current symmetry as well). Moreover it is not difficult to show, using unitarity constraints on representations of the corresponding algebra, that the central charge c of the Virasoro algebra one obtains is $3D$, where D is the maximal axial charge in the chiral ring. This statement is true whether or not the $(2,2)$ conformal theory arises from a Landau–Ginzburg theory.

For the case of supersymmetric sigma models on Calabi–Yau manifolds, as we have seen the Kähler class does not flow. It is believed that in these cases the actual metric of the Calabi–Yau manifold flows to a unique metric compatible with conformal invariance. In the large volume limit the one-loop analysis we performed shows that this is the Ricci-flat metric. In general, however, the metric corresponding to the conformal fixed point is not the Ricci-flat metric. Nevertheless it is believed that for a fixed complex structure and fixed (complexified) Kähler class, there is a unique metric on the Calabi–Yau manifold corresponding to a superconformal sigma model. The c for such conformal theories is given by $3D$ where D is the complex dimension of the Calabi–Yau, as the highest R-charge of the chiral ring for the Calabi–Yau is its complex dimension.

It turns out that $(2,2)$ superconformal theories with $c < 3$, or equivalently $D < 1$, can be classified and all correspond to Landau–Ginzburg theories with quasi-homogeneous superpotential. Moreover they are in 1-1 correspondence with ADE singularities of C^2/Γ where Γ is a discrete subgroup of $SU(2)$. The A-series corresponds to cyclic subgroups, the D-series corresponds to dihedral subgroups and the E-series correspond to the three exceptional subgroups of $SU(2)$. The corresponding W's are given by

$$W = x^n + y^2 + z^2, \qquad A_{n-1},\ n > 1, \tag{14.80}$$

$$W = x^n + xy^2 + z^2, \qquad D_{n+1},\ n > 2, \tag{14.81}$$

and three exceptional cases:

$$W = x^3 + y^4 + z^2, \qquad E_6, \tag{14.82}$$

$$W = x^3 + xy^3 + z^2, \qquad E_7, \tag{14.83}$$

$$W = x^3 + y^5 + z^2, \qquad E_8. \tag{14.84}$$

The relation to ADE singularities is that they correspond to the $W = 0$ hypersurface in $\mathbb{C}^3$.

CHAPTER 15

Linear Sigma Models

In this section we study a class of supersymmetric gauge theories in $1+1$ dimensions, called linear sigma models. They provide us with a global description of non-linear sigma models, which have been described patchwise up till now. This enables us to distinguish the parameters of the sigma model that enter into F-terms and twisted F-terms. Furthermore, we will learn that a linear sigma model has different "phases" with various kinds of low energy theories, not just the non-linear sigma model. This leads to an interesting relation between the non-linear sigma models and Landau–Ginzburg models. Most importantly, the linear sigma model is the essential tool for the proof of mirror symmetry, as will be elaborated in later sections.

15.1. The Basic Idea

Let us consider a field theory of a number of real scalar fields $\phi^1, \dots, \phi^n$ with the Lagrangian

$$L = -\frac{1}{2}\sum_{i=1}^{n}(\partial_\mu \phi^i)^2 - U(\phi). \tag{15.1}$$

$U(\phi)$ is a function of the ϕ^i's and serves as the potential of the theory, which we assume to be bounded from below. An example is

$$U(\phi) = \frac{e^2}{4}\left(\sum_{i=1}^{n}(\phi^i)^2 - r\right)^2. \tag{15.2}$$

A classical vacuum of this system is a constant (independent of the world-sheet spatial coordinate) value of ϕ at a minimum of the potential, $U(\phi)$. In general there can be many classical vacua. The set of classical vacua (considered as a subset of $\mathbb{R}^n$) is called the vacuum manifold and will be denoted M_{vac}. In the example given by Eq. (15.2) with $r \le 0$, there is a unique minimum at the origin: $\phi^i = 0$ for all i. Thus, M_{vac} is a point. For $r > 0$, however, the minimum is attained for any $\phi = (\phi^i)$ which obeys $\sum_{i=1}^{n}(\phi^i)^2 = r$. In this case M_{vac} is the sphere S^{n-1} of size $\sqrt{r}$. At each point of M_{vac} the first derivative of $U(\phi)$ vanishes: $\partial_i U(\phi) = 0$. Therefore, the second derivative matrix, or the Hessian

$$\partial_i \partial_j U(\phi), \tag{15.3}$$

is well defined as a second rank symmetric tensor of $\mathbb{R}^n$ at such a point. It can be diagonalized by an orthogonal transformation and the eigenvalues are non-negative since $U(\phi)$ attains its minimum at M_{vac}. In the perturbative treatment of the theory where one expands the fields at such a point, the eigenvalues determine the masses of the fields. The fields tangent to M_{vac} are of course zero mass fields. In some cases, these are the only massless fields and the fields transverse to M_{vac} have positive masses. In some other cases, however, some massless fields do not correspond to a tangent direction of M_{vac}). In the example Eq. (15.2) with $r < 0$, all ϕ^i have a positive mass, $e\sqrt{|r|}$. At $r = 0$ all of them become massless but do not correspond to flat directions. For $r > 0$, all modes tangent to S^{n-1} are massless but the radial mode has a positive mass $e\sqrt{2r}$.

Let us assume that all transverse modes to M_{vac} are massive. The theory of massless modes is the non-linear sigma model on the vacuum manifold M_{vac}, if the massive modes are neglected. The metric of M_{vac} is the one induced from the Euclidean metric of $\mathbb{R}^n$ which appears in the kinetic term of Eq. (15.1). For instance, in the example given by Eq. (15.2) with $r > 0$ we have a sigma model on S^{n-1} of size $\sqrt{r}$. Of course one cannot ignore the massive modes altogether. However, as we have seen in the previous section, if we are interested in the behavior of the system at an energy much smaller than the masses of all the transverse modes, it is appropriate to integrate them out from the path-integral. Alternatively, one can take a limit of the parameters in $U(\phi)$ (like e in the example Eq. (15.2)) where the masses of the transverse modes go to infinity compared to the scale we are interested in, in which case they are completely frozen. In either case, we will obtain an effective theory in terms of the massless modes only.

Integrating out the massive modes will affect the theory of massless modes. It may change the metric of M_{vac} or even the topology of M_{vac}. There also will appear terms with four or more derivatives in the effective action. Such terms are not in the non-linear sigma model Lagrangian, but are *irrelevant* in the sense that they are negligible when the masses of the transverse modes are taken to infinity. We should note that the non-linear

sigma model itself is scale-dependent, as we have seen in the previous section. In order to have a "standard" non-linear sigma model, the scale dependence of the metric of M_{vac} should also be matched.

Since the kinetic term for the scalar fields in this example is that of the Euclidean or the linear space $\mathbb{R}^n$, this model is called the *linear sigma model.*[1]

15.1.1. Gauge Symmetry. We have seen at least classically that the non-linear sigma models on submanifolds of $\mathbb{R}^n$ can be obtained from the standard scalar field theory with a potential. Can we obtain in a similar fashion the non-linear sigma models on their *quotient* by some group action? A standard example of a manifold realized as a quotient is $\mathbb{CP}^{N-1}$ which is the $U(1)$ quotient of the sphere S^{2N-1} in $\mathbb{C}^N \cong \mathbb{R}^{2N}$, where the action of $U(1)$ is the uniform phase rotation of the coordinates of $\mathbb{C}^N$. What is the linear sigma model for $\mathbb{CP}^{N-1}$? We start with the one for S^{2N-1} which is described in terms of N complex scalar fields $\phi_1, \ldots, \phi_N$ with the Lagrangian

$$L = -\sum_{i=1}^{N} |\partial_\mu \phi_i|^2 - U(\phi), \tag{15.4}$$

with

$$U(\phi) = \frac{e^2}{2}\left(\sum_{i=1}^{N} |\phi_i|^2 - r\right)^2. \tag{15.5}$$

The vacuum manifold for $r > 0$ is indeed S^{2N-1}. Note that this Lagrangian is invariant under the constant phase rotation

$$(\phi_1(x), \ldots, \phi_N(x)) \to (\,\mathrm{e}^{i\gamma}\phi_1(x), \ldots, \mathrm{e}^{i\gamma}\phi_N(x)). \tag{15.6}$$

Namely, this is a global symmetry of the system. (There is a larger symmetry which will be mentioned later.) We want to change the theory now so that the vacuum manifold is the quotient of S^{2N-1} by this $U(1)$ action. In other words, we want the map $\phi(x) = \{\phi_i(x)\}$ to be physically equivalent to $\phi'(x) = (\,\mathrm{e}^{i\gamma(x)}\phi_i(x))$ for an arbitrary $\mathrm{e}^{i\gamma(x)}$ which can depend on the space-time coordinates x^μ. It may appear that the only thing one has to do is to

[1] Historically, the linear sigma model was introduced first. The *non-linear sigma model* was later developed when studying questions involving the vacuum geometry of linear sigma models. Sometimes in the mathematics of mirror symmetry, the words "linear sigma model" have yet another meaning.

declare the configurations related by

$$(\phi_1(x), \ldots, \phi_N(x)) \to (e^{i\gamma(x)}\phi_1(x), \ldots, e^{i\gamma(x)}\phi_N(x)). \tag{15.7}$$

to be physically equivalent. However, we note that the Lagrangian shown in Eq. (15.4) is not invariant under Eq. (15.7) unless the phase $e^{i\gamma}$ is a constant; the derivative term is not invariant since $\partial_\mu\phi_i$ transforms inhomogeneously as $\partial_\mu\phi_i \to e^{i\gamma(x)}(\partial_\mu + i\partial_\mu\gamma)\phi_i$. The standard recipe to make it invariant is to introduce a vector field (or a one-form field) v_μ which transforms as

$$v_\mu(x) \to v_\mu(x) - \partial_\mu\gamma(x), \tag{15.8}$$

so that $D_\mu\phi_i := (\partial_\mu + iv_\mu)\phi_i$ transforms homogeneously: $D_\mu\phi_i \to e^{i\gamma}D_\mu\phi_i$. Then the modified Lagrangian

$$L = -\sum_{i=1}^{N} |D_\mu\phi_i|^2 - U(\phi) \tag{15.9}$$

is invariant under Eqs. (15.7)–(15.8) for an arbitrary phase-valued function $e^{i\gamma(x)}$. We now declare that the configurations of $(\phi_i(x), v_\mu(x))$ related by the transformation Eqs. (15.7)–(15.8) are physically equivalent. This is the proposal.

The transformation Eqs. (15.7)–(15.8) is a kind of symmetry of the new theory since it keeps the Lagrangian from Eq. (15.9) invariant. However, it is not an ordinary symmetry since the transformed configuration is regarded as physically equivalent to the original one. There is a redundancy in the description of the theory, and the transformation Eqs. (15.7)–(15.8) simply relates those redundant configurations. Such a symmetry is called a *gauge symmetry*. The procedure starting from the theory shown in Eq. (15.4) with a global $U(1)$ symmetry and obtaining a theory given by Eq. (15.9) with a $U(1)$ gauge symmetry is called *gauging*. It usually involves introducing another field v_μ (and this is the reason why the connection form is called the *gauge field*). A theory with a gauge symmetry is called a *gauge theory*. Mathematically what this means is that we have a $U(1)$ vector bundle where v_μ is the connection and ϕ_i are sections of the bundle, and $D_\mu\phi_i$ is the covariant derivative of the section.

Now let us see whether the theory of the massless modes is equivalent to the sigma model on the $U(1)$ quotient of S^{2N-1}. Let us first look at the gauge field v_μ. We did not introduce a kinetic term for this field (although

we could introduce one; we will indeed do so shortly). Thus, the field v_μ acts like an auxiliary field and can be eliminated by solving its equation of motion, which reads as

$$\sum_{i=1}^{N} (D_\mu\overline{\phi_i}\phi_i - \overline{\phi_i}D_\mu\phi_i) = 0. \tag{15.10}$$

This equation is solved by

$$v_\mu = \frac{i}{2}\frac{\sum_{i=1}^N (\overline{\phi_i}\partial_\mu\phi_i - \partial_\mu\overline{\phi_i}\phi_i)}{\sum_{i=1}^N |\phi_i|^2}. \tag{15.11}$$

Once this is plugged into the Lagrangian from Eq. (15.9), everything is written in terms of $\phi_i(x)$ and the gauge transformation is implemented simply by Eq. (15.7).

EXERCISE 15.1.1. *Show that Eq. (15.7) induces the transformation in Eq. (15.8) via Eq. (15.11).*

Now it is clear that the theory of the massless modes is the sigma model on the $U(1)$ quotient of S^{2N-1}, which is $\mathbb{CP}^{N-1}$. What is the metric of this target space? To see this we first fix a configuration $\phi(x) = \{\phi_i(x)\}$ that defines a map to S^{2N-1} and represents a map to $\mathbb{CP}^{N-1}$. Let us pick a tangent vactor ξ^μ on the worldsheet. This is mapped by ϕ to a tangent vector of $\mathbb{CP}^{N-1}$, and we want to measure its length. From the Lagrangian in Eq. (15.9), we see that its length squared is measured as $\sum_{i=1}^N |\xi^\mu D_\mu\phi_i|^2$, namely the length squared of the vector $\xi^\mu D_\mu\phi_i$ in $\mathbb{C}^N$ measured by the standard Euclidean metric of $\mathbb{C}^N$. Here v_μ in $D_\mu\phi_i = (\partial_\mu + iv_\mu)\phi_i$ is given by Eq. (15.11) so that $D_\mu\phi_i$ obeys (15.10). Eq. (15.10) says that the vector $\xi^\mu D_\mu\phi_i$ is orthogonal to the orbit of the $U(1)$ gauge group action. This means that the length of a vector in $\mathbb{CP}^{N-1}$ is measured by first lifting it to a tangent vector of S^{2N-1} orthogonal to the $U(1)$ gauge orbit, and then measuring its length using the metric of S^{2N-1} or of the Euclidean metric of $\mathbb{C}^N$. This is a standard way to construct a metric on the quotient manifold. It turns out that the metric we obtain in the present example is r times the Fubini-Study metric. The Fubini-Study metric is expressed in terms of the inhomogeneous coordinates $z_i = \phi_i/\phi_N$ $(i = 1, \ldots, N-1)$ as

$$g^{\mathrm{FS}} = \frac{\sum_{i=1}^{N-1}|dz_i|^2}{1+\sum_{i=1}^{N-1}|z_i|^2} - \frac{\sum_{i=1}^{N-1}|\overline{z}_i dz_i|^2}{(1+\sum_{i=1}^{N-1}|z_i|^2)^2}. \tag{15.12}$$

15.1.2. Symmetry Breaking. Goldstone Bosons. Suppose the Lagrangian in Eq. (15.1) is invariant under an action of a group G on the coordinates ϕ^i.[2] This in particular means that the potential $U(\phi)$ is G-invariant, $U(g\phi) = U(\phi)$. In particular, a point that minimizes $U(\phi)$ is sent by G to points minimizing $U(\phi)$, so the vacuum manifold M_{vac} is invariant under the G action: G acts on M_{vac}.

Let us pick a classical vacuum, a point ϕ_0 in M_{vac}. In general only a proper subgroup H_0 of G fixes ϕ_0. This situation is described by saying the symmetry group G is *spontaneously broken* to the subgroup H_0 by the choice of a vacuum ϕ_0. The subgroup H_0 is said to be the unbroken subgroup of G at the vacuum ϕ_0. The "broken directions," G/H_0, span an orbit of G through ϕ_0. Since G keeps the potential $U(\phi)$ invariant, this orbit lies in the vacuum manifold M_{vac}. In particular, the modes tangent to this orbit (naturally identified as the vectors in $\mathrm{Lie}(G)/\mathrm{Lie}(H_0)$) are massless (i.e., the Hessian of U is null along those directions). These massless modes are called *Goldstone modes* (sometimes also *Goldstone bosons*, as the fields ϕ are bosonic).

In the example of Eq. (15.2), the global symmetry group is $G = O(n)$. For $r > 0$, any choice of vacuum breaks $O(n)$ to a subgroup isomorphic to $O(n-1)$. The vacuum manifold S^{n-1} consists of a single orbit $O(n)/O(n-1)$ and all the massless modes are the Goldstone bosons. For $r \leq 0$, the whole symmetry $O(n)$ remains unbroken at the (unique) vacuum. Therefore there is no Goldstone boson. In the example from Eq. (15.9), the Lagrangian itself is invariant under $U(N)$ but its $U(1)$ subgroup is a gauge symmetry and should not be counted as a part of the global symmetry. The global symmetry group is thus the quotient group $U(N)/U(1) = SU(N)/\mathbb{Z}_N$. For $r > 0$, any choice of a vacuum breaks $SU(N)/\mathbb{Z}_N$ to $U(N-1)/\mathbb{Z}_N$ and the vacuum manifold $\mathbb{CP}^{N-1}$ consists of a single orbit. Thus, all the massless modes are the Goldstone bosons.

The above discussion was in the context of the classical theories, but Goldstone's theorem states that the story is similar even in quantum theories. Let us consider a quantum field theory with a global symmetry G. Suppose the ground state $|0\rangle$ spontaneously breaks the symmetry group G

[2] In the present discussion, the kinetic term of ϕ^i does not have to be the one corresponding to the Euclidean metric. The Euclidean space $\mathbb{R}^n$ can be replaced by an arbitrary Riemannian manifold.

to the subgroup H_0. This means that $\langle g\mathcal{O}\rangle = \langle \mathcal{O}\rangle$ for all $\mathcal{O}$'s only when g belongs to H_0. Then Goldstone's theorem says that there is a massless scalar field associated to each Lie algebra generator of G that does not belong to H_0.

Higgs Mechanism. What if a *gauge symmetry* is broken by a choice of classical vacuum? We consider this problem in a specific example; The system of a complex scalar field ϕ and a gauge field v_μ with the Lagrangian

$$L = -|D_\mu\phi|^2 - \frac{1}{2e^2}v_{\mu\nu}^2 - \frac{e^2}{2}\left(|\phi|^2 - r\right)^2, \tag{15.13}$$

where $v_{\mu\nu} = \partial_\mu v_\nu - \partial_\nu v_\mu$ is the curvature of the gauge field v_μ. We consider the symmetry

$$\phi(x) \to e^{i\gamma(x)}\phi(x),\ v_\mu(x) \to v_\mu(x) - \partial_\mu\gamma(x), \tag{15.14}$$

of the Lagrangian as the gauge symmetry. This is almost the same as the system shown in Eq. (15.9) with $N = 1$. The only difference is that now we have the kinetic term for the gauge field. (We also call the procedure *gauging* when the gauge kinetic term is added like this.)

The classical vacua are at $|\phi|^2 = r$ (we assume $r > 0$), and if we choose one, say $\phi = \sqrt{r}$, the gauge symmetry is completely broken. Let us look at the theory near this vacuum. We use the polar coordinates for the complex field $\phi = \rho e^{i\varphi}$ which is non-singular there. If the $U(1)$ symmetry were not gauged, the angular variable φ would have been the Goldstone mode. Now, the derivative $D_\mu\phi$ is written as $D_\mu\phi = e^{i\varphi}(\partial_\mu\rho + i\rho(v_\mu + \partial_\mu\varphi))$, and the Lagrangian is expressed as

$$L = -(\partial_\mu\rho)^2 - \frac{e^2}{2}(\rho^2 - r)^2 - \frac{1}{2e^2}v_{\mu\nu}^2 - \rho^2(v_\mu + \partial_\mu\varphi)^2. \tag{15.15}$$

The gauge transformation shifts φ so that the combination $v'_\mu = v_\mu + \partial_\mu\varphi$ is gauge invariant. ρ and v'_μ are the only fields that appear in the Lagrangian (note that $v'_{\mu\nu} = v_{\mu\nu}$). The field φ is absorbed by the gauge field v_μ, or more precisely, φ and v_μ combine to make one vector field v'_μ. In terms of the variables (ρ, v'_μ), the system has no gauge symmetry and there is no redundancy in the description. We expand the Lagrangian at the vacuum $\rho = \sqrt{r}$ in terms of the shifted variable $\rho = \sqrt{r} + \varepsilon$;

$$L = -(\partial_\mu\varepsilon)^2 - 2e^2 r\varepsilon^2 - \frac{1}{2e^2}(v'_{\mu\nu})^2 - r(v'_\mu)^2 + \cdots, \tag{15.16}$$

where $+\cdots$ are the terms at least cubic in the fields ε, v'_μ. We see from this that the field ε has a mass. This is what we have seen already; the modes transverse to the vacuum manifold is massive in this case. More surprisingly, because of the fourth term, the vector field v'_μ also has a mass. In a sense, the gauge field acquires a mass by "eating" one Goldstone mode. This is what happens when a gauge symmetry is spontaneously broken. This is called the *Higgs mechanism*.

15.1.3. Symmetry Restoration in $1+1$ Dimensions. Goldstone's theorem says that the breaking of a continuous global symmetry yields massless scalar fields. One can actually use this to exclude the possibility of global symmetry breaking in the quantum theories in $1+1$ dimensions. The basic physical idea can be illustrated in the context of the sigma model: classically the configurations with least energy correspond to constant maps from the worldsheet to a point on the target space. However, quantum mechanically, with little cost in action the image of the worldsheet can spread out and this means that we cannot "freeze" the vacuum of the quantum theory to correspond to a fixed point on the target space. In this sense the sigma models are interesting for (1+1)-dimensional QFT and in a sense probe the full geometry of the target space (this fails to be the case in higher dimensions where the vacuum corresponds to maps to the vicinity of a given point in the target space).

The basic idea behind this lack of freezing in 1+1 dimensions is that a massless scalar field is not allowed in quantum field theories in two dimensions. For suppose there is a massless scalar field, $\phi(x)$. The two-point function would be given by

$$\langle\phi(x)\phi(y)\rangle = \int \frac{d^2k}{(2\pi)^2}\frac{\mathrm{e}^{ik(x-y)}}{k^2}. \tag{15.17}$$

This integral does not make sense since it has a logarithmic divergence at $k=0$, the infrared divergence. (Even if we cut off the integral near $|k|=0$, we would obtain $\langle\phi(x)\phi(y)\rangle = -(1/4\pi)\log|x-y| + \mathrm{c}$ where c is a constant associated with the choice of the cut-off and is not positive definite.)[3] This shows that a massless scalar field cannot be a good operator in the quantum theory. On the other hand, Goldstone's theorem shows that once a continuous global symmetry is broken, a massless scalar field operator appears. This proves that a continuous global symmetry cannot be broken in 1+1 dimensions. However, this argument does not exclude the breaking of discrete symmetry, which would not lead to Goldstone modes. Indeed there are many examples where a discrete symmetry is spontaneously broken. We will see some of them in this chapter.

[3]We have used the propagator from Eq. (15.17) as the basic element of the sigma model perturbation theory. Was that a lie? No. As we explained, our approach (called the Wilsonian approach) was to obtain an effective action at a finite (nonzero) energy μ from the theory at a higher energy. Thus, the k integral is cut off from below at μ. In most interesting cases, the sigma model perturbation theory breaks down before μ comes close to zero. This is another manifestation of the infrared singularity in $1+1$ dimensions. At very low energies, we have to find a different description.

EXERCISE 15.1.2. *Show why the above argument does not apply to QFTs in more than two dimensions.*

Thus, even if the classical theory exhibits continuous global symmetry breaking, that symmetry must be restored in the quantum theory. Also, classically massless fields (like Goldstone modes) must disappear or acquire a mass in the quantum theory. How that happens is an interesting question. In some cases it is not hard to see, but in some other cases it involves very subtle dynamics of the quantum theory. We now describe a simple example, but later in the chapter we will see a more interesting example.

Let us revisit the sigma model with the target space being a circle S^1 described by a periodic scalar field $\phi \equiv \phi + 2\pi$. This system has a symmetry of shifting ϕ, the translations in the target space S^1. The would-be Goldstone mode is nothing but the field ϕ itself, but it is not a single-valued field and is excluded from the list of local operators. To see that the symmetry is indeed not broken, let us look at the one-point function of the operator $\mathrm{e}^{in\phi}$ that transforms non-trivially under the translation of ϕ. As we have seen in Sec. 11.1.3, this operator increases the target space momentum by n. However, the Hilbert space is decomposed as in Eq. (11.66) with respect to the momentum l and winding number m. What we have just said means that the operator $\mathrm{e}^{in\phi}$ sends the ground state $|0,0\rangle$ in $\mathcal{H}_{(0,0)}$ to a state in an orthogonal space $\mathcal{H}_{(n,0)}$. Thus, the one-point function vanishes: $\langle\, \mathrm{e}^{in\phi}\rangle = 0$. This shows that the symmetry is not broken in the quantum theory. In effect, $|0,0\rangle$ represents a superposition of all possible vacuum values of ϕ.

15.2. Supersymmetric Gauge Theories

We will now consider *supersymmetric* linear sigma models. First we must introduce the supersymmetric version of gauge field, gauge transformation, gauge invariant Lagrangian, etc. We consider here only the case where the gauge group is abelian.

15.2.1. Vector Multiplet. We first recall how we introduced gauge symmetries and gauge fields. Let us consider a field theory of a complex scalar field $\phi(x)$ with the Lagrangian

$$L = |\partial_\mu \phi|^2. \tag{15.18}$$

This Lagrangian is invariant under the phase rotation

$$\phi(x) \mapsto e^{i\alpha}\phi(x), \tag{15.19}$$

where α is a constant. If we let α depend on x, the derivative $\partial_\mu\phi$ transforms to $e^{i\alpha}(\partial_\mu + i\partial_\mu\alpha)\phi$ and the Lagrangian shown in Eq. (15.18) is not invariant under the phase rotation. However, if we introduce a vector field (one-form field) v_μ that transforms as $v_\mu \to v_\mu - \partial_\mu\alpha$, the modified derivative $D_\mu\phi := (\partial_\mu + iv_\mu)\phi$ transforms under the phase rotation as

$$D_\mu\phi(x) \mapsto e^{i\alpha(x)} D_\mu\phi(x) \tag{15.20}$$

and the Lagrangian

$$L = |D_\mu\phi|^2 \tag{15.21}$$

is invariant.

Now let us consider a supersymmetric field theory of a chiral superfield Φ with the Lagrangian

$$L = \int d^4\theta\, \overline{\Phi}\Phi, \tag{15.22}$$

which is invariant under the constant phase rotation $\Phi \to e^{i\alpha}\Phi$. If we replace α by a chiral superfield $A = A(x^\mu, \theta^\pm, \overline{\theta}^\pm)$, the transformation

$$\Phi \to e^{iA}\Phi \tag{15.23}$$

sends a chiral superfield to a chiral superfield. However, $\overline{\Phi}\Phi$ transforms to $\overline{\Phi}\, e^{-i\overline{A}+iA}\Phi$ and the Lagrangian is not invariant. Now, as in the case above, we introduce a *real* superfield $V = V(x^\mu, \theta^\pm, \overline{\theta}^\pm)$ that transforms as

$$V \to V + i(\overline{A} - A). \tag{15.24}$$

Then the modified Lagrangian

$$L = \int d^4\theta\, \overline{\Phi}\, e^V \Phi \tag{15.25}$$

is invariant under the transformation Eqs. (15.23)–(15.24).

A real scalar superfield V that transforms as in Eq. (15.24) under a gauge transformation is called a vector superfield. Using the gauge transformations one can eliminate the lower components of the theta-expansion of V and express it in the form

$$\begin{aligned} V =&\theta^-\overline{\theta}^-(v_0 - v_1) + \theta^+\overline{\theta}^+(v_0 + v_1) - \theta^-\overline{\theta}^+\sigma - \theta^+\overline{\theta}^-\overline{\sigma} \\ &+ i\theta^-\theta^+(\overline{\theta}^-\overline{\lambda}_- + \overline{\theta}^+\overline{\lambda}_+) + i\overline{\theta}^+\overline{\theta}^-(\theta^-\lambda_- + \theta^+\lambda_+) + \theta^-\theta^+\overline{\theta}^+\overline{\theta}^- D. \end{aligned} \tag{15.26}$$

EXERCISE 15.2.1. *Show that by a suitable gauge transformation V can be expressed as above.*

Since V is a Lorentz singlet, v_0 and v_1 define a one-form field, σ defines a complex scalar field, $\lambda_\pm$ and $\overline{\lambda}_\pm$ define a Dirac fermion field, and D is a real scalar field. The gauge in which V is represented as Eq. (15.26) is called the Wess–Zumino gauge. The residual gauge symmetry (gauge transformations that keep the form Eq. (15.26)) is the one with $A = \alpha(x^\mu)$ which transforms

$$v_\mu(x) \to v_\mu(x) - \partial_\mu\alpha(x), \tag{15.27}$$

with all the other component fields unchanged. The supersymmetry variation is given by $\delta = \epsilon_+\mathcal{Q}_- - \epsilon_-\mathcal{Q}_+ - \overline{\epsilon}_+\overline{\mathcal{Q}}_- + \overline{\epsilon}_-\overline{\mathcal{Q}}_+$ where $\mathcal{Q}_\pm$ and $\overline{\mathcal{Q}}_\pm$ are the differential operators given in Eqs. (12.6)–(12.7). The Wess–Zumino gauge is not in general preserved by this variation. In order to find the supersymmetry transformation of the component fields σ, $\lambda_\pm$, v_μ and D, we need to amend it with a gauge transformation that brings δV back into the Wess–Zumino gauge. It turns out that the required gauge transformation is the one with

$$\begin{aligned} A =&i\theta^+\left(\overline{\epsilon}_+\overline{\sigma} + \overline{\epsilon}_-(v_0 + v_1)\right) - i\theta^-\left(\overline{\epsilon}_-\sigma + \overline{\epsilon}_+(v_0 - v_1)\right) \\ &+ \theta^+\theta^-\left(\overline{\epsilon}_-\overline{\lambda}_+ - \overline{\epsilon}_+\overline{\lambda}_-\right) + \cdots, \end{aligned} \tag{15.28}$$

where $+\cdots$ are the derivative terms to make A chiral. In this way we find the following supersymmetry transformation for the component fields in Wess–Zumino gauge: For the vector multiplet fields it is

$$
\begin{aligned}
\delta v_\pm &= i\overline{\epsilon}_\pm \lambda_\pm + i\epsilon_\pm \overline{\lambda}_\pm,\\
\delta\sigma &= -i\overline{\epsilon}_+ \lambda_- - i\epsilon_- \overline{\lambda}_+,\\
\delta D &= -\overline{\epsilon}_+ \partial_- \lambda_+ - \overline{\epsilon}_- \partial_+ \lambda_- + \epsilon_+ \partial_- \overline{\lambda}_+ + \epsilon_- \partial_+ \overline{\lambda}_-,\\
\delta\lambda_+ &= i\epsilon_+ (D + iv_{01}) + 2\epsilon_- \partial_+ \overline{\sigma},\\
\delta\lambda_- &= i\epsilon_- (D - iv_{01}) + 2\epsilon_+ \partial_- \sigma.
\end{aligned}
$$

For the charged chiral multiplet fields it is

$$
\begin{aligned}
\delta\phi &= \epsilon_+ \psi_- - \epsilon_- \psi_+,\\
\delta\psi_+ &= i\overline{\epsilon}_- (D_0 + D_1)\phi + \epsilon_+ F - \overline{\epsilon}_+ \overline{\sigma}\phi,\\
\delta\psi_- &= -i\overline{\epsilon}_+ (D_0 - D_1)\phi + \epsilon_- F + \overline{\epsilon}_- \sigma\phi,\\
\delta F &= -i\overline{\epsilon}_+ (D_0 - D_1)\psi_+ - i\overline{\epsilon}_- (D_0 + D_1)\psi_-\\
&\quad + \overline{\epsilon}_+ \overline{\sigma}\psi_- + \overline{\epsilon}_- \sigma\psi_+ + i(\overline{\epsilon}_- \overline{\lambda}_+ - \overline{\epsilon}_+ \overline{\lambda}_-)\phi,
\end{aligned}
$$

where $D_\mu\phi$ and $D_\mu\psi_\pm$ are the covariant derivatives

$$D_\mu := \partial_\mu + iv_\mu, \tag{15.29}$$

with respect to the connection defined by v_μ.

The superfield

$$\Sigma := \overline{D}_+ D_- V \tag{15.30}$$

is invariant under the gauge transformation $V \to V + i(\overline{A} - A)$. It is a twisted chiral superfield

$$\overline{D}_+ \Sigma = D_- \Sigma = 0 \tag{15.31}$$

which is expressed as

$$\Sigma = \sigma(\widetilde{y}) + i\theta^+ \overline{\lambda}_+(\widetilde{y}) - i\overline{\theta}^- \lambda_-(\widetilde{y}) + \theta^+ \overline{\theta}^- [D(\widetilde{y}) - iv_{01})(\widetilde{y})], \tag{15.32}$$

in terms of the component fields in the Wess–Zumino gauge as shown by Eq. (15.26). In the above expressions $\widetilde{y}^\pm := x^\pm \mp i\theta^\pm \overline{\theta}^\pm$ and v_{01} is the field-strength of v_μ (or the curvature)

$$v_{01} := \partial_0 v_1 - \partial_1 v_0. \tag{15.33}$$

The superfield Σ is called the *super-field-strength* of V.

15.2.2. Supersymmetric Lagrangians. Let us present a supersymmetric Lagrangian for the vector multiplet V and the charged chiral multiplet Φ.

The gauge invariant Lagrangian in Eq. (15.25) is supersymmetric. In terms of the component fields it is written as

$$
\begin{aligned}
L_{\text{kin}} &= \int d^4\theta\, \overline{\Phi}\, e^V \Phi\\
&= -D^\mu\overline{\phi} D_\mu \phi + i\overline{\psi}_- (D_0 + D_1)\psi_- + i\overline{\psi}_+ (D_0 - D_1)\psi_+\\
&\quad + D|\phi|^2 + |F|^2 - |\sigma|^2|\phi|^2 - \overline{\psi}_- \sigma\psi_+ - \overline{\psi}_+ \overline{\sigma}\psi_- - i\overline{\phi}\lambda_- \psi_+\\
&\quad + i\overline{\phi}\lambda_+ \psi_- + i\overline{\psi}_+ \overline{\lambda}_- \phi - i\overline{\psi}_- \overline{\lambda}_+ \phi.
\end{aligned}
\tag{15.34}
$$

This contains the kinetic terms for the fields ϕ and $\psi_\pm$. They are *minimally* coupled to the gauge field v_μ via the covariant derivative in Eq. (15.29). They are also coupled to the scalar and fermionic components of the vector mutiplet.

The kinetic terms for the vector multiplet fields can be described in terms of the super-field-strength Σ as

$$
\begin{aligned}
L_{\text{gauge}} &= -\frac{1}{2e^2}\int d^4\theta\, \overline{\Sigma}\Sigma\\
&= \frac{1}{2e^2}\left(-\partial^\mu\overline{\sigma}\partial_\mu\sigma + i\overline{\lambda}_- (\partial_0 + \partial_1)\lambda_- + i\overline{\lambda}_+ (\partial_0 - \partial_1)\lambda_+ + v_{01}^2 + D^2\right).
\end{aligned}
\tag{15.35}
$$

Here e^2 is the gauge coupling constant and has dimensions of mass.

One can also write twisted F-terms for twisted superpotentials involving Σ. The twisted superpotential that will be important later is the linear one

$$\widetilde{W}_{FI,\theta} = -t\Sigma \tag{15.36}$$

where t is a complex parameter

$$t = r - i\theta. \tag{15.37}$$

The twisted F-term is written as

$$L_{FI,\theta} = \frac{1}{2}\left(-t\int d^2\widetilde{\theta}\ \Sigma\ +\ c.c.\right) = -rD + \theta v_{01}. \tag{15.38}$$

The parameter r is called the Fayet–Iliopoulos parameter and θ is called the theta angle. These are dimensionless parameters.

Now let us consider a supersymmetric and gauge invariant Lagrangian which is simply the sum of the above three terms

$$L = \int d^4\theta \left(\overline{\Phi} e^V \Phi - \frac{1}{2e^2}\overline{\Sigma}\Sigma \right) + \frac{1}{2}\left(-t \int d^2\widetilde{\theta}\,\Sigma \;+\; c.c. \right). \tag{15.39}$$

This Lagrangian is invariant under the vector and axial R-rotations under assigning the $U(1)_V \times U(1)_A$ charges $(0,2)$ to Σ. Thus the classical system has both $U(1)_V$ and $U(1)_A$ R-symmetries. The fields D and F have no kinetic term and can be eliminated using the equation of motion. After elimination of these auxiliary fields we obtain the Lagrangian for the other component fields

$$\begin{aligned} L = &- D^\mu\overline{\phi} D_\mu \phi + i\overline{\psi}_-(D_0 + D_1)\psi_- + i\overline{\psi}_+(D_0 - D_1)\psi_+ \\ &- \frac{e^2}{2}\left(|\phi|^2 - r\right)^2 - |\sigma|^2|\phi|^2 - \overline{\psi}_-\sigma\psi_+ - \overline{\psi}_+\overline{\sigma}\psi_- \\ &- i\overline{\phi}\lambda_-\psi_+ + i\overline{\phi}\lambda_+\psi_- + i\overline{\psi}_+\overline{\lambda}_-\phi - i\overline{\psi}_-\overline{\lambda}_+\phi \\ &+ \frac{1}{2e^2}\left(-\partial^\mu\overline{\sigma}\partial_\mu\sigma + i\overline{\lambda}_-(\partial_0+\partial_1)\lambda_- + i\overline{\lambda}_+(\partial_0 - \partial_1)\lambda_+ + v_{01}^2\right) \\ &+ \theta v_{01}. \end{aligned} \tag{15.40}$$

In particular the potential energy for the scalar fields ϕ and σ is given by

$$U = |\sigma|^2|\phi|^2 + \frac{e^2}{2}\left(|\phi|^2 - r\right)^2. \tag{15.41}$$

It is straightforward to generalize the above construction to the cases where there are many $U(1)$ gauge groups and many charged matter fields. Suppose the gauge group is $U(1)^k = \prod_{a=1}^k U(1)_a$, and there are N matter fields Φ_i, $i = 1, \ldots, N$, with charges Q_{ia} under the group $U(1)_a$ (meaning $\Phi_i \to e^{iQ_{ia}A_a}\Phi_i$). Then the generalization of the above Langrangian is

$$\begin{aligned} L = &\int d^4\theta \left(\sum_{i=1}^N \overline{\Phi}_i e^{Q_{ia}V_a}\Phi_i - \sum_{a,b=1}^k \frac{1}{2e_{a,b}^2}\overline{\Sigma}_a\Sigma_b \right) \\ &+\frac{1}{2}\left(\int d^2\widetilde{\theta} \sum_{a=1}^k(-t_a\Sigma_a) \;+ c.c. \right), \end{aligned} \tag{15.42}$$

where in the exponent of the Φ_i kinetic term the sum over $a = 1, \ldots, k$ is assumed. This is invariant under $U(1)_V \times U(1)_A$ R-rotations under the charge assignment $(0,2)$ to each Σ_a. If one can find a polynomial $W(\Phi_i)$ of Φ_i which is invariant under the gauge transformations, one can also find an

F-term

$$L_W = \int d^2\theta\, W(\Phi_i) \;+\; c.c. \tag{15.43}$$

The Lagrangian $L + L_W$ is still $U(1)_A$-invariant but $U(1)_V$-invariance holds only if $W(\Phi_i)$ is quasi-homogeneous. After eliminating the auxiliary fields D_a and F_i, we obtain the Lagrangian with the potential energy for the scalar fields being

$$\begin{aligned} U \;=\; &\sum_{i=1}^N |Q_{ia}\sigma_a|^2|\phi_i|^2 + \sum_{a,b=1}^k \frac{(e^{a,b})^2}{2}\left(Q_{ia}|\phi_i|^2 - r_a\right)\left(Q_{jb}|\phi_j|^2 - r_b\right) \\ &+\sum_{i=1}^k \left|\frac{\partial W}{\partial \phi_i}\right|^2, \end{aligned} \tag{15.44}$$

where $(e^{a,b})^2$ is the inverse matrix of $1/e_{a,b}^2$ and the summations over a and i, j are implicit.

15.3. Renormalization and Axial Anomaly

Let us consider the simplest model — $U(1)$ gauge theory with a single chiral superfield of charge 1. We consider here the effective theory at a high but finite energy scale μ. This is obtained by integrating out the modes of the fields with the frequencies in the range $\mu \le |k| \le \Lambda_{\rm UV}$, where $\Lambda_{\rm UV}$ is the ultraviolet cut-off. Let us look at the terms in the Lagrangian involving the D field

$$\frac{1}{2e^2}D^2 + D(|\phi|^2 - r_0). \tag{15.45}$$

Here r_0 is the FI parameter at the cut-off scale. Integrating out the modes of ϕ, the term $D|\phi|^2$ is replaced by $D\langle|\phi|^2\rangle$, where $\langle|\phi|^2\rangle$ is the one-point correlation function of $|\phi|^2$. The ϕ-propagator can be read from the ϕ-quadratic term in the action[4] $\frac{1}{2\pi}\int d^2x \phi^\dagger D^\mu D_\mu\phi$ and is given by

$$\langle \phi(x)\phi^\dagger(y)\rangle = \int \frac{d^2k}{(2\pi)^2}\frac{2\pi}{k^2}.$$

Thus, the one-point function in question is

$$\langle|\phi|^2\rangle = \int_{\mu\le|k|\le\Lambda_{\rm UV}} \frac{d^2k}{(2\pi)^2}\frac{2\pi}{k^2} = \log\left(\frac{\Lambda_{\rm UV}}{\mu}\right). \tag{15.46}$$

[4]We choose the action here to be related to the Lagrangian by $S = \frac{1}{2\pi}\int d^2x L$.

The momentum integral is restricted to $\mu \leq |k| \leq \Lambda_{\rm UV}$ because we are only integrating out the modes of frequencies within that range. Thus, the D-dependent terms in the effective action at the scale μ are given by

$$\frac{1}{2e^2}D^2 + D\left(\log\left(\frac{\Lambda_{UV}}{\mu}\right) - r_0\right). \tag{15.47}$$

Since the logarithm diverges in the continuum limit $\Lambda_{\rm UV} \to \infty$, in order to make the effective action finite we must give the following $\Lambda_{\rm UV}$ dependence to the bare FI parameter r_0,

$$r_0 = r + \log\left(\frac{\Lambda_{UV}}{\mu}\right). \tag{15.48}$$

r here is the renormalized FI parameter at the scale μ. Its μ dependence for a fixed theory (e.g., fixed $\Lambda_{\rm UV}$ and a fixed r_0) must be given by

$$r(\mu) = \log\left(\frac{\mu}{\Lambda}\right). \tag{15.49}$$

Λ is a finite parameter of mass dimension that determines the renormalization group flow of the FI parameter. Thus, by the quantum correction a mass scale Λ is dynamically generated. Namely, the dimensionless parameter r of the classical theory is replaced by the scale parameter Λ in the quantum theory. This is the phenomenon called *dimensional transmutation* and Λ is called a *renormalization group invariant dynamical scale.*

A related quantum effect is the anomaly of the axial R-symmetry. Recall that the classical Lagrangian is invariant under the axial R-rotation with the axial R-charge of Σ being 2 (but the charge of Φ being arbitrary). This symmetry is broken by an anomaly since there is a charged fermion. The fermion kinetic term on the Euclidean torus is

$$-2i\overline{\psi}_- D_{\bar{z}}\psi_- + 2i\overline{\psi}_+ D_z\psi_+. \tag{15.50}$$

In a gauge field background[5] with

$$k := \frac{i}{2\pi}\int i\, v_{12}dx^1dx^2 \neq 0, \tag{15.51}$$

the number of ψ_- zero modes (resp. $\overline{\psi}_+$ zero modes) is larger by k than the number of $\overline{\psi}_-$ zero modes (resp. ψ_- zero modes). The reader will note that

[5]Here we are working in the Euclidean space. The path from the Minkowski space is given by the Wick rotation, $x^0 \to -ix^2$, which also yields $v_{01} \to iv_{21} = -iv_{12}$.

$k = c_1(E)$, where E is the $U(1)$ bundle on which v is the connection. Thus, the path-integral measure changes as

$$\mathcal{D}\psi\mathcal{D}\overline{\psi} \longrightarrow e^{-2ki\alpha}\mathcal{D}\psi\mathcal{D}\overline{\psi}. \tag{15.52}$$

Since the theta angle term in the Euclidean action is $i(\theta/2\pi)\int v_{12}dx^1dx^2 = -ik\theta$, and therefore the path-integral weight is $e^{ik\theta}$, the rotation shown in Eq. (15.52) amounts to the shift in theta angle

$$\theta \longrightarrow \theta - 2\alpha. \tag{15.53}$$

Thus, the $U(1)_A$ R-symmetry of the classical system is broken to $\mathbb{Z}_2$ ($\psi \to -\psi$) in the quantum theory. One important consequence of this is that the physics does not depend on the theta angle θ since a shift of θ can be absorbed by the axial rotation, or a field redefinition.

Thus, the dimensionless parameters r and θ of the classical theory are no longer parameters of the quantum theory. They are replaced by the single scale parameter Λ.

One can repeat this argument in the case where there are N chiral superfields Φ_i of charge Q_i ($i = 1,\dots,N$). The term $D|\phi|^2$ in Eq. (15.45) is now replaced by $D\sum_{i=1}^N Q_i|\phi_i|^2$, and thus the renormalization group flow of the FI parameter is given by

$$r(\mu) = \sum_{i=1}^N Q_i \log\left(\frac{\mu}{\Lambda}\right). \tag{15.54}$$

The axial rotation shifts the theta angle as

$$\theta \longrightarrow \theta - 2\sum_{i=1}^N Q_i\alpha. \tag{15.55}$$

Thus, if $b_1 := \sum_{i=1}^N Q_i \neq 0$, the dimensional transmutation occurs and the $U(1)_A$ symmetry is anomalously broken to $\mathbb{Z}_{2b_1}$. The FI and theta parameters are replaced by the single scalar parameter Λ. If $b_1 = 0$, the FI parameter does not run as a function of the scale and the full $U(1)_A$ symmetry is unbroken. The FI and theta parameters r and θ remain as the parameters of the quantum theory.

Let us finally consider the case with the gauge group $U(1)^k = \prod_{a=1}^k U(1)_a$ and N matter fields Φ_i of charge Q_{ia}. Let us put

$$b_{1,a} := \sum_{i=1}^N Q_{ia}. \tag{15.56}$$

The FI parameters run as

$$r_a(\mu) = b_{1,a} \log\left(\frac{\mu}{\Lambda}\right) + \widetilde{r}_a \tag{15.57}$$

and the axial R-rotation shifts the theta angles as

$$\theta_a \to \theta_a - 2b_{1,a}\alpha. \tag{15.58}$$

Thus, if $b_{1,a}$ vanishes for all a, all r_a do not run and $U(1)_A$ R-symmetry is anomaly free. Thus, the FI-theta parameters $t_a = r_a - i\theta_a$ are the parameters of the theory. If $b_{1,a} \neq 0$ for some a, the parameters of the quantum theory are one scale parameter and $2k-2$ dimensionless parameters. Namely Λ, $\widetilde{r}_a$ and θ_a modulo the relation

$$(\Lambda, \widetilde{r}_a, \theta_a) \equiv (\Lambda\, e^{\delta_1}, \widetilde{r}_a + b_{1,a}\delta_1, \theta_a + b_{1,a}\delta_2). \tag{15.59}$$

The above argument applies independently of whether or not the superpotential term $\int d^2\theta\, W(\Phi_i)$ is present. The interaction induced from this does not yield divergences that renormalize the FI parameters, which is the content of the decoupling theorem presented before. Furthermore, the superpotential $W(\Phi_i)$ itself is not renormalized as long as we keep all the fields.

15.4. Non-Linear Sigma Models from Gauge Theories

Here we show that the gauged linear sigma models realize non-linear sigma models on a certain class of target Kähler manifolds. The discussion separates into two parts. In the first part, we do not turn on the superpotential for the charged matter fields. This will give us the sigma models on a class of manifolds (called *toric manifolds*) with commuting $U(1)$ isometries. In the second part, we do turn on certain types of superpotentials. This will give us the sigma model on submanifolds of toric manifolds. We start our discussion with the basic example of the $\mathbb{CP}^{N-1}$ sigma model.

15.4.1. $\mathbb{CP}^{N-1}$. Let us consider the $U(1)$ gauge theory with N chiral superfields $\Phi_1, \dots, \Phi_N$ with the Lagrangian

$$L = \int d^4\theta \left(\sum_{i=1}^N \overline{\Phi}_i\, e^V \Phi_i - \frac{1}{2e^2}\overline{\Sigma}\Sigma \right) + \frac{1}{2}\left(-t\int d^2\widetilde{\theta}\,\Sigma \; + \; c.c. \right). \tag{15.60}$$

After eliminating the auxiliary fields D and F_i, we find the following component expression of the Lagrangian:

$$\begin{aligned} L = \sum_{j=1}^N \Big[& -D^\mu\overline{\phi}_j D_\mu\phi_j + i\overline{\psi}_{j-}(D_0+D_1)\psi_{j-} + i\overline{\psi}_{j+}(D_0-D_1)\psi_{j+} \\ & -|\sigma|^2|\phi_j|^2 - \overline{\psi}_{j-}\sigma\psi_{j+} - \overline{\psi}_{j+}\overline{\sigma}\psi_{j-} - i\overline{\phi}_j\lambda_-\psi_{j+} \\ & + i\overline{\phi}_j\lambda_+\psi_{j-} + i\overline{\psi}_{j+}\overline{\lambda}_-\phi_j - i\overline{\psi}_{j-}\overline{\lambda}_+\phi_j \Big] \\ & + \frac{1}{2e^2}\Big(-\partial^\mu\overline{\sigma}\partial_\mu\sigma + i\overline{\lambda}_-(\partial_0+\partial_1)\lambda_- + i\overline{\lambda}_+(\partial_0-\partial_1)\lambda_+ + v_{01}^2 \Big) \\ & + \theta v_{01} - \frac{e^2}{2}\left(\sum_{i=1}^N |\phi_i|^2 - r \right)^2 . \end{aligned} \tag{15.61}$$

Let us look at classical supersymmetric vacua given by configurations where the potential energy

$$U = \sum_{i=1}^N |\sigma|^2|\phi_i|^2 + \frac{e^2}{2}\left(\sum_{i=1}^N |\phi_i|^2 - r \right)^2 \tag{15.62}$$

vanishes. If r is positive, $U=0$ is attained by a configuration which obeys $\sigma = 0$ and

$$\sum_{i=1}^N |\phi_i|^2 = r. \tag{15.63}$$

If $r = 0$, $U = 0$ requires all $\phi_i = 0$ but σ is free. If r is negative, $U > 0$ for every configuration, and since there could then be no zero-energy ground state, the supersymmetry appears to be spontaneously broken.

Let us examine the case of $r > 0$ in more detail. The set of all supersymmetric vacua modulo the $U(1)$ gauge group action forms the vacuum manifold. It is nothing but the complex projective space of dimension $N-1$;

$$\mathbb{CP}^{N-1} = \left\{ (\phi_1, \dots, \phi_N) \;\middle|\; \sum_{i=1}^N |\phi_i|^2 = r \right\} \Big/ U(1) \,. \tag{15.64}$$

The modes of ϕ_i's tangent to this vacuum manifold are massless. The field σ and the modes of ϕ_i's transverse to $\sum_{i=1}^N |\phi_i|^2 = r$ have mass $e\sqrt{2r}$ as can be seen by minimizing the potential in Eq. (15.62). The gauge field v_μ acquires mass $e\sqrt{2r}$ by eating the Goldstone mode — namely, the Higgs mechanism is at work. For fermions, Eq. (15.61) tells us that the modes of

$\psi_{i\pm}$ and $\overline{\psi}_{i\pm}$ obeying

$$\sum_{i=1}^{N} \overline{\phi}_i \psi_{i\pm} = 0, \ \sum_{i=1}^{N} \overline{\psi}_{i\pm} \phi_i = 0, \tag{15.65}$$

are massless (non-derivative fermion bilinear terms vanish). Other modes including the fermions in the vector multiplet have mass $e\sqrt{2r}$. The Eq. (15.65) means that the vectors $\psi_\pm = (\psi_{j\pm}, \overline{\psi}_{j\pm})$ are tangent to $\sum_{j=1}^N |\phi_j|^2 = r$ and are orthogonal to the gauge orbit $\delta(\phi_j, \overline{\phi}_j) = (i\phi_j, -i\overline{\phi}_j)$. Namely, they are tangent vectors to the vacuum manifold $\mathbb{CP}^{N-1}$ at ϕ_i. These together with the tangent modes of the ϕ_i's constitute massless supermultiplets. The massive bosonic and fermionic modes constitute a supermultiplet of mass $e\sqrt{2r}$. The latter multiplet emerges by the supersymmetric version of the Higgs mechanism — *superHiggs mechanism* — where a vector multiplet acquires mass by eating a part of the chiral multiplet.

In the limit

$$e \to \infty, \tag{15.66}$$

the massive modes decouple and the classical theory reduces to that of the massless modes only. We now show that the reduced theory can be identified as the non-linear sigma model on the vacuum manifold $\mathbb{CP}^{N-1}$. At the classical level, the resulting theory is the same as the one without the vector multiplet kinetic term, $-\frac{1}{2e^2}\int d^4\theta\, |\Sigma|^2 \to 0$. Then the vector multiplet fields are non-dynamical and the equations of motion simply yield algebraic constraints. The equations of motion for D and $\lambda_\pm$ yield the constraints Eqs. (15.63)–(15.65) on the matter fields. The equations for v_μ and σ give constraints on themselves:

$$v_\mu = \frac{i}{2} \frac{\sum_{i=1}^N \left(\overline{\phi}_i \partial_\mu \phi_i - \partial_\mu \overline{\phi}_i \phi_i\right)}{\sum_{j=1}^N |\phi_j|^2}, \tag{15.67}$$

$$\sigma = -\frac{\sum_{i=1}^N \overline{\psi}_{i+} \psi_{i-}}{\sum_{j=1}^N |\phi_j|^2}. \tag{15.68}$$

The kinetic terms for ϕ_i and $\psi_{i\pm}$ are equal to the kinetic terms of the supersymmetric non-linear sigma model on $\mathbb{CP}^{N-1}$. The metric of $\mathbb{CP}^{N-1}$ can be read off from the scalar kinetic term and is given by

$$ds^2 = \frac{1}{2\pi} \sum_{i=1}^N |D\phi_i|^2, \tag{15.69}$$

where D is the covariant derivative determined by Eq. (15.67). Since Eq. (15.67) solves the equation $\sum_{j=1}^N (\overline{\phi}_j D_\mu \phi_j - D_\mu \overline{\phi}_j \phi_j) = 0$ which states that $D_\mu \phi_j$ is orthogonal to the gauge orbit $\delta_g \phi_j = i\phi_j$, the metric $\sum_{i=1}^N |D\phi_i|^2$ measures the length of a tangent vector of $\mathbb{CP}^{N-1}$ by lifting it to a tangent vector of $\{\sum_{i=1}^N |\phi_i|^2 = r\}$ in $\mathbb{C}^N$ orthogonal to the gauge orbit. This is equal to r times the normalized Fubini–Study metric g^{FS} and thus

$$ds^2 = \frac{r}{2\pi} g^{\mathrm{FS}}. \tag{15.70}$$

The gauge field in Eq. (15.67) is the pull-back of a gauge field A on $\mathbb{CP}^{N-1}$. The gauge field A is the connection of a line bundle whose first Chern class generates the integral cohomology group $H^2(\mathbb{CP}^{N-1}, \mathbb{Z})$ (this line bundle is commonly denoted as $\mathcal{O}(1)$). The first Chern class $c_1(\mathcal{O}(1))$ is represented by the differential form $-\frac{1}{2\pi} dA$ which is equal to $\frac{1}{2\pi}$ times the Kähler form ω^{FS} of the Fubini–Study metric. Thus, the theta term $(\theta/2\pi)\int dv$ is equal to $(\theta/2\pi)\int d(\phi^* A) = -(\theta/2\pi)\int \phi^*\omega$ which is the B-field coupling with

$$B = \frac{\theta}{2\pi}\omega^{\mathrm{FS}}. \tag{15.71}$$

For the complex line $C \cong \mathbb{CP}^1$ in $\mathbb{CP}^{N-1}$ defined by (say) $\phi_1 = \cdots = \phi_{N-2} = 0$ it has a period

$$\int_C B = \theta. \tag{15.72}$$

Finally, the background value from Eq. (15.68) for σ yields the four-fermi term of the non-linear sigma model. Thus, the classical theory reduces in the limit $e \to \infty$ to the supersymmetric non-linear sigma model whose target space is the vacuum manifold $\mathbb{CP}^{N-1}$ with the metric (15.70) and the B-field as shown by Eq. (15.71).

Let us examine whether the quantum theory reduces to the non-linear sigma model as well. First of all, since the FI parameter is renormalized so that the bare or classical FI coupling r_0 is always positive (and large), there is no worry about the supersymmetry breaking associated with $r < 0$. We can focus our discussion on the $r > 0$ case as long as we look at the theory at high energies compared to the dynamical scale Λ. The effective theory of the massless modes is obtained by integrating out the massive modes. Since the latter have mass $M = e\sqrt{2r}$ this is justified when we look at the theory at the energy scale $\mu \ll e\sqrt{r}$. (To obtain the effective theory we also integrate out the part of the massless modes with frequencies in $\mu < |k| < \Lambda_{\mathrm{UV}}$.) The finite parts of the loop integrals of massive modes induce terms suppressed by powers of μ/M. There is one divergent loop which renormalizes the FI parameter in the way analyzed above. As we change the scale μ, the FI parameter runs according to the renormalization group flow. Applying Eq. (15.54), we see that the FI parameter r' at a lower energy scale μ' is obtained from the FI parameter r at the scale μ by

$$r = r' + N\log\left(\frac{\mu}{\mu'}\right). \tag{15.73}$$

On the other hand, the RG flow of the metric in the non-linear sigma model is determined by Eq. (14.59). The Ricci tensor of Eq. (15.70) is independent of the scale factor $r/2\pi$ and is equal to that of the Fubini–Study metric. The Fubini–Study metric is known to be an Einstein metric with cosmological constant N: $R^{\rm FS}_{i\bar{j}} = Ng^{\rm FS}_{i\bar{j}}$. Thus we find

$$R_{i\bar{j}} = Ng^{\rm FS}_{i\bar{j}}, \tag{15.74}$$

and Eq. (14.59) shows that the metric $g'_{i\bar{j}}$ at the lower scale μ' is determined by

$$g_{i\bar{j}} = g'_{i\bar{j}} + \frac{1}{2\pi}N\log\left(\frac{\mu}{\mu'}\right)g^{\rm FS}_{i\bar{j}}. \tag{15.75}$$

Since $g_{i\bar{j}} = (r/2\pi)g^{\rm FS}_{i\bar{j}}$ as in Eq. (15.70), this means that the metric $g'_{i\bar{j}}$ is again proportional to the Fubini–Study metric

$$g'_{i\bar{j}} = \frac{1}{2\pi}\left(r - N\log\left(\frac{\mu}{\mu'}\right)\right)g^{\rm FS}_{i\bar{j}}. \tag{15.76}$$

This is precisely the metric obtained from the linear sigma model at the scale μ' where the FI parameter is given by r' in Eq. (15.73). Thus, the RG flow of the linear sigma model matches precisely to the RG flow of the non-linear sigma model, at least to the one-loop level. Thus, we see that the linear sigma model indeed reduces at energies much smaller than $e\sqrt{r}$ to the non-linear sigma model on $\mathbb{CP}^{N-1}$. Furthermore, we have observed that the scale parameters of the two theories are in a simple relationship.

We note that the theory is parametrized by the dynamical scale Λ only and it can always be identified as the non-linear sigma model on $\mathbb{CP}^1$. This is true even though the classical theory has three "phases", $r > 0$, $r = 0$ and $r < 0$, where the interpretations are different. In particular, the supersymmetry breaking suggested by the classical analysis for $r < 0$ does not occur. This is the important effect of the renormalization. Rather, as

analyzed in Sec. 13.3, we expect that the theory has two supersymmetric vacua because $\dim H^*(\mathbb{CP}^1) = 2$. This will be examined in a later discussion in this section and also in later chapters.

We also note here one of the most important facts that can be derived by the use of the linear sigma model. If we denote by ω the Kähler class for the metric shown in Eq. (15.70), it is proportional to the Fubini–Study form: $\omega = (r/2\pi)\omega^{\rm FS}$. It follows from this and Eq. (15.71) that the complex combination $\omega - iB = \frac{r-i\theta}{2\pi}\omega^{\rm FS}$ is proportional to the complex parameter $t = r - i\theta$. This in particular means that the complexified Kähler class is given by

$$[\omega] - i[B] = \frac{t}{2\pi}[\omega^{\rm FS}], \tag{15.77}$$

or equivalently

$$\int_{\mathbb{CP}^1}\Big([\omega] - i[B]\Big) = \frac{t}{2\pi}. \tag{15.78}$$

As we have just seen, this remains true even after the effect of the renormalization is taken into accout. Since the parameter $t = r - i\theta$ is a parameter that enters into the twisted F-term, this means that *the complexified Kähler class is a twisted chiral parameter.* This parameter is a global parameter of the non-linear sigma model and the above conclusion cannot be easily obtained by the patch-wise definition which was used prior to this section.

15.4.2. Toric Manifolds. It is straightforward to generalize the above argument to more complicated examples. Let us consider the $U(1)^k = \prod_{a=1}^k U(1)_a$ gauge theory with N chiral matter fields $\Phi_1, \ldots \Phi_N$ of charges $Q_{1a}, \ldots, Q_{Na}$ under $U(1)_a$ with the Lagrangian given in Eq. (15.42). We do not consider here a superpotential term given by Eq. (15.43) and we also set $1/e^2_{a,b} = \delta_{a,b}(1/e^2_a)$.

Classical Theory. The potential for the scalar fields is given by

$$U = \sum_{i=1}^N |Q_{ia}\sigma_a|^2|\phi_i|^2 + \sum_{a=1}^k \frac{e_a^2}{2}\left(\sum_{i=1}^N Q_{ia}|\phi_i|^2 - r_a\right)^2. \tag{15.79}$$

As in the previous case, let us look at the supersymmetric vacua where U vanishes. The analysis depends on the values of the FI parameters r_a. The vacuum equation $U = 0$ requires that ϕ_i satisfy

$$\sum_{i=1}^N Q_{ia}|\phi_i|^2 = r_a, \quad a = 1, \ldots, k. \tag{15.80}$$

For now we focus on a region of r_a's where $U = 0$ imposes all $\sigma_a = 0$ via the nonzero values of some ϕ_i's required by Eq. (15.80). Then the vacuum manifold is the space of solutions to Eq. (15.80) modulo the action of the $U(1)^k$ gauge group:

$$X_r = \left\{ (\phi_1, \dots, \phi_N) \;\middle|\; \sum_{i=1}^{N} Q_{ia}|\phi_i|^2 = r_a \ (\forall a) \right\} \Big/ U(1)^k. \tag{15.81}$$

As in the previous case, the classical theory reduces in the limit $e \to \infty$ to the non-linear sigma model on X_r. The metric g of the sigma model is the one induced from the flat metric $\frac{1}{2\pi}\sum_{i=1}^{N}|d\phi_i|^2$ on $\mathbb{C}^N$: given a tangent vector of X_r, lift it to $\{\sum_{i=1}^{N} Q_{ia}|\phi_i|^2 = r_a\}$ in $\mathbb{C}^N$ as a tangent vector orthogonal to the $U(1)^k$ gauge orbit, and then measure its length using the flat metric of $\mathbb{C}^N$. The Kähler form $\omega_{\mathbb{C}^N}$ on $\mathbb{C}^N$, which can be considered as a symplectic form on $\mathbb{C}^N$, descends by the same procedure to a symplectic form ω on X_r.[6] The manifold X_r also inherits a complex structure from $\mathbb{C}^N$. As a complex manifold X_r is the same as the quotient $X_{\mathcal{P}}$ of $(\mathbb{C}^N - \mathcal{P})$ by the action of $(\mathbb{C}^\times)^k$ which is the complexification of the gauge group $U(1)^k$. Here $\mathcal{P} \subset \mathbb{C}^N$ is the locus where the $(\mathbb{C}^\times)^k$ orbit does not contain a solution to Eq. (15.80). The locus $\mathcal{P}$ depends on the choice of the values of the r_a's.[7] The complex structure is compatible with the metric and the symplectic form ω defines a Kähler form. Namely, $X_r \cong X_{\mathcal{P}}$ is a Kähler manifold and the sigma model indeed has $(2,2)$ supersymmetry as it should.

The standard $U(1)^N$ (resp. $(\mathbb{C}^\times)^N$) action on the coordinates of $\mathbb{C}^N$ descends to the $U(1)^{N-k}$ holomorphic isometry (resp. $(\mathbb{C}^\times)^{N-k}$ holomorphic automorphism) on $X_r \cong X_{\mathcal{P}} = (\mathbb{C}^N - \mathcal{P})/(\mathbb{C}^\times)^k$. The $(\mathbb{C}^\times)^{N-k}$ action is free and transitive on an open dense submanifold of $X_{\mathcal{P}} = (\mathbb{C}^N - \mathcal{P})/(\mathbb{C}^\times)^k$. Such a complex manifold is called a *toric manifold* and actually any "good" toric manifold can be realized in this way. As we noted above, the locus $\mathcal{P} \subset \mathbb{C}^N$ depends on the choice of $r = \{r_a\}$. Therefore, for another choice $r' = \{r'_a\}$, it is possible that $X_{r'}$ is isomorphic to a different complex manifold $(\mathbb{C}^N - \mathcal{P}')/(\mathbb{C}^\times)^k$ with a different $\mathcal{P}'$. More generally, for a "wrong" choice of $r = (r_a)$, the manifold X_r could be of lower dimension or even empty (like the $r \le 0$ cases of the $U(1)$ theory with N charge 1 matter fields). Given the toric manifold $X_{\mathcal{P}}$, the region of $r = (r_a)$ such that $X_r \cong X_{\mathcal{P}}$ is called the *Kähler cone* of $X_{\mathcal{P}}$. It is given as follows. We quote here a standard fact in toric geometry that holds under a certain mild assumption (See Sec. 7.4)[8]: a choice of basis of the gauge group $U(1)^k$ corresponds to a choice of basis $e_1, \dots, e_k$ of the homology group $H_2(X_{\mathcal{P}}, \mathbb{Z})$ in such a way that

$$Q_{ia} = c_1(H_i) \cdot e_a. \tag{15.82}$$

Here H_i is the line bundle over $X_{\mathcal{P}}$ having ϕ_i as a global section. Then

The FI parameter $r = (r_a)$ is in the Kähler cone of $X_{\mathcal{P}}$ if and only if

$$\sum_{a=1}^{k} m^a r_a > 0, \tag{15.83}$$

for any $m = (m^a)$ such that $\sum_{a=1}^{k} m^a e_a$ represents a holomorphic curve in $X_{\mathcal{P}}$.

The set of homology classes generated by the classes of holomorphic curves span a cone in $H_2(X_{\mathcal{P}}, \mathbb{Z})$ called the *Mori cone*, and there is a systematic way to find the generators of the cone.

Let us take $r = (r_a)$ to be in the Kähler cone of $X_{\mathcal{P}}$. One can show that the Kähler class $[\omega]$ is linear in the FI parameters r_a:

$$[\omega] = \sum_{a=1}^{k} r_a\, c_1(\mathcal{L}_a), \tag{15.84}$$

where $c_1(\mathcal{L}_a)$ is the first Chern class of the complex line bundle $\mathcal{L}_a$ over $X_{\mathcal{P}}$. The line bundle $\mathcal{L}_a$ is defined as the quotient of $(\mathbb{C}^N - \mathcal{P}) \times \mathbb{C}$ by the action of $(\mathbb{C}^\times)^k$ which acts on the last factor by $(\lambda_1, \dots, \lambda_k) : c \mapsto \lambda_a c$. In the limit $e_a \to \infty$ the worldsheet gauge fields v_μ^a are fixed to be $v_\mu^a = \frac{i}{2}M^{ab}\sum_{i=1}^{N} Q_{ib}(\overline{\phi_i}\partial_\mu\phi_i - \partial_\mu\overline{\phi_i}\phi_i)$ where M^{ab} is the inverse of the matrix $M_{ab} = \sum_{i=1}^{N} Q_{ia}Q_{ib}$. By looking at the gauge transformation property of v_μ^b under the phase rotation of the ϕ_i's, we find that v_μ^b is the pull back of the gauge field on X_r that defines a connection of the line bundle $\mathcal{L}_a$. Thus, the theta terms $\sum_{a=1}^{k}(\theta_a/2\pi)\int dv^a$ turn into the B-field coupling, where the

[6] This procedure is called the *symplectic reduction* of $(\mathbb{C}^N, \omega_{\mathbb{C}^N})$ with respect to the action of $U(1)^k$ and the moment map $\mu_a = \sum_{i=1}^{N} Q_{ia}|\phi_i|^2 - r_a$.

[7] This quotient is called the *Geometric Invariant Theory quotient* (G.I.T. quotient) of $\mathbb{C}^N$ with respect to the $\mathbb{C}^k$ action. The equivalence of X_r and $X_{\mathcal{P}} = (\mathbb{C}^N - \mathcal{P})/(\mathbb{C}^\times)^k$ as a complex manifold is a standard fact: equivalence of symplectic and G.I.T. quotients.

[8] The assumption is $S = \Sigma(1)$ in the notation of Sec. 7.4. We will see an example where this assumption is violated.

cohomology class of the B-field is given by

$$[B] = \sum_{a=1}^{k} \theta_a \, c_1(\mathcal{L}_a). \tag{15.85}$$

To summarize, in the limit $e_a \to \infty$ with the values of r_a obeying Eq. (15.83) for the choice of basis of $U(1)^k$ specified above, the classical theory reduces to the supersymmetric non-linear sigma model on the toric manifold, $X_{\mathcal{P}}$ where the cohomology classes of the Kähler form and the B field are given in Eqs. (15.84)–(15.85).

Effect of Renormalization. In the quantum theory, the FI parameters run along the RG flow as Eq. (15.57) or

$$r_a(\mu) = r_a(\mu') + b_{1,a} \log\left(\frac{\mu}{\mu'}\right), \tag{15.86}$$

where $b_{1,a} = \sum_{i=1}^{N} Q_{ia}$. In particular, the bare FI parameters $r_{0,a} = r_a(\Lambda_{\rm UV})$ are in the region where $r_{0,a} \sim b_{1,a}\lambda$ with $\lambda \gg 1$. We will show shortly that this bare FI parameter or $r = (r_a(\mu))$ at a sufficiently high energy μ is in the Kähler cone of a toric manifold $X_{\mathcal{P}}$, provided the sigma model on $X_{\mathcal{P}}$ is asymptotically free. We are interested in the effective theory of the massless modes which are tangent to the vacuum manifold X_r. In the effective theory at a scale μ much smaller than any of $e_a\sqrt{|r_a|}$, one can integrate out the massive modes. The loop integral of these massive modes and the massless modes with high frequencies induce new terms. The finite loops of the massive modes are suppressed in powers of $\mu/e_a\sqrt{|r_a|}$ and do not contribute. The divergent loop simply renormalizes the FI parameters as Eq. (15.86).

On the other hand, the Kähler class of the non-linear sigma model on $X = X_{\mathcal{P}}$ is renormalized as Eq. (14.20) or

$$[\omega](\mu) = [\omega](\mu') + \log\left(\frac{\mu}{\mu'}\right) c_1(X), \tag{15.87}$$

where $c_1(X)$ is the first Chern class of the holomorphic tangent bundle T_X. Let us compute $c_1(X)$. We first note that the holomorphic tangent bundle appears in the exact sequence

$$0 \longrightarrow \mathcal{O}^{\oplus k} \longrightarrow H_1 \oplus H_2 \oplus \cdots \oplus H_N \longrightarrow T_X \longrightarrow 0,$$

where H_i is the line bundle having ϕ_i as a global section (which appears in Eq. (15.82)). The first map sends $(l_a)_{a=1}^k$ to $(l_a Q_{ia}\phi_i)_{i=1}^N$, and the second

map sends $(f_i)_{i=1}^N$ to $\sum_{i=1}^N f_i\partial/\partial\phi_i$. Thus, the first Chern class is given by $c_1(X) = c_1(H_1 \oplus \cdots \oplus H_N) - c_1(\mathcal{O}^{\oplus k}) = \sum_{i=1}^N c_1(H_i)$. Since $c_1(H_i) = \sum_{a=1}^k Q_{ia}c_1(\mathcal{L}_a)$, this leads to the formula

$$c_1(X) = \sum_{i=1}^{N} c_1(H_i) = \sum_{a=1}^{k} b_{1,a}\, c_1(\mathcal{L}_a). \tag{15.88}$$

Noting the expression in Eq. (15.84) of the Kähler class in terms of the FI parameters, we find that the RG flow from Eq. (15.87) matches precisely with the RG flow shown in Eq. (15.86) of the FI parameter.

We learn that the complexified Kähler class is given by

$$[\omega] - i[B] = \sum_{a=1}^{k} \frac{t_a}{2\pi} c_1(\mathcal{L}_a), \tag{15.89}$$

where $t_a = r_a - i\theta_a$ is the complex parameter that enters into the twisted F-term of the Lagrangian given by Eq. (15.42). This remains true even after the renormalization effect is taken into account. Thus, as in the $\mathbb{CP}^1$ case, *the complexifield Kähler class is a twisted chiral parameter* of the theory.

Finally, as promised, we show that the FI parameters at high energies determined by Eq. (15.86) indeed form the Kähler cone of a toric manifold $X = X_{\mathcal{P}}$ provided the sigma model on X is asymptotically free. As we noted before, the non-linear sigma model is well defined only when the Ricci tensor is non-negative, or equivalently the first Chern class given by Eq. (15.88) is represented by a positive semi-definite form. A manifold is called a *Fano manifold* when it has a positive definite first Chern class and therefore the sigma model is asymptotically free. (A manifold with positive semi-definite first Chern class is called a *nef manifold*.) It follows from Eq. (15.88) that

$$b_{1,a} = c_1(X) \cdot e_a, \tag{15.90}$$

where the e_a's define a basis of $H_2(X_{\mathcal{P}}, \mathbb{Z})$ (see Eq. (15.82)). If X is a Fano manifold, $c_1(X)$ is positive definite on any holomorphic curve and therefore

$$\sum_{a=1}^{k} m^a b_{1,a} > 0, \tag{15.91}$$

for any $m = \sum_{a=1}^k m^a e_a$ representing a holomorphic curve (i.e., for any m in the Mori cone). Thus, the renormalization group flow from Eq. (15.86) shows that if X is a Fano manifold, $\sum_{a=1}^k m^a r_a(\mu)$ is positive for any such m

at sufficiently high energies. Namely, the FI parameter $r = (r_a)$ at the cut-off scale is in the Kähler cone of X and the vacuum manifold is indeed X. Therefore the quantum linear sigma model always realizes the non-linear sigma model on X. This is true even if the classical theory has different regions of $r = (r_a)$ where X_r is not isomorphic to X. The quantum theory consists of a single "phase". If X is nef but not Fano, $\sum_{a=1}^k m^a b_{1,a}$ vanishes for some curve classes $m = \sum_{a=1}^k m^a e_a$. Then the FI parameter $r = (r_a)$ at the cut-off scale is not always in the Kähler cone of $X = X_{\mathcal{P}}$ for the given $\mathcal{P}$. In order to obtain the non-linear sigma model on $X_{\mathcal{P}}$, the FI parameters in such directions should be chosen to be positive. For other choices r', $X_{r'}$ is isomorphic to $X_{\mathcal{P}'}$ with another $\mathcal{P}'$. The quantum theory consists of multiple "phases".

Condition for Asymptotic Freedom. The condition of a toric manifold to be Fano can be stated using a set of combinatorial data called a *fan*. The description of a toric variety using a fan is given in Sec. 7.2. We follow the terminology of that section as much as possible.

In Sec. 7.2, it is explained what a fan is and how to construct a toric variety $X_{\mathbf{\Sigma}}$ from a fan $\mathbf{\Sigma}$. $X_{\mathbf{\Sigma}}$ is a quotient of $\mathbb{C}^n - Z(\mathbf{\Sigma})$ by a group G defined in Sec. 7.3. In Sec. 7.3, it is explained how to construct a fan $\mathbf{\Sigma}$ from the data (Q_{ia}, r_a). Then the subset $\mathcal{P} \subset \mathbb{C}^N$ introduced above is equal to $Z(\mathbf{\Sigma})$, so that $X \cong X_{\mathbf{\Sigma}} = (\mathbb{C}^N - \mathcal{P})/(\mathbb{C}^\times)^k$.

The criterion for a toric manifold to be Fano is described as follows. Let us denote by $\Delta_{\mathbf{\Sigma}}$ the convex hull of the vertices $\mathbf{\Sigma}(1)$ and call it the polytope associated with the fan $\mathbf{\Sigma}$.[9] If the fan consists of the cones over the faces of the associated polytope $\Delta_{\mathbf{\Sigma}}$, then, $X_{\mathbf{\Sigma}}$ is Fano if and only if the polytope $\Delta_{\mathbf{\Sigma}}$ is *reflexive*. (For the definition of reflexivity, see Def. 7.10.1.)

Examples.

$U(1)$ **theories.** Consider a theory with chiral matter fields $\Phi_1, \cdots, \Phi_N$ with charges $Q_1, \cdots, Q_N$.

If all Q_i are positive, the FI parameter is large and positive at high energies. In the limit $e\sqrt{r} \to \infty$ the theory reduces to the non-linear sigma model on the vacuum manifold for positive r which is the complex weighted projective space

$$X = \mathbb{C}\mathrm{WP}^{N-1}_{[Q_1,\ldots,Q_N]}. \tag{15.92}$$

[9] In Ch. 7, $\Delta_{\mathbf{\Sigma}}$ is denoted by Δ°.

The case where Q_i are all negative is identical to this; we only have to flip the sign of r.

When there are both positive and negative Q_i's the vacuum manifold is non-compact. Let us assume $Q_1, \ldots, Q_l$ are positive and $Q_{l+1}, \ldots, Q_N$ are negative. The behavior of the theory depends on whether $\sum_{i=1}^N Q_i$ is zero or not. We discuss the two cases separately.

- $\sum_{i=1}^N Q_i > 0$

In this case, the FI parameter r is large and positive at the cut-off scale. The vacuum manifold is the $U(1)$ quotient of $\sum_{i=1}^l Q_i|\phi_i|^2 = r + \sum_{j=l+1}^N |Q_j||\phi_j|^2$, which is a non-compact manifold. This can be identified as the total space of a vector bundle over the weighted projective space;

$$X = \left[\bigoplus_{j=l+1}^N \mathcal{L}^{Q_j} \longrightarrow \mathbb{C}\mathrm{WP}^{l-1}_{[Q_1,\ldots,Q_l]}\right]. \tag{15.93}$$

- $\sum_{i=1}^N Q_i = 0$

In this case, the FI parameter r does not run and both $r > 0$ and $r < 0$ are possible. For $r > 0$, it is the sigma model on the total space of the vector bundle

$$X = \left[\bigoplus_{j=l+1}^N \mathcal{L}^{Q_j} \longrightarrow \mathbb{C}\mathrm{WP}^{l-1}_{[Q_1,\ldots,Q_l]}\right], \tag{15.94}$$

whereas for $r < 0$ it is the sigma model on the total space of another vector bundle on another weighted projective space

$$X = \left[\bigoplus_{j=1}^l \mathcal{L}^{-Q_i} \longrightarrow \mathbb{C}\mathrm{WP}^{N-l-1}_{[|Q_{l+1}|,\ldots,|Q_N|]}\right]. \tag{15.95}$$

There is a singularity classically at $r = 0$ where a new branch of free σ develops. The locus of the singularity is actually shifted by a quantum effect, as will be discussed later.

Two Examples with $\sum_{i=1}^N Q_i = 0$.

(i) $\mathcal{O}(-N)$ over $\mathbb{CP}^{N-1}$ vs $\mathbb{C}^N/\mathbb{Z}_N$.

Let us consider a $U(1)$ gauge theory with N chiral matter fields of charge 1, $\Phi_1, \ldots, \Phi_N$, and one chiral field P of charge $-N$. For $r \gg 0$ the theory describes the sigma model on the total space of the line bundle $\mathcal{L}^{-N}$ over $\mathbb{CP}^{N-1}$. This is the canonical bundle of $\mathbb{CP}^{N-1}$ and is often referred to as $\mathcal{O}(-N)$, since its first Chern class is $-N[H]$ where $[H] \in$

$H^2(\mathbb{CP}^{N-1}, \mathbb{Z})$ is the hyperplane class (Poincaré dual to a hyperplane). For $r \ll 0$, the vacuum manifold is the $U(1)$ quotient of $(\phi_1, \dots, \phi_N, p)$ obeying $N|p|^2 = |r| + \sum_{i=1}^N |\phi_i|^2$, which we can think of as a constraint on p. Since $p \neq 0$ the gauge symmetry $U(1)$ is always broken. When $\phi \neq 0$ it is completely broken, and when $\phi = 0$ it is broken to the subgroup $\mathbb{Z}_N \subset U(1)$ which fixes p but acts on the ϕ_i's as the phase rotation by N-th roots of unity. This consideration leads us to see that the vacuum manifold is $\mathbb{C}^N/\mathbb{Z}_N$.

Thus, for $r \gg 0$ the theory describes a sigma model on the total space of $\mathcal{O}(-N)$ over $\mathbb{CP}^{N-1}$, while for $r \ll 0$ it is the sigma model on the orbifold $\mathbb{C}^N/\mathbb{Z}_N$. The parameter r for the $r \gg 0$ case corresponds to the size of the base $\mathbb{CP}^{N-1}$ while the parameter r for the $r \ll 0$ case seems to have no geometric meaning. It is not clear at this moment how the theory for $r \ll 0$ depends on this parameter, but the gauge theory description of the sigma model predicts the existence of such a parameter which classically appears to be forbidden. (The dependence of the theory on this parameter will become clear when we look at the mirror description of the theory.) This is an important observation. Unlike in the case the of sigma model on a smooth manifold, the sigma model on a singular manifold can have extra parameters like this. We refer to such a theory, in the limit where $r \ll 0$ as a theory in an "orbifold phase".

(ii) $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ over $\mathbb{CP}^1$.

The next example is associated with the $U(1)$ gauge theory with four matter fields of charge $1, 1, -1, -1$. The vacuum equation reads as

$$|\phi_1|^2 + |\phi_2|^2 - |\phi_3|^2 - |\phi_4|^2 = r. \tag{15.96}$$

The vacuum manifold at $r \gg 0$ is the total space of $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ over $\mathbb{CP}^1$ where the base is that of (ϕ_1, ϕ_2), while the vacuum manifold at $r \ll 0$ is also the total space of $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ over $\mathbb{CP}^1$ where the base is that of (ϕ_3, ϕ_4). In both cases, $|r|$ parametrizes the size of the base $\mathbb{CP}^1$. At $r = 0$ (where the interpretation of the theory is not clear) the size of $\mathbb{CP}^1$ becomes zero and the vacuum manifold is singular. Denoting the gauge invariant coordinates by $x = \phi_1\phi_3$, $y = \phi_1\phi_4$, $z = \phi_2\phi_3$ and $w = \phi_2\phi_4$, the singular vacuum manifold is described by

$$xw = yz. \tag{15.97}$$

This is the so-called conifold singularity.

Toric del Pezzo Surfaces. We consider here toric Fano manifolds of dimension 2. Two-dimensional Fano manifolds are called del Pezzo surfaces and there are ten of them; $\mathbb{CP}^2$, $\mathbb{CP}^1 \times \mathbb{CP}^1$, and blow-ups of $\mathbb{CP}^2$ up to eight points. The first five of them are realized as toric manifolds. The fan and the vertices are depicted in Fig. 1.

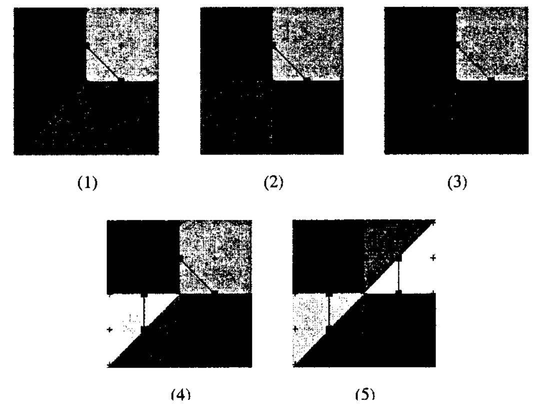


FIGURE 1. Fans for the five toric del Pezzo surfaces: The black dots are the vertices. The broken segments passing through the vertices are the boundary of the polytope

(1) $\mathbb{CP}^2$.

This has been discussed already. This is realized by $U(1)$ gauge theory with three chiral multiplets of charge 1. The Kähler cone is $r \geq 0$. A single FI-theta parameter t corresponds to the complexified Kähler parameter of $\mathbb{CP}^2$.

(2) $\mathbb{CP}^1 \times \mathbb{CP}^1$.

This is realized by $U(1)^2$ gauge theory with four chiral multiplets, two of charge $(1,0)$ and two of charge $(0,1)$. The Kähler cone is defined by $r_1 \geq 0$ and $r_2 \geq 0$. The theory at the cut-off scale is indeed in the Kähler cone since $\sum_{i=1}^4 Q_{i1} = \sum_{i=1}^4 Q_{i2} = 2$. The two FI-theta parameters t_1 and t_2 correspond to the complexified Kähler parameters of the two $\mathbb{CP}^1$ factors.

(3) One-point blow-up of $\mathbb{CP}^2$.

This is realized by $U(1)^2$ gauge theory with four chiral multiplets with charge $(1,1)$, $(1,0)$, $(1,1)$ and $(1,-1)$. The Kähler cone is given by $r_1 - r_2 \geq 0$ and $r_2 \geq 0$. The FI parameter at the cut-off scale is indeed in the Kähler cone since $\sum_{i=1}^4 Q_{i1} = 3$ and $\sum_{i=1}^4 Q_{i2} = 1$. The map

$$[\phi_1,\, \phi_2,\, \phi_3,\, \phi_4] \mapsto [\phi_1\phi_4,\, \phi_2,\, \phi_3\phi_4] \tag{15.98}$$

defines a map to $\mathbb{CP}^2$. It is an isomorphism except over the point $[0,1,0] \in \mathbb{CP}^2$, whose pre-image is a curve E isomorphic to $\mathbb{CP}^1$. The second cohomology group is generated by this "exceptional divisor" E, and by the inverse image H of the complex line in $\mathbb{CP}^2$. These are realized by $\phi_4 = 0$ and $\phi_2 = 0$, respectively. The FI parameters r_1 and r_2 are the sizes of H and E. There is another curve $H - E$ realized by $\phi_1 = 0$ (or $\phi_3 = 0$) which has size $r_1 - r_2$. The Mori cone is spanned by $H - E$ and E (that is why the Kähler cone is given by $r_1 - r_2 \geq 0$ and $r_2 \geq 0$). This surface can also be considered as a $\mathbb{CP}^1$ bundle over $\mathbb{CP}^1$ by the map

$$[\phi_1,\, \phi_2,\, \phi_3,\, \phi_4] \mapsto [\phi_1,\, \phi_3]. \tag{15.99}$$

The class of the fiber is the class of, say, the curve $\phi_1 = 0$, and therefore is $H - E$.

(4) Two-point blow-up of $\mathbb{CP}^2$.

This is realized by a $U(1)^3$ gauge theory with five chiral multiplets with charge $(1,1,0)$, $(1,0,1)$, $(1,1,1)$, $(0,-1,0)$ and $(0,0,-1)$. The Kähler cone is given by $r_1 - r_2 - r_3 \geq 0$, $r_2 \geq 0$, and $r_3 \geq 0$. The FI parameter at the cut-off scale is indeed in the Kähler cone since $\sum_{i=1}^4 Q_{i1} = 3$, $\sum_{i=1}^4 Q_{i2} = 1$, and $\sum_{i=1}^4 Q_{i3} = 1$. The map

$$[\phi_1,\, \phi_2,\, \phi_3,\, \phi_4,\, \phi_5] \mapsto [\phi_1\phi_4,\, \phi_2\phi_5,\, \phi_3\phi_4\phi_5] \tag{15.100}$$

shows that it is the blow-up of $\mathbb{CP}^2$ at two points $[0,1,0]$ and $[1,0,0]$. The second cohomology group is generated by the respective exceptional curves E_1 and E_2 plus the line H from $\mathbb{CP}^2$. The FI parameters r_2, r_3 and r_1 are the sizes of these curves.

(5) Three-point blow up of $\mathbb{CP}^2$.

This is realized by $U(1)^4$ gauge theory with six chiral multiplets whose charges are $(1,1,0,1)$, $(1,0,1,1)$, $(1,1,1,0)$, $(0,-1,0,0)$, $(0,0,-1,0)$ and $(0,0,0,-1)$. The Kähler cone is given by $r_1 - r_a - r_b \geq 0$, $r_a \geq 0$ for $a, b = 2,3,4$ and $a \neq b$. Since $\sum_{i=1}^4 Q_{i1} = 3$ and $\sum_{i=1}^4 Q_{ia} = 1$ for $a = 2,3,4$, the FI parameter at the cut-off scale is indeed in the Kähler cone. The map

$$[\phi_1,\, \phi_2,\, \phi_3,\, \phi_4,\, \phi_5,\, \phi_6] \mapsto [\phi_1\phi_4\phi_6,\, \phi_2\phi_5\phi_6,\, \phi_3\phi_4\phi_5] \tag{15.101}$$

shows that it is the blow-up of $\mathbb{CP}^2$ at three points $[0,1,0]$, $[1,0,0]$ and $[0,0,1]$. The second cohomology group is generated by the respective exceptional curves E_1, E_2 and E_3 plus the line H from $\mathbb{CP}^2$. The FI parameters r_2, r_3, r_4 and r_1 are the sizes of these curves.

EXERCISE 15.4.1. *Extend all the above examples to a non-compact Calabi–Yau threefold by introducing an extra field suitably charged under the $U(1)$'s. Also give the geometric interpretation of the extra field as a fiber coordinate.*

Nef but not Fano: An Example.

Here we consider toric manifolds which are nef but not Fano. These are the series of complex surfaces F_n $(n = 0,1,2,\ldots)$ called Hirzebruch surfaces. F_n is realized as a toric manifold with four vertices given by $v_1 = (1,0)$, $v_2 = (0,1)$, $v_3 = (-1,-n)$ and $v_4 = (0,-1)$. As one can see easily by looking at the fans in Fig. 1, F_0 is $\mathbb{CP}^1 \times \mathbb{CP}^1$ and F_1 is the one-point blow-up of $\mathbb{CP}^2$. We depict in Fig. 2 the fan for the next one F_2. As one can

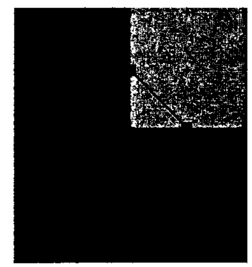


FIGURE 2. The fan for F_2

easily see, for $n \geq 3$ the set of vertices is not convex and the dual polytope is not an integral polytope.

From the vertices, we read of that the linear sigma model is the $U(1) \times U(1)$ gauge theory with four matter fields of charges $Q_1 = (0,1)$, $Q_2 = (1,0)$, $Q_3 = (0,1)$, $Q_4 = (1,-2)$. Since $\sum_{i=1}^4 Q_{i1} = 2$ the parameter r_1 is very large at the cut-off scale, but one can choose the value of r_2 since $\sum_{i=1}^4 Q_{i2} = 0$. The vacuum equations are

$$|\phi_2|^2 + |\phi_4|^2 = r_1, \tag{15.102}$$

$$|\phi_1|^2 + |\phi_3|^2 - 2|\phi_4|^2 = r_2. \tag{15.103}$$

The Mori cone for F_2 is spanned for this basis of $U(1)^2$ by e_1 and e_2. Thus, the Kähler cone for F_2 is

$$r_1 \geq 0 \text{ and } r_2 \geq 0. \tag{15.104}$$

Also, the first Chern class $c_1(F_2)$ obeys

$$c_1(F_2)\cdot e_1 = \sum_{i=1}^{4} Q_{i1} = 2,\ c_1(F_2)\cdot e_2 = \sum_{i=1}^{4} Q_{i2} = 0. \tag{15.105}$$

Thus, F_2 does not have positive first Chern class, i.e., it is not Fano. Since $c_1(F_2)$ is still non-negative it is nef. At the cut-off scale r_1 is always positive but one must choose r_2 to be positive in order to realize the sigma model on F_2.

What if we had chosen $r_2 < 0$? As long as $2r_1 + r_2 > 0$ there are solutions to the vacuum equations. The vacuum manifold is the blow-down of F_2 along the curve $\phi_4 = 0$ and has an A_1 singularity at $\phi_1 = \phi_3 = 0$.[10] Thus, the theory with $r_2 < 0$ is identified as the sigma model of the blow down of F_2 along the curve $\phi_4 = 0$. However, this variety has only one Kähler parameter if the A_1 singularity is not blown up, and there is no obvious geometric interpretation of the parameter r_2. This is actually the same as the situation encountered when we discussed $U(1)$ gauge theory with charge $1,\ldots,1,-N$ matter fields. As in that case, since the target space has a singularity (A_1 singularity which is of the type $\mathbb{C}^2/\mathbb{Z}_2$) the theory is in an "orbifold phase" and depends on a "hidden" parameter r_2 (or t_2 to be more precise). Precisely how it depends on t_2 is most explicitly seen in the mirror description.

A_{N-1} **ALE Space.** This is in a sense a generalization of the example of $\mathcal{O}(-2)$ over $\mathbb{CP}^1$. Let us consider a fan as in Fig. 3 with the vertices $v_1 = (1,0)$, $v_2 = (1,1), \ldots, v_N = (1, N-1)$, and $v_{N+1} = (1, N)$. The corresponding linear sigma model is the $U(1)^{N-1}$ theory with $N+1$ matter fields with the charges $Q_{i1} = (1,-2,1,0,\ldots,0)$, $Q_{i2} = (0,1,-2,1,\ldots,0), \ldots, Q_{i(N-1)} =$

[10]There is no solution to the vacuum equation with $\phi_4 = 0$. This means that the vertex $v_4 = (1,-2)$ must be eliminated from $\Sigma(1)$. Then the relation Eq. (15.82) does not hold in this case.

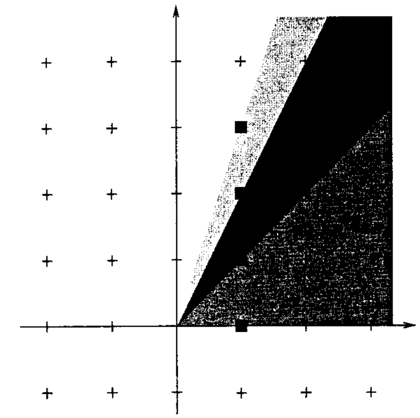


FIGURE 3. The fan for the A_3 ALE space

$(0,\ldots,1,-2,1)$. The gauge invariant variables are

$$x = 1\cdot\phi_2\cdot\phi_3^2\cdots\phi_N^{N-1}\cdot\phi_{N+1}^N, \tag{15.106}$$

$$y = \phi_1^N\cdot\phi_2^{N-1}\cdots\phi_N\cdot 1, \tag{15.107}$$

$$z = \phi_1\cdot\phi_2\cdots\phi_N\cdot\phi_{N+1}, \tag{15.108}$$

which are related by

$$xy = z^N. \tag{15.109}$$

The subvariety of $\mathbb{C}^3$ defined by this equation is a singular surface with a A_{N-1} singularity at $x = y = z = 0$. This singularity can also be represented by the orbifold $\mathbb{C}^2/\mathbb{Z}_N$. If all the FI parameters are positive, $r_a \gg 0$, the vacuum manifold is the A_{N-1} ALE space which is the minimal resolution of the singular surface shown in Eq. (15.109). There are curves $C_1,\ldots,C_{N-1}$ with the intersection relations dictated by the A_{N-1} Dynkin diagram. The curve C_a is defined by $\phi_{a+1} = 0$ and its size is given by r_a. (The equations $\phi_1 = 0$ and $\phi_{N+1} = 0$ define non-compact curves that project to $y = 0$ and $x = 0$ respectively.) If some r_a are negative, the curve C_a is blown down and the surface obtains an A_1 singularity. If all r_a are negative, all C_a are blown down and the vacuum manifold is the singular surface given by Eq. (15.109) itself. If some successive $r_a, r_{a+1},\ldots,r_{a+l-1}$ are negative, the corresponding curves shrink to zero size and the surface has an A_l singularity.

15.4.3. Hypersurfaces and Complete Intersections. So far, we have been considering gauge theories without F-terms, which reduce at low enough energies to the non-linear sigma models on toric manifolds. We can actually obtain non-linear sigma models on a certain class of submanifolds of toric manifolds by turning on a certain type of superpotential. We focus on the basic example of hypersurfaces of $\mathbb{CP}^{N-1}$, which captures the essential point. We also briefly discuss the complete intersections of hypersurfaces in $\mathbb{CP}^{N-1}$. We leave the cases of hypersurfaces or complete intersections in more general toric manifolds as exercises for the reader.

Hypersurfaces in $\mathbb{CP}^{N-1}$. Let us consider a degree d polynomial of $\phi_1, \dots, \phi_N$;

$$G(\phi_1, \dots, \phi_N) = \sum_{i_1, \dots, i_d} a_{i_1 \dots i_d} \phi_{i_1} \cdots \phi_{i_d}. \tag{15.110}$$

We assume that $G(\phi_i)$ is generic in the sense that

$$G = \frac{\partial G}{\partial \phi_1} = \cdots = \frac{\partial G}{\partial \phi_N} = 0 \text{ implies } \phi_1 = \cdots = \phi_N = 0. \tag{15.111}$$

Then the complex hypersurface M of $\mathbb{CP}^{N-1}$ defined by

$$G(\phi_1, \dots, \phi_N) = 0 \tag{15.112}$$

is a smooth complex manifold of complex dimension $N-2$. The Kähler form of $\mathbb{CP}^{N-1}$ restricts to a Kähler form on M. It is known that the second cohomology group is one-dimensional and is generated by the restriction of the class $[H] := c_1(\mathcal{O}(1))$ which is represented by a positive-definite Kähler form (up to normalization). The first Chern class of M is equal to

$$c_1(M) = (N-d)[H]|_M. \tag{15.113}$$

So, M is Ricci positive for $d < N$, Calabi–Yau for $d = N$, and Ricci negative for $d > N$. The non-linear sigma model on M is asymptotically free, scale invariant, and infrared free, respectively.

Now, let us consider a $U(1)$ gauge theory with $N+1$ chiral multiplets $\Phi_1, \cdots, \Phi_N, P$ of charge $1, \cdots, 1, -d$. Then the superpotential

$$W = P\, G(\Phi_1, \dots, \Phi_N) \tag{15.114}$$

is gauge invariant. We consider the following Lagrangian

(15.115)
$$\begin{aligned} L = \int d^4\theta \left(\sum_{i=1}^N \overline{\Phi}_i \, e^V \Phi_i + \overline{P} \, e^{-dV} P - \frac{1}{2e^2} \overline{\Sigma}\Sigma \right) \\ + \frac{1}{2} \left(\int d^2\widetilde{\theta} (-t\Sigma) + c.c. \right) + \frac{1}{2} \left(\int d^2\theta \, P G(\Phi_1, \cdots, \Phi_N) + c.c. \right). \end{aligned}$$

The potential for the scalar field is given by

$$\begin{aligned} U =& |\sigma|^2 \sum_{i=1}^N |\phi_i|^2 + |\sigma|^2 d^2 |p|^2 + \frac{e^2}{2} \left(\sum_{i=1}^N |\phi_i|^2 - d|p|^2 - r \right)^2 \\ &+ \frac{1}{4} \Big| G(\phi_1, \dots, \phi_N) \Big|^2 + \frac{1}{4} \sum_{i=1}^N |p|^2 |\partial_i G|^2. \end{aligned} \tag{15.116}$$

Let us analyze the spectrum of the classical theory. The structure of the classical supersymmetric vacuum manifold $U = 0$ is different for $r > 0$ and $r < 0$, and we will treat these two cases (along with the case $r = 0$) separately.

$r > 0$, $U = 0$ requires some $\phi_i \neq 0$ and therefore $\sigma = 0$. If $p \neq 0$, $U = 0$ further requires $G = \partial_1 G = \cdots = \partial_N G = 0$ which implies by the condition Eq. (15.111) that all $\phi_i = 0$. However, this contradicts $\phi_i \neq 0$ for some i. Thus p must be zero. We thus find that $U = 0$ is attained by $\sigma = p = 0$ and

$$\sum_{i=1}^N |\phi_i|^2 = r, \;\; G(\phi_1, \dots, \phi_N) = 0. \tag{15.117}$$

The vacuum manifold is the set of (ϕ_i) obeying these equations, divided by the $U(1)$ gauge group action. This is nothing but the hypersurface M. The modes of ϕ_i tangent to the manifold M are massless. Other modes are massive. Some have mass of order $e\sqrt{r}$ as in the case without superpotential, but some others have mass determined by G or its coefficients $a_{i_1 \dots i_d}$ in Eq. (15.110). If we send e and $a_{i_1 \dots i_d}$ to infinity by an overall scaling, all the massive modes decouple and the classical theory reduces to the non-linear sigma model on the hypersurface M, with the complexified Kähler class given by $[\omega] - i[B] = (t/2\pi)[\omega^{\mathrm{FS}}]|_M$.

$r < 0$, $U = 0$ requires $p \neq 0$ and thus $\sigma = 0$. Under the condition Eq. (15.111), $U = 0$ then requires all $\phi_i = 0$. p is thus constrained to the circle

$$|p| = \sqrt{|r|/d}. \tag{15.118}$$

The gauge group acts on this circle transitively, and therefore the vacuum manifold is a single point. A choice of vacuum value of p, say $\langle p\rangle = \sqrt{|r|/d}$, breaks the $U(1)$ gauge symmetry. Therefore, the vector multiplet fields together with the P-multiplet fields have mass $e\sqrt{|r|/d}$ by the superHiggs mechanism. The fields Φ_i are all massless as long as the degree d of the polynomial $G(\Phi_i)$ is larger than two, $d > 2$. If we take the limit $e \to \infty$, the classical theory reduces to the theory of Φ_i's only. It is the Landau–Ginzburg theory with the superpotential

$$W = \langle p\rangle G(\Phi_1, \ldots, \Phi_N), \tag{15.119}$$

where $\langle p\rangle$ is the vacuum value of p (say $\langle p\rangle = \sqrt{|r|/d}$). We should notice, however, that the gauge group $U(1)$ is not completely broken by the choice of the value of p. Since p has charge $-d$, the discrete subgroup $\mathbb{Z}_d \subset U(1)$ remains unbroken. This discrete subgroup acts non-trivially as the gauge symmetry of the low energy theory of the charge 1 massless modes Φ_i. Thus, the Landau–Ginzburg theory is not the ordinary one, but a $\mathbb{Z}_d$ gauge theory. Physical fields must be invariant under the $\mathbb{Z}_d$ action, and the configuration must be single-valued only up to the $\mathbb{Z}_d$ action. Such a gauge theory is usually called an *orbifold theory* and our low energy theory is called a *Landau–Ginzburg orbifold*.

$r = 0$, $U = 0$ requires $\sum_{i=1}^N |\phi_i|^2 = d|p|^2$. If $p \neq 0$, some $\phi_i \neq 0$. However, $U = 0$ with $p \neq 0$ requires $G = \partial_1 G = \cdots = \partial_N G = 0$ which means by the condition Eq. (15.111) $\phi_1 = \cdots = \phi_N = 0$, a contradiction. Thus p must be zero and $\phi_i = 0$. Then σ is free. The vacuum manifold is the complex σ-plane. Σ multiplet fields are always massless. At $\sigma \neq 0$ other modes are massive, but they become massless at $\sigma = 0$.

In the quantum theory, we must take into account the renormalization of the FI parameter r. It depends on whether $b_1 = N - d$ is positive, zero, or negative. We separate the discussion into these three cases.

- $d < N$.

In this case, the theory is parametrized by the dynamically generated scale Λ which determines the RG flow of the FI parameter

$$r(\mu) = (N - d)\log(\mu/\Lambda). \tag{15.120}$$

At the cut-off scale $\Lambda_{\rm UV}$ or at a scale much larger than Λ, the FI parameter is positive and very large: $r \gg 1$. Thus, the first case of the above argument applies. In particular, by taking the limit where $e/\Lambda \to \infty$ and $a_{i_1\ldots i_d}/\Lambda \to \infty$, the theory reduces to the non-linear sigma model on the hypersurface M. Since

$$c_1(M) = (N - d)[H]|_M \tag{15.121}$$

is positive, the sigma model is asymptotically free. The logarithmic running of the Kähler parameter of the non-linear sigma model is proportional to Eq. (15.121) and matches precisely with the logarithmic running in Eq. (15.120) of the FI parameter.

- $d = N$.

In this case, the FI parameter does not run and the theory is parametrized by $t = r - i\theta$. In particular, we can choose the value of r as we wish. We separate the discussion into three cases.

For $r \gg 0$, the theory reduces in the limit $e\sqrt{r} \to \infty$ and $a_{i_1\ldots i_d} \to \infty$ to the non-linear sigma model on the hypersurface M. Since M is a Calabi–Yau manifold, $c_1(M) = 0$, the Kähler class of the sigma model does not run, which agrees with the fact that r does not run, either. The complexified Kähler class is identified as t at large r.

For $r \ll 0$, the theory reduces in the limit $e\sqrt{|r|} \to \infty$ to the LG orbifold.

For $r = 0$, the σ branch develops. It is a non-compact flat direction and the theory must exhibit some kind of singularity when approached from $r \gg 0$ or $r \ll 0$. The behavior of the theory near $r = 0$ is modified by several quantum effects and the theta angle θ plays an important role. This will be discussed later in this chapter.

- $d > N$.

In this case, the FI parameter at the cut-off scale is large and negative. Thus, the theory at high energies does not describe the non-linear sigma model on the hypersurface M but looks to be closer to the LG orbifold. The LG orbifold itself is a superconformal field theory and must preserve the axial R-symmetry. On the other hand, the gauge theory preserves only the discrete subgroup $\mathbb{Z}_{2(d-N)}$ and contains a running coupling (the FI parameter). Thus, it would be appropriate to identify the model as the LG orbifold perturbed by a relevant operator that breaks the $U(1)$ axial R-symmetry to $\mathbb{Z}_{2(d-N)}$.

Complete Intersections in $\mathbb{CP}^{N-1}$. Let us consider degree d_r polynomials G_r of $\phi_1, \ldots, \phi_N$ $(r = 1, \ldots, \ell)$. We assume that these functions are generic in the sense that $G_r = 0$ for all r and $\sum_{r=1}^{\ell} p_r\partial_i G_r$ for some p_r and for all i

require $\phi_1 = \cdots = \phi_N = 0$. Then the submanifold

$$M = \{\ G_r(\phi_1, \ldots, \phi_N) = 0\ \forall r\ \} \subset \mathbb{CP}^{N-1} \tag{15.122}$$

is a smooth complex manifold of dimension $N - \ell - 1$. This submanifold M is called the *complete intersection* of $\{G_1 = 0\}, \ldots, \{G_\ell = 0\}$ in $\mathbb{CP}^{N-1}$. The second cohomology group includes a positive class $[H]|_M$ (which in fact generates $H^2(M, \mathbb{Z})$ if $\dim M \geq 3$) and the first Chern class is given by

$$c_1(M) = (N - d_1 - \cdots - d_\ell)[H]|_M. \tag{15.123}$$

Let us consider a $U(1)$ gauge theory with chiral multiplets $\Phi_1, \ldots, \Phi_N$, $P_1, \ldots, P_\ell$ of charge $1, \ldots, 1, -d_1, \ldots, -d_\ell$. We consider the action that includes the F-term with the superpotential

$$W = \sum_{r=1}^{\ell} P_r\, G_r(\Phi_1, \ldots, \Phi_N). \tag{15.124}$$

Then one can show that at $r > 0$ the vacuum manifold is the complete intersection M.

15.5. Low Energy Dynamics

In the previous discussion, we have identified the gauge theories as non-linear sigma models (and Landau–Ginzburg models in some cases) by looking at energies which are smaller than the coupling $e\sqrt{r}$ but are considered as high energies from the point of view of the non-linear sigma models. We now attempt to describe the physics of the linear sigma models at much lower energies in order to learn about the low energy dynamics of the non-linear sigma models. In the case where the theory undergoes a dimensional transmutation we will look at energies μ smaller than the dynamical scale Λ. In the scale invariant theories, we will probe the region where r is close to zero.

15.5.1. The Behaviour at Large Σ. It turns out that it is useful in many ways to look at the behavior of the theory where the lowest component σ of the super-field-strength Σ is taken to be large and slowly varying. Let us look at the σ-dependent terms in the kinetic term of the charged matter field Φ. From Eq. (15.34) we read off that it is

$$-|\sigma|^2|\phi|^2 - \overline{\psi}_-\sigma\psi_+ - \overline{\psi}_+\overline{\sigma}\psi_-. \tag{15.125}$$

We see that σ plays the role of the mass for the field Φ. Taking σ large means making Φ heavy. We are thus looking at the region in the field space where there are heavy charged matter fields.

$(1+1)$-Dimensional Gauge Theory with Heavy Charged Particles. To be specific, let us consider a $U(1)$ gauge theory with several charged chiral superfields Φ_i. At large σ the charged matter fields are heavy and the massless degrees of freedom are only the Σ multiplet itself. The theory is that of a $U(1)$ gauge theory in $1+1$ dimensions with heavy charged fields.

Let us compute the vacuum energy of the system. Since the Φ_i's are heavy, they are frozen at the zero expectation value and one can set $\Phi_i = 0$ classically. Then the potential energy is given by

$$U_r = \frac{e^2}{2} r^2. \tag{15.126}$$

The contributions to the vacuum energy from σ and $\lambda_\pm$ cancel against each other because of the supersymmetry. There is actually a contribution to the energy density from the gauge field v_μ. The terms in the action that depend on the gauge field are

$$S = \frac{1}{2\pi}\int d^2x \left(\frac{1}{2e^2}v_{01}^2 + \theta v_{01}\right). \tag{15.127}$$

Let us quantize the system by compactifying the spatial direction on S^1 so that x^1 is a periodic coordinate of period 2π, $x^1 \equiv x^1 + 2\pi$. By using gauge transformations $v_\mu \to v_\mu - \partial_\mu\gamma$, one can set

$$v_0 = 0,\ v_1 = a(t), \tag{15.128}$$

where $a(t)$ depends only on $t = x^0$. The gauge transformation $\gamma = mx^1$ preserves this form. This is an allowed gauge transformation provided m is an integer since $e^{i\gamma} = e^{imx^1}$ is single-valued if $m \in \mathbb{Z}$. Thus there is a gauge equivalence relation

$$a(t) \equiv a(t) + m,\ m \in \mathbb{Z}. \tag{15.129}$$

In terms of this variable, the action is given by

$$S = \int dt \left(\frac{1}{2e^2}\dot{a}^2 + \theta\dot{a}\right). \tag{15.130}$$

The transition amplitude (see Eq. (10.11)) from a state Ψ_i at time t_i to a state Ψ_f at time t_f is given by the path-integral

(15.131)
$$\langle \Psi_f, \mathrm{e}^{-i(t_f-t_i)H}\Psi_i\rangle = \int da_f da_i \Psi_f^*(a_f) \int_{\substack{a(t_f)=a_f\\ a(t_i)=a_i}} \mathcal{D}a\, \mathrm{e}^{iS}\, \Psi_i(a_i)$$
$$= \int da_f da_i \Psi_f^*(a_f)\, \mathrm{e}^{i\theta a_f} \int_{\substack{a(t_f)=a_f\\ a(t_i)=a_i}} \mathcal{D}a\, \mathrm{e}^{i\int_{t_i}^{t_f}\frac{1}{2e^2}\dot{a}^2 dt}\, \mathrm{e}^{-i\theta a_i}\Psi_i(a_i)$$

This shows that the Hamiltonian acts on the phase-rotated wave-functions $\widetilde{\Psi}(a) = \mathrm{e}^{-i\theta a}\Psi(a)$ as $\frac{e^2}{2}\left(-i\frac{d}{da}\right)^2$. Namely, it acts on the ordinary wave-functions as

$$H\Psi(a) = \frac{e^2}{2}\left(-i\frac{d}{da} - \theta\right)^2 \Psi(a). \tag{15.132}$$

We recall that a is a periodic variable as shown by Eq. (15.129). Thus, single-valued wave-functions $\Psi(a)$ are expressed as linear combinations of the Fourier modes $\mathrm{e}^{2\pi nia}$ with $n \in \mathbb{Z}$. These Fourier modes are actually the energy eigenfunctions. Thus, the spectrum is

$$E_n = \frac{e^2}{2}(2\pi n - \theta)^2. \tag{15.133}$$

The ground state energy is therefore given by

$$E_{\rm vac} = \frac{e^2}{2}\hat{\theta}^2 \tag{15.134}$$

where $\hat{\theta}^2$ is defined by

$$\hat{\theta}^2 := \min_{n\in\mathbb{Z}}\{(\theta - 2\pi n)^2\}. \tag{15.135}$$

This total energy $E_{\rm vac}$ can be considered also as the vacuum energy density since $\frac{1}{2\pi}\int dx^1 = 1$ in the present set-up. What is the value of the field strength at the ground state? To see this, we note that the conjugate momentum for a is given by $p_a = \frac{\partial L}{\partial \dot{a}} = \frac{\dot{a}}{e^2} + \theta$. [11] From this we see that

$$v_{01} = -e^2\theta + e^2 p_a. \tag{15.136}$$

[11] In fact a naive canonical quantization also leads to the result from Eq. (15.132); $H = p_a\dot{a} - L = \frac{e^2}{2}(p_a - \theta)^2$.

Namely, the field strength $v_{01} = \dot{a}$ is equal to $-e^2\theta$ up to integer multiples of $2\pi e^2$. In particular, the magnitude of the vacuum value of v_{01} is

$$|v_{01}|_{\rm vac} = e^2|\hat{\theta}|. \tag{15.137}$$

The vacuum value of v_{01} is thus discontinuous as a function of θ. There is an intuitive understanding of this discontinuity, due to Coleman, which applies when the theory is formulated on $\mathbb{R}^2$. We assume that the mass M of the charged particle is much larger than the gauge coupling, $M \gg e$, so that the charged particles can be treated semi-classically. If we put a charged particle of charge Q at $x^1 = 0$, it generates a field strength v_{01} which obeys

$$\partial_1 v_{01} = 2\pi Q e^2 \delta(x^1). \tag{15.138}$$

Namely, it generates a gap of v_{01} by $2\pi Q e^2$. Now suppose θ is positive but smaller than π. Then there is a unique ground state with the field strength $v_{01} = -e^2\theta$ and the energy density $U = \frac{e^2}{2}\theta^2$. One cannot have a single charged particle since that would make $v_{01}(+\infty)$ different from $v_{01}(-\infty)$ but v_{01} is required to take the (unique) vacuum value at both spatial infinities. However, one can have particles of total charge zero. For instance, let us consider the situation where we have one with charge 1 at $x^1 = -L/2$ and one with charge -1 at $x^1 = L/2$. Outside the interval $-L/2 \le x^1 \le L/2$ the field strength takes the vacuum value $-e^2\theta$, while it takes the value $-e^2\theta + 2\pi e^2$ inside that interval. The energy of that configuration compared to the one for the vacuum state with $v_{01} \equiv -e^2\theta$ is

$$\Delta E = \left(\frac{e^2}{2}(2\pi - \theta)^2 - \frac{e^2}{2}\theta^2\right) L. \tag{15.139}$$

As long as $\theta < \pi$, this energy is positive and is proportional to the separation L. To decrease the energy, the separation L is reduced to zero. Namely, there is an attractive force between the particles of opposite charge. Charged particles cannot exist in isolation; they are *confined*. Now let us increase θ so that $\theta > \pi$. Then ΔE is negative. It is now energetically favorable for the separation L to be larger and there is a replusive force. Eventually, the two particles are infinitely separated and disappear to the negative and positive infinities in x^1. What is left is the field strength with the value $v_{01} = -e^2\theta + 2\pi e^2$. The absolute value is nothing but $e^2|\hat{\theta}|$ for θ in the range $\pi < \theta < 3\pi$. Even if we start without the particles of opposite charges, they can be created and go off to infinity. Creating a pair costs an energy $2M$, but the negative energy ΔE for large L is enough to cancel it. Effectively, the field strength is reduced by $2\pi e^2$. This is the intuitive explanation of the discontinuity. A similar thing happens when θ goes beyond 3π, $5\pi, \ldots$ or when θ decreases in the negative direction and goes below $-\pi$, $-3\pi, \ldots$.[12]

The total energy density is thus the sum of Eqs. (15.126)–(15.134)

$$U = \frac{e^2}{2}\left(r^2 + \hat{\theta}^2\right) = \frac{e^2}{2}|\hat{t}|^2. \tag{15.140}$$

We notice that this expression is almost the same as the potential energy of the Landau–Ginzburg model obtained by setting Φ_i to zero and considering Σ as the ordinary twisted chiral superfield having the twisted superpotential

$$W(\Sigma) = -t\Sigma. \tag{15.141}$$

That Σ is not really an ordinary twisted chiral superfield but the superfield-strength (the imaginary part of the auxiliary field is the curvature v_{01}) has only a minor effect: the shift in θ by 2π times an integer.

This story, however, can be further modified by quantum effects. In the above discussion we considered Φ_i to be totally frozen. But of course we must take into account the quantum fluctuation of the Φ_i's. What it does is to modify the FI-theta parameter as a function of σ. Let us now analyze this.

Effective Action for Σ. Let us first consider the basic example of the $U(1)$ gauge theory with a single chiral superfield Φ of charge 1, without F-term (which is not allowed in this case). Let us take σ to be slowly varying and large compared to the energy scale μ where we look at the effective theory. The Φ multiplet has a mass of order $\sigma \gg \mu$ and therefore it is appropriate to describe the effective theory in terms of the low frequency modes of Σ only. Thus, the effective action at energy μ is obtained by integrating out all the modes of Φ and the modes of Σ with frequencies in the range $\mu \le |k| \le \Lambda_{\mathrm{UV}}$. By supersymmetry, the terms with at most two derivatives and not more

[12]Here we are assuming that there is a matter field of charge 1, or the greatest common divisor of the charges Q_i is 1. If the g.c.d. of Q_i's is $q > 1$, the critical value of θ is $q\pi$ (times an odd integer) and the definition of $\hat{\theta}^2$ is replaced by

$$\hat{\theta}^2 := \min_{n\in\mathbb{Z}}\{(\theta + 2\pi q n)^2\}.$$

Thus, in such a case the physics is periodic in θ with period $2\pi q$.

than four fermions are constrained to be of the form

$$S_{\mathrm{eff}}(\Sigma) = \int d^4\theta(-K_{\mathrm{eff}}(\Sigma,\overline{\Sigma})) + \frac{1}{2}\left(\int d^2\widetilde{\theta}\ \widetilde{W}_{\mathrm{eff}}(\Sigma) + c.c.\right). \tag{15.142}$$

We try to compute these terms in two steps: integrate out Φ first, then the high frequency modes of Σ. Since the action $S(\Sigma,\Phi)$ is quadratic in Φ, the first step can be carried out exactly by the one-loop computation

$$e^{iS^{(1)}_{\mathrm{eff}}(\Sigma)} = \int \mathcal{D}\Phi\ e^{iS(\Sigma,\Phi)}. \tag{15.143}$$

As we will see, the effective superpotential $\widetilde{W}^{(1)}_{\mathrm{eff}}(\Sigma)$ will not be further corrected by the second step (a non-renormalization theorem). Thus, the focus will be on obtaining $\widetilde{W}_{\mathrm{eff}}$ by the first step.

Since $\Sigma = \sigma + \theta^+\overline{\theta}^-(D - iv_{01}) + \cdots$, the dependence of the effective action on D and v_{01} is as follows. From the D-term we obtain

$$\int d^4\theta(-K_{\mathrm{eff}}(\Sigma,\overline{\Sigma})) = \partial_\sigma\partial_{\overline{\sigma}}K_{\mathrm{eff}}(\sigma,\overline{\sigma})|D - iv_{01}|^2 + \cdots. \tag{15.144}$$

From the twisted F-terms we have

(15.145)
$$\begin{aligned}\frac{1}{2}\left(\int d^2\widetilde{\theta}\widetilde{W}_{\mathrm{eff}}(\Sigma)\ +\ c.c.\right) =& \frac{1}{2}\left(\partial_\sigma\widetilde{W}_{\mathrm{eff}}(\sigma)(D - iv_{01}) + c.c.\right)\\ =& D\mathrm{Re}\left[\partial_\sigma\widetilde{W}_{\mathrm{eff}}(\sigma)\right] + v_{01}\mathrm{Im}\left[\partial_\sigma\widetilde{W}_{\mathrm{eff}}(\sigma)\right] + \cdots.\end{aligned}$$

Thus, in order to determine $\widetilde{W}_{\mathrm{eff}}$ it is enough to look at the D and v_{01} linear terms in the effective action. The Kähler potential can be determined by the D-quadratic term. To simplify the computation one can set $\lambda_\pm = \overline{\lambda}_\pm = 0$ without losing any information. In this case the Φ-dependent part of the (Euclidean) classical action is

$$\begin{aligned}L^E_{\mathrm{kin}} =& |D_\mu\phi|^2 + |\sigma|^2|\phi|^2 - D|\phi|^2\\ &- 2i\overline{\psi}_-D_{\overline{z}}\psi_- + 2i\overline{\psi}_+D_z\psi_+ + \overline{\psi}_-\sigma\psi_+ + \overline{\psi}_+\overline{\sigma}\psi_-.\end{aligned} \tag{15.146}$$

We are going to evaluate

$$e^{-\frac{1}{2\pi}\int \Delta L^{(1)}_E d^2x} := \int \mathcal{D}^2\phi\mathcal{D}\psi\mathcal{D}\overline{\psi}\, e^{-\frac{1}{2\pi}\int L^E_{\mathrm{kin}}d^2x}. \tag{15.147}$$

The dependence of $\Delta L^{(1)}_E$ on the phase of $\sigma = |\sigma|e^{i\gamma}$ is easy to obtain. At the classical level, this phase can be absorbed by the phase rotation of the fermions

$$\psi_\pm \to e^{\mp i\gamma/2}\psi_\pm,\ \overline{\psi}_\pm \to e^{\pm i\gamma/2}\overline{\psi}_\pm. \tag{15.148}$$

However, this is the axial rotation which is anomalous. The effect is thus the shift in the theta angle noted before. In other words the effective action for σ is related to that for $|\sigma|$ by

$$\Delta L_E^{(1)}(\sigma) = \Delta L_E^{(1)}(|\sigma|) - i\gamma v_{12} = \Delta L_E^{(1)}(|\sigma|) - i\arg(\sigma)v_{12}. \tag{15.149}$$

Now, $\Delta L_E^{(1)}(|\sigma|)$ is given by

$$e^{-\frac{1}{2\pi}\int \Delta L_E^{(1)}(|\sigma|)d^2x} = \frac{\det\begin{pmatrix} |\sigma| & 2iD_{\bar z} \\ 2iD_z & -|\sigma| \end{pmatrix}}{\det\Big(-D_\mu D_\mu + |\sigma|^2 - D\Big)}. \tag{15.150}$$

The square of the Dirac operator is

$$\begin{aligned} \begin{pmatrix} |\sigma| & 2iD_{\bar z} \\ 2iD_z & -|\sigma| \end{pmatrix}^2 &= \begin{pmatrix} -4D_{\bar z}D_z + |\sigma|^2 & 0 \\ 0 & -4D_zD_{\bar z} + |\sigma|^2 \end{pmatrix} \\ &= \begin{pmatrix} -D_\mu D_\mu + |\sigma|^2 - v_{12} & 0 \\ 0 & -D_\mu D_\mu + |\sigma|^2 + v_{12} \end{pmatrix}, \end{aligned} \tag{15.151}$$

where we have used the relation $D_{\bar z}D_z = \frac{1}{2}(D_{\bar z}D_z + D_zD_{\bar z}) + \frac{1}{2}[D_{\bar z}, D_z] = \frac{1}{4}D_\mu D_\mu + \frac{1}{2}iv_{z\bar z}$. Thus, we obtain

$$\begin{aligned} \frac{1}{2\pi}\int \Delta L_E^{(1)}(|\sigma|)d^2x &= \log\det(-D_\mu D_\mu + |\sigma|^2 - D) \\ &\quad - \frac{1}{2}\log\det(-D_\mu D_\mu + |\sigma|^2 - v_{12}) \\ &\quad - \frac{1}{2}\log\det(-D_\mu D_\mu + |\sigma|^2 + v_{12}). \end{aligned} \tag{15.152}$$

There is no v_{12} linear term in this relation but there is a D-linear term. It is given by

$$\frac{1}{2\pi}\int \Delta L_E^{(1)}(|\sigma|)d^2x\Bigg|_{D-\text{linear}} = \mathrm{Tr}\left(\frac{-D}{-\partial_\mu\partial_\mu + |\sigma|^2}\right). \tag{15.153}$$

Namely, we have

(15.154)
$$\Delta L_E^{(1)}(|\sigma|)\Bigg|_{D-\text{linear}} = -D\int\limits_{|k|\le\Lambda_{\rm UV}} \frac{d^2k}{(2\pi)^2}\frac{2\pi}{k^2+|\sigma|^2} = -\frac{1}{2}D\log\left(\frac{\Lambda_{\rm UV}^2+|\sigma|^2}{|\sigma|^2}\right).$$

Similarly we can read the D-quadratic term from Eq. (15.152) as

$$\begin{aligned} \Delta L_E^{(1)}(|\sigma|)\Bigg|_{D-\text{quadratic}} &= -\frac{1}{2}D^2\int\limits_{|k|\le\Lambda_{\rm UV}} \frac{d^2k}{(2\pi)^2}\frac{2\pi}{(k^2+|\sigma|^2)^2} \\ &= -\frac{1}{4|\sigma|^2}D^2\frac{1}{1+\frac{|\sigma|^2}{\Lambda_{\rm UV}^2}}. \end{aligned} \tag{15.155}$$

To summarize, we have

$$\Delta L_E^{(1)}(\sigma) = -\log\left(\frac{\Lambda_{\rm UV}}{|\sigma|}\right)D - i\arg(\sigma)v_{12} - \frac{1}{4|\sigma|^2}(D^2 - v_{12}^2) + \cdots, \tag{15.156}$$

where $+\cdots$ are the terms which are neither linear nor quadratic in (D, v_{12}), and we have neglected the powers of $|\sigma|/\Lambda_{\rm UV}$ which vanish in the continuum limit. Noting the relation of the Euclidean and Minkowski Lagrangians $L_E = -L|_{x^0=-ix^2}$, we see that

$$\partial_\sigma \Delta\widetilde{W}^{(1)} = \log\left(\frac{\Lambda_{\rm UV}}{|\sigma|}\right) - i\arg(\sigma) = \log\left(\frac{\Lambda_{\rm UV}}{\sigma}\right), \tag{15.157}$$

$$\partial_\sigma\partial_{\bar\sigma}\Delta K^{(1)} = \frac{1}{4|\sigma|^2}. \tag{15.158}$$

Thus we find

$$\partial_\sigma \widetilde{W}_{\rm eff}^{(1)} = \log\left(\frac{\Lambda_{UV}}{\sigma}\right) - t_0 = \log\left(\frac{\mu}{\sigma}\right) - t(\mu), \tag{15.159}$$

$$\partial_\sigma\partial_{\bar\sigma} K_{\rm eff}^{(1)} = \frac{1}{2e^2} + \frac{1}{4|\sigma|^2}. \tag{15.160}$$

In Eq. (15.159), the dependence on the ultraviolet cut-off $\Lambda_{\rm UV}$ has cancelled against the one from the bare coupling t_0. Similarly, it is independent of the choice of the scale μ; the $\log(\mu)$ dependence is cancelled by the $\log(\mu)$ dependence of $t(\mu)$ induced by the RG flow. In terms of $\mathit{\Lambda} := \mu\, e^{-t(\mu)} = \Lambda\, e^{i\theta}$, the complexified RG invariant scale parameter, Eq. (15.159) can be written as

$$\partial_\sigma \widetilde{W}_{\rm eff}^{(1)}(\sigma) = \log(\mathit{\Lambda}/\sigma). \tag{15.161}$$

This effective superpotential captures the axial anomaly of the system. The axial rotation $\Sigma \to e^{2i\beta}\Sigma$ shifts the theta angle as $\theta \to \theta - 2\beta$ (or $\widetilde{W}_{\rm eff}^{(1)}(\Sigma)$ has the correct axial charge 2 if we let the axial R-rotation shift the Theta angle as $\theta \to \theta + 2\beta$).

We have yet to integrate out the high frequency modes of the Σ multiplet fields. This will definitely affect the Kähler potential. However, it cannot affect the twisted superpotential. The correction should involve the gauge coupling constant e but that parameter cannot enter into $\widetilde{W}_{\text{eff}}$. To elaborate on this point, we first note that the standard requirements (symmetry, holomorphy, asymptotic condition) constrain the form of the superpotential. Here we use the axial R-symmetry with $\sigma \to e^{2i\beta}\sigma$, $\Lambda \to e^{2i\beta}\Lambda$ and the condition that $\widetilde{W}_{\text{eff}}(\Sigma)$ approaches Eq. (15.161) at $\sigma/\Lambda \to \infty$. The constrained form is such that

$$\partial_\sigma \widetilde{W}_{\text{eff}}(\sigma) = \log(\Lambda/\sigma) + \sum_{n=1}^{\infty} a_n (\Lambda/\sigma)^n. \tag{15.162}$$

The correction terms take the form of non-perturbative corrections. However, in the present computation, we are simply integrating out the high frequency modes of Σ in a theory without a charged field, and there is no room for non-perturbative effects. Thus, we conclude that all $a_n = 0$. This establishes that Eq. (15.161) remains the same at lower energies.

We thus see that the effective superpotential is given by

$$\widetilde{W}_{\text{eff}}(\Sigma) = -\Sigma\left(\log\left(\frac{\Sigma}{\Lambda}\right) - 1\right). \tag{15.163}$$

We consider its first derivative as the effective FI-theta parameter that varies as a function of σ:

$$t_{\text{eff}}(\sigma) := -\partial_\sigma \widetilde{W}_{\text{eff}}(\sigma) = \log(\sigma/\Lambda). \tag{15.164}$$

Using Eq. (15.140) we find that the energy density is given by

$$U = \frac{e_{\text{eff}}^2}{2}\,|\,\hat{t}_{\text{eff}}(\sigma)\,|^2, \tag{15.165}$$

where $(1/2e_{\text{eff}}^2) = \partial_\sigma\partial_{\bar\sigma}K_{\text{eff}}$, and the hat on $\hat{t}_{\text{eff}}$ stands for the shift by $2\pi n$ that is explained above. This shift resolves the apparent problem of the superpotential shown in Eq. (15.163) not being single-valued.

It is straightforward to generalize the above result to more general cases. If there are N chiral superfields of charge 1, the effective action is simply obtained by multiplying $\Delta L(\sigma)$ by N. Thus, the effective superpotential is

$$\widetilde{W}_{\text{eff}}(\Sigma) = -\Sigma\left[N\left(\log\left(\frac{\Sigma}{\mu}\right) - 1\right) + t(\mu)\right] = -N\Sigma\left(\log\left(\frac{\Sigma}{\Lambda}\right) - 1\right), \tag{15.166}$$

where $\Lambda := \mu\, e^{-t(\mu)/N}$ is the complexified RG invariant dynamical scale. For the most general case where the gauge group is $U(1)^k = \prod_{a=1}^k U(1)_a$ with the chiral matter fields Φ_i of charge Q_{ia}, the effective superpotential is

$$\widetilde{W}_{\text{eff}}(\Sigma_1, \ldots, \Sigma_k) = -\sum_{a=1}^k \Sigma_a\left[\sum_{i=1}^N Q_{ia}\left(\log\left(\frac{\sum_{b=1}^k Q_{ib}\Sigma_b}{\mu}\right) - 1\right) + t_a(\mu)\right]. \tag{15.167}$$

This is derived exactly by using one-loop computations in the case where there is no superpotential term for the Φ_i's. However, even if there is such an F-term, by the decoupling theorem of F-terms and twisted F-terms, the result Eq. (15.167) will not be affected.

15.5.2. The $\mathbb{CP}^{N-1}$ Model. Now let us study the low energy behavior of the $\mathbb{CP}^{N-1}$ model. As we have seen, this is realized by the $U(1)$ gauge theory with N chiral superfields of charge 1. The axial R-symmetry $U(1)_A$ is anomalously broken to $\mathbb{Z}_{2N}$ and the theory dynamically generates the scale parameter Λ. We look at the effective theory at energy $\mu \ll \Lambda$. The region in the field space where σ is slowly varying compared to $1/\mu$ and much larger than μ is described by the theory of the Σ multiplet determined above. Namely, the effective twisted superpotential is given by Eq. (15.166) with the effective FI-theta parameter

$$t_{\text{eff}}(\sigma) := -\partial_\sigma \widetilde{W}_{\text{eff}}(\sigma) = N\log(\sigma/\Lambda). \tag{15.168}$$

The supersymmetric ground states are found by looking for the value of σ that satisfy $U = (e_{\text{eff}}^2/2)|\hat{t}_{\text{eff}}(\sigma)|^2 = 0$. Namely, we look for solutions to $t_{\text{eff}}(\sigma) \in 2\pi i\mathbb{Z}$ or equivalently

$$e^{t_{\text{eff}}(\sigma)} = 1. \tag{15.169}$$

This is solved by

$$\sigma = \Lambda \cdot e^{2\pi i n/N},\ n = 0, \ldots, N-1. \tag{15.170}$$

Since the scale μ is taken to be much smaller than Λ, these vacua are in the region where the analysis is valid. Thus, we find N supersymmetric vacua in this region. The $\mathbb{Z}_{2N}$ axial R-symmetry cyclically permutes these N vacua. Namely, a choice of a vacuum spontaneously breaks the axial R-symmetry to $\mathbb{Z}_2$:

$$\mathbb{Z}_{2N} \rightsquigarrow \mathbb{Z}_2. \tag{15.171}$$

From this analysis alone, however, we cannot exclude the possibility of other vacua in the region with small σ. To describe the physics in such a region, we need to use a completely different set of variables. If we use the full variables Φ_i's and Σ, we need to find a minimum where the potential U in Eq. (15.62) vanishes. However, if $\mu \ll \Lambda$, $r(\mu)$ is large and negative and $U = 0$ cannot be attained by any configuration. This is one indication that there is no other vacuum state. Also, the above number, N, saturates the number of supersymmetric vacua

$$\dim H^*(\mathbb{CP}^{N-1}) = N \tag{15.172}$$

found from the direct analysis of the non-linear sigma model (done in Sec. 13.3). This also indicates that there is no other vacuum. However, to find the decisive answer we need more information, which will be provided when we will prove the mirror symmetry of the $\mathbb{CP}^{N-1}$ model and the LG model of affine Toda superpotential. The determination of the supersymmetric vacua of the latter model is straightforward and it tells us that the above N vacua are indeed the whole set.

The Dynamics at Large N. We have seen that σ has nonzero expectation values at these N vacua. This shows that the matter fields Φ_i, which include massless modes (the Goldstone modes for $SU(N)/\mathbb{Z}_N \rightsquigarrow U(N-1)/\mathbb{Z}_N$) classically, acquire a mass

$$m_\Phi \simeq \Lambda, \tag{15.173}$$

at the quantum level. Since there are no Goldstone bosons, the global symmetry $SU(N)/\mathbb{Z}_N$ cannot be broken.

Let us try to analyze the gauge dynamics of these massive charged fields. For this we need to know also the gauge kinetic terms, not only the superpotential terms. From Eq. (15.160) we see that the effective gauge coupling constant at the one-loop level is given by

$$\frac{1}{e_{\rm eff}^{(1)2}} = \frac{1}{e^2} + \frac{N}{2|\sigma|^2}. \tag{15.174}$$

As we noted above, this is further corrected by Σ-integrals and we do not know the actual form of the effective gauge coupling constant. However, there is a limit in which one can actually use Eq. (15.174) to analyze the dynamics. It is the large N limit. Since there are N matter fields of the same charge, the matter integral simply yields N times $\Delta L^{(1)}(\Sigma)$. Thus, any correction to Eq. (15.174) is suppressed by powers of $1/N$. Also, the gauge coupling near the vacua is of order $\Lambda/\sqrt{N}$ and can be made as small

as one wishes, no matter how large is the bare gauge coupling e (this is particularly useful for our purpose — the $e \to \infty$ limit). In particular, in this limit, the mass of the charged matter fields is very large compared to the gauge coupling constant,

$$m_\Phi / e_{\rm eff} \sim \sqrt{N} \gg 1. \tag{15.175}$$

Thus, we can treat the charged matter fields semi-classically.

Suppose the Φ_i or $\overline{\Phi}_i$ particles are located at $x^1 = x^1, \ldots, x^1_s$. Then the equation of motion for the gauge field is given by

$$\frac{\partial}{\partial x^1}\left(\frac{v_{01}}{e_{\rm eff}^2} + \theta_{\rm eff}\right) = 2\pi \sum_{i=1}^{s} \epsilon_i \delta(x^1 - x^1_i), \tag{15.176}$$

where $\theta_{\rm eff}$ is the effective theta angle

$$\theta_{\rm eff} = \mathrm{Im}\,(t_{\rm eff}(\sigma)) = N \arg(\sigma/\Lambda), \tag{15.177}$$

and $\epsilon_i = \pm 1$ is the charge of the particle at $x^1 = x^1_i$. Thus, $v_{01}/e^2_{\rm eff} + \theta_{\rm eff}$ has a gap of $\pm 2\pi$ at the location of the particles. At any of the N vacua we have $v_{01} = e^2_{\rm eff}|\hat{\theta}_{\rm eff}|^2 = 0$, which means $\theta_{\rm eff} = 2\pi n$ for some $n \in \mathbb{Z}$. Thus, in order to have a finite energy configuration, we need

$$\left.\begin{array}{c} v_{01} \to 0 \\ \theta_{\rm eff} \to 2\pi n_\pm \end{array}\right\} \text{ at } x^1 \to \pm\infty, \tag{15.178}$$

where $n_\pm$ are some integers. For an arbitrary distribution of particles, we can find a solution to Eq. (15.176) obeying this condition. In particular, a Φ_i particle (or a $\overline{\Phi}_i$ particle) can exist by itself. In the presence of a Φ_i particle, the vacuum at the left infinity $x^1 \to -\infty$ is not the same as the vacuum at the right infinity $x^1 \to +\infty$. This is because

$$\begin{aligned} \theta_{\rm eff}\Big|_{x^1=+\infty} - \theta_{\rm eff}\Big|_{x^1=-\infty} \\ = \int_{-\infty}^{+\infty} \frac{\partial}{\partial x^1}\left(\frac{v_{01}}{e^2_{\rm eff}} + \theta_{\rm eff}\right) dx^1 = \int 2\pi\delta(x^1 - x^1_0) = 2\pi, \end{aligned} \tag{15.179}$$

where we have used $v_{01} \to 0$ at $x^1 \to \pm\infty$. If the left infinity is at $\sigma = \Lambda$, then the right infinity is at $\sigma = \Lambda\, e^{2\pi i/N}$. A configuration connecting different vacua is called a *soliton*. We have shown that Φ_i is a soliton. We will see later that this soliton preserves a part of the supersymmetry and that its mass can be computed exactly.

If one Φ_i particle and one $\overline{\Phi}_i$ particle are located at $x^1 = -L/2$ and $x^1 = L/2$ respectively, Eq. (15.176) can be solved by a configuration

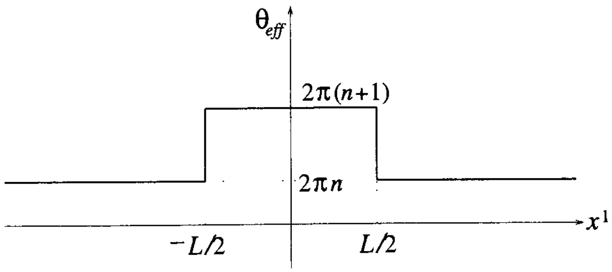


FIGURE 4. The configuration of $\theta_{\text{eff}} = N \arg(\sigma/\Lambda)$ for a pair of particles, charge 1 at $x^1 = -L/2$ and charge -1 at $x^1 = L/2$

as shown in Fig. 4. The configuration is at the vacuum in the region $-L/2 < x^1 < L/2$ and the total energy does not grow linearly as a function of the separation L. Thus, there is no long range force between them. Namely, *charged particles are not confined in this theory*. This is essentially the effect of the coupling

$$N \arg(\sigma/\Lambda) v_{01}. \tag{15.180}$$

This coupling screens the long range interaction between the charged particles.

Thus, the Φ_i particle exists as a particle state in the quantum Hilbert space. From the classical story, we expect that these states constitute the fundamental representation of the group $SU(N)$. Note that $SU(N)$ is not quite the same as the classical global symmetry group $SU(N)/\mathbb{Z}_N$. The symmetry group of the quantum theory is not $SU(N)/\mathbb{Z}_N$ but its universal covering group. Such a phenomenon is common in quantum field theories (known as *charge fractionalization*). In the present case this happens because there appeared a state transforming nontrivially under the "overlap" $\mathbb{Z}_N$ of $SU(N)$ and the gauge group, $U(1)$. Whether such a thing happens or not depends on the gauge dynamics. If the Φ_i particles were confined (as in the case without $\arg(\sigma)$-v_μ coupling), there would not be a state charged under the $U(1)$ gauge group, and therefore all the states would be neutral under $\mathbb{Z}_N = SU(N) \cap U(1)$. In that case, the global symmetry group would remain as $SU(N)/\mathbb{Z}_N$.

Other Toric Sigma Models. It is straightforward to generalize this analysis to the linear sigma model of the $U(1)^k$ gauge group and N matter fields of charge Q_{ia} that corresponds to the non-linear sigma model on a general toric manifold X. To find the supersymmetric vacua, the equations to solve are (in the case where g.c.d. of $(Q_{ia})_i$ is 1 for all a)

$$\exp\left(i t_{a,\text{eff}}(\sigma_\bullet)\right) = 1, \ a = 1, \dots, k, \tag{15.181}$$

where $t_{a,\text{eff}}(\sigma_\bullet) = -\partial_{\sigma_a} \widetilde{W}_{\text{eff}}(\sigma_\bullet)$ for Eq. (15.167). This equation reads as

$$\prod_{i=1}^{N} \left(\frac{1}{\mu} \sum_{b=1}^{k} Q_{ib}\sigma_b \right)^{Q_{ia}} = e^{-t_a(\mu)}, \ a = 1, \dots, k. \tag{15.182}$$

One may be able to find the solution case-by-case, but in general it is a nontrivial task even to find the number of solutions. More importantly, it is not clear from this analysis itself whether or not there are other supersymmetric vacua. Again, one can use the mirror symmetry which will be proved later to show that there are no other solutions. Also, one can actually compute the number of supersymmetric ground states using the mirror description. These turn out to be exactly the Euler number $\chi(X)$, which is known to be the same as the dimension of the cohomology group $H^*(X)$.

15.5.3. The "Phases".

Let us consider a $U(1)$ gauge theory with several chiral superfields $\Phi_1, \dots, \Phi_M$ with charges $Q_1, \dots, Q_M$ that sum to zero:

$$\sum_{i=1}^{M} Q_i = 0. \tag{15.183}$$

In this case, the axial R-symmetry $U(1)_A$ is an exact symmetry of the quantum theory, and the FI parameter does not run along the RG flow. We have in mind two classes of theories: one is the linear sigma model for compact Calabi–Yau hypersurfaces in $\mathbb{CP}^{N-1}$ or weighted projective spaces; the other is the theory without F-terms, which yields the non-linear sigma model on non-compact Calabi–Yau manifolds.

Since the FI parameter does not run, one can choose r to have whatever value one wants. As we have seen in the previous discussion, the theory at $r \gg 0$ and the theory at $r \ll 0$ have completely different interpretations, and also at $r = 0$ the theory becomes singular due to a development of a new branch of vacuum manifold where σ is unconstrained. Thus, it appears that the parameter space is completely separated by a singular point $r = 0$ into two regions with different physics.

This picture is considerably modified when the theta angle θ is taken into account. The actual parameter of the theory (in addition to the real and chiral parameters that enter into D-terms and F-terms) is $t = r - i\theta$ and the parameter space is a complex torus or a cylinder. It may appear that the parameter space is still separated into two regions by the circle at $r = 0$. However, this turns out not to be the case when we think about the origin of the singularity at $r = 0$. The singularity is expected when there is a new branch of vacua where new massless degrees of freedom appear. In the classical analysis at $r = 0$, this is identified as the Σ multiplet since there is a non-compact flat direction where σ is free. However, at large σ, as we have determined, the actual energy density also receives a contribution from the electric field or theta angle as in Eq. (15.140). Taking into account the more refined quantum correction, the energy density at large σ is

$$U = \frac{e_{\text{eff}}^2}{2}\left(r_{\text{eff}}^2 + \hat{\theta}_{\text{eff}}^2\right) = \frac{e_{\text{eff}}^2}{2}|\hat{t}_{\text{eff}}|^2, \tag{15.184}$$

where

$$t_{\text{eff}} = -\partial_\sigma \widetilde{W}_{\text{eff}}(\sigma) = t + \sum_{i=1}^{M} Q_i \log Q_i. \tag{15.185}$$

Here we have used the formula from Eq. (15.167) for $\widetilde{W}_{\text{eff}}$, where the Σ/μ dependence disappears because of Eq. (15.183). Thus, the energy at large σ vanishes at $r = -\sum_{i=1}^{M} Q_i \log Q_i$ and at *a single value of* θ which is 0 or π (mod 2π) depending on the Q_i's. Thus, except at a single point in the cylinder, there is no flat direction of σ. This means that the singularity is expected only at the single point. This yields a significant change in our picture; the two regions, $r \gg 0$ and $r \ll 0$, are no longer separated by a singularity, but are smoothly connected along a path avoiding the singular point. These two regions can be considered as a sort of analytic continuation of each other.

This change of picture has two important applications. One is the correspondence between Calabi–Yau sigma models and Landau–Ginzburg orbifold models. The other is the relation between sigma models on topologically distinct manifolds. We now describe them here.

Calabi–Yau/Landau–Ginzburg Correspondence. Let us consider the $U(1)$ gauge theory with chiral superfields $\Phi_1, \ldots, \Phi_N, P$ of charges $1, \ldots, 1, -N$, with the Lagrangian shown in Eq. (15.115) where $G(\Phi_1, \ldots, \Phi_N)$ is a generic degree N polynomial. As we have seen, the theory at $r \gg 0$ is identified as the non-linear sigma model on the Calabi–Yau hypersurface $G=0$ of $\mathbb{CP}^{N-1}$, whereas the theory at $r \ll 0$ is identified as the LG orbifold with group $\mathbb{Z}_N$ and the superpotential $W = \langle p \rangle G(\Phi_1, \ldots, \Phi_N)$. Thus, the Calabi–Yau sigma model and the LG orbifold are smoothly connected to each other. In other words, the LG orbifold is in the moduli space of the Calabi–Yau sigma model, or the Calabi–Yau sigma model is in the moduli space of the LG orbifold. The two are interpretations of different regions of a single family of theories.

A similar correspondence holds even when we take $G = 0$. In such a case the theory at $r \gg 0$ describes the sigma model on the total space of $\mathcal{O}(-N)$ over $\mathbb{CP}^{N-1}$, which is a non-compact Calabi–Yau manifold. On the other hand, as we have seen earlier, the theory at $r \ll 0$ is the sigma model on the orbifold $\mathbb{C}^N/\mathbb{Z}_N$. In the "orbifold phase" the parameter r has no geometric meaning. Thus, the sigma model on the total space of $\mathcal{O}(-N)$ over $\mathbb{CP}^{N-1}$ and the one on the orbifold $\mathbb{C}^N/\mathbb{Z}_N$ are in the same moduli space of theories.

Topology Change. Consider the $U(1)$ gauge theory with chiral superfields $\Phi_1, \ldots, \Phi_N$ of charge $Q_1, \ldots, Q_l > 0 > Q_{l+1}, \ldots, Q_N$ (obeying $\sum_{i=1}^N Q_i = 0$) without F-term. As we have analyzed in the examples of Sec. 15.4.2, the theory at $r \gg 0$ is identified as the sigma model on a non-compact CY manifold which is the total space of a certain vector bundle of a weighted projective space, whereas the theory at $r \ll 0$ is identified as the sigma model on the total space of another vector bundle on another weighted projective space that is generically different from the one at $r \gg 0$. Thus, the two sigma models whose target spaces are (generically) different are smoothly connected to each other.

15.5.4. Landau–Ginzburg Orbifold as an IR fixed Point. As another example, let us consider the non-linear sigma model on a hypersurface of $\mathbb{CP}^{N-1}$ of degree d less than N, so that the sigma model is asymptotically free. As we have seen, this theory is realized in the linear sigma model as a $U(1)$ gauge group with chiral superfields $\Phi_1, \ldots, \Phi_N, P$ of charge $1, \ldots, 1, -d$. The Lagrangian of the theory is given by Eq. (15.115) where $G(\Phi_1, \ldots, \Phi_N)$ is a generic degree d polynomial. The axial R-symmetry

$U(1)_A$ is anomalously broken to $\mathbb{Z}_{2(N-d)}$ and the theory dynamically generates the scale parameter Λ.

We have analyzed the effective theory at energy $\mu \ll \Lambda$ at large and slowly varying Σ. It is the theory of a $U(1)$ gauge multiplet with the effective FI-theta parameter given by

$$t_{\text{eff}}(\sigma) = (N-d)\log(\sigma/\Lambda) - d\log(-d). \tag{15.186}$$

The supersymmetric vacua of this theory are determined by finding solutions to $e^{it_{\text{eff}}(\sigma)} = 1$ or

$$\sigma^{N-d} = (-d)^d \Lambda^{N-d}, \tag{15.187}$$

and we find $(N-d)$ of them in the admissible region. These are massive, and a choice of vacuum spontaneously breaks the axial R-symmetry as $\mathbb{Z}_{2(N-d)} \rightsquigarrow \mathbb{Z}_2$.

Now let us ask whether these $(N-d)$ are the whole set of vacua. There is an obvious reason to doubt it; direct analysis of the non-linear sigma model shows that the number of vacua is equal to the dimension of the cohomology group $H^*(M)$, which is larger than $(N-d)$. (It is not smaller than $N-1$ since the powers of $[H]|_M$ are non-trivial.) How can we find the rest? They must be in the region where the large σ analysis does not apply. Let us examine the potential from Eq. (15.116) in terms of the full set of variables once again, now at low energies. At $\mu \ll \Lambda$ the FI parameter is negative, and the analysis of supersymmetric vacua $U = 0$ is completely different from that at high energies. It is more like in the $d = N$ case with $r < 0$ and we find a single supersymmetric vacuum at $\sigma = 0$, $\phi_i = 0$ and $|p| = \sqrt{|r|/d}$ where the axial R-symmetry group $\mathbb{Z}_{2(N-d)}$ is not spontaneously broken. Thus, we find at least one extra supersymmetric vacuum besides those found at $\sigma \sim \Lambda$. The theory around this vacuum is described by the LG orbifold of the fields $\Phi_1, \ldots, \Phi_N$ with the group $\mathbb{Z}_d$ and the superpotential $W \sim G(\Phi_1, \ldots, \Phi_N)$. For $d > 2$ this LG orbifold is expected to flow to a non-trivial superconformal field theory where the axial $\mathbb{Z}_{2(N-d)}$ discrete R-symmetry enhances to the full $U(1)$ symmetry (or actually further to affine symmetry). One can actually analyze the spectrum of the supersymmetric vacua of this LG orbifold, which in fact shows that the number of vacua is $\dim H^*(M) - (N-d)$, and the total number saturates the one derived from the direct analysis. Thus, we expect that this extra (degenerate) vacuum really exists in the quantum theory and is the only one that was missed by the large σ analysis. Of course, to be decisive we need more information. Again, we will see that mirror symmetry (which we will give an argument for) shows that this is in fact correct.

15.5.5. A Flow from Landau–Ginzburg Orbifold. As a final example, let us consider the case $d > N$ of the $U(1)$ gauge theory considered directly above. As we have seen, the FI parameter at the cut-off scale is negative and the theory at high energy describes the LG orbifold perturbed by an operator that breaks the $U(1)$ axial R-symmetry to $\mathbb{Z}_{2(d-N)}$.

The large σ analysis shows that there are $(d-N)$ vacua determined by Eq. (15.187), each of which breaks $\mathbb{Z}_{2(d-N)}$ to $\mathbb{Z}_2$. We may also find supersymmetric vacua near $\sigma = 0$. In fact, the FI parameter becomes positive at low energies and we find the degree d hypersurface M in $\mathbb{CP}^{N-1}$ as the vacuum manifold at $\sigma = 0$. The non-linear sigma model on M is IR free and we expect this to be one of the IR fixed points of the theory.

CHAPTER 16

Chiral Rings and Topological Field Theory

In this chapter, we study the chiral rings of $(2,2)$ supersymmetric field theories. The chiral ring is a basic property of the $(2,2)$ theories, somewhat like the Witten index. Just like the Witten index, it is independent of infinitely many supersymmetric deformations of the theory. However, unlike the Witten index it *does depend on a finite number of deformations captured by holomorphic (or anti-holomorphic) parameters capturing the F-terms.* The Witten index turns out to be related to the number of basis elements in the ring. The ring structure itself requires more data that depend on the choice of the F-terms. A closely related idea is to consider a slight change of the conventional $(2,2)$ theories to obtain what are called "topological field theories." Topological theories coincide with ordinary $(2,2)$ theories on flat worldsheets, but differ from them on curved Riemann surfaces (known as "topological twisting") in a way that leads to preserving half of the supersymmetries. The chiral ring is captured by the correlation functions of the observables of the topological field theory on a sphere. We study several classes of twisted theories in detail and carry out the computation of the chiral ring in some examples.

16.1. Chiral Rings

Let Q be either

$$Q_A = \overline{Q}_+ + Q_- \text{ or } Q_B = \overline{Q}_+ + \overline{Q}_-. \tag{16.1}$$

As we have seen before, if we assume that the central charges vanish, $Z = \widetilde{Z} = 0$, then Q is nilpotent:

$$Q^2 = 0. \tag{16.2}$$

We have used this fact to consider the Q-cohomology of states, which is isomorphic to the space of supersymmetric ground states. Instead, here we consider the Q-cohomology of operators. This will lead us to the notion of chiral rings.

An operator $\mathcal{O}$ is called a *chiral operator* if

$$[Q_B, \mathcal{O}] = 0, \tag{16.3}$$

and a *twisted chiral operator* if

$$[Q_A, \mathcal{O}] = 0. \tag{16.4}$$

As we have seen before in Eq. (12.81), the lowest component ϕ of a chiral superfield Φ obeys

$$[\overline{Q}_\pm, \phi] = 0, \tag{16.5}$$

and is therefore a chiral operator. Similarly, the lowest component v of a twisted chiral superfield U obeys Eq. (12.83)

$$[\overline{Q}_+, v] = [Q_-, v] = 0, \tag{16.6}$$

and is a twisted chiral operator. It follows from the supersymmetry algebra that if $\mathcal{O}$ is a chiral (twisted chiral) operator, then its worldsheet derivative is Q_B-exact (Q_A exact). For instance, if $\mathcal{O}$ is a chiral operator, then

(16.7)

$$\begin{aligned}\frac{i}{2}\left(\frac{\partial}{\partial x^0}+\frac{\partial}{\partial x^1}\right)\mathcal{O} =&[(H+P), \mathcal{O}] = [\{Q_+, \overline{Q}_+\}, \mathcal{O}]\\ =&\{[Q_+, \mathcal{O}], \overline{Q}_+\} + \{Q_+, [\overline{Q}_+, \mathcal{O}]\}\\ =&\{[Q_+, \mathcal{O}], \overline{Q}_+\} - \{Q_+, [\overline{Q}_-, \mathcal{O}]\}\\ =&\{[Q_+, \mathcal{O}], \overline{Q}_+\} - [\{Q_+, \overline{Q}_-\}, \mathcal{O}] + \{\overline{Q}_-, [Q_+, \mathcal{O}]\}\\ =&\{Q_B, [Q_+, \mathcal{O}]\}.\end{aligned}$$

Similarly,

$$\frac{i}{2}\left(\frac{\partial}{\partial x^0}-\frac{\partial}{\partial x^1}\right)\mathcal{O} = \{Q_B, [Q_-, \mathcal{O}]\}. \tag{16.8}$$

Thus, the worldsheet translation does not change the Q_B (Q_A) cohomology class of a (twisted) chiral operator.

If $\mathcal{O}_1$ and $\mathcal{O}_2$ are two (twisted) chiral operators, then the product $\mathcal{O}_1\mathcal{O}_2$ is also a (twisted) chiral operator. Thus the Q-cohomology classes of operators form a ring, called the *chiral ring* for $Q = Q_B$ and the *twisted chiral ring* for $Q = Q_A$.

Let $\{\phi_i\}_{i=0}^M$ be a basis of the Q-cohomology group of operators. Since it is a ring, the product of two elements is expanded as

$$\phi_i\phi_j = \phi_k C_{ij}^k + [Q, \Lambda]. \tag{16.9}$$

The coefficients C_{ij}^k are the structure constants of the ring with respect to the basis $\{\phi_i\}_{i=0}^M$.

It is symmetric or antisymmetric in $i \leftrightarrow j$ depending on whether ϕ_i and ϕ_j are both fermionic or not:

$$C_{ji}^k = (-1)^{|i||j|} C_{ij}^k. \tag{16.10}$$

Since the operator product is associative, $\phi_i(\phi_j\phi_k) = (\phi_i\phi_j)\phi_k$, the structure constants obey

$$C_{il}^m C_{jk}^l = C_{lk}^m C_{ij}^l. \tag{16.11}$$

Also, the identity operator $\mathcal{O} = 1$ represents a Q-cohomology class whose product with other elements is trivial. We choose $\phi_0 = 1$ so that

$$C_{0j}^k = C_{j0}^k = \delta_j^k. \tag{16.12}$$

16.2. Twisting

So far in this chapter we have assumed that the two-dimensional worldsheet is a flat manifold — the flat Minkowski or Euclidean plane $\mathbb{R}^2$, a flat cylinder $\mathbb{R} \times S^1$, or a flat torus T^2. However, there are many reasons for formulating the theories on a curved Riemann surface. One motivation is to find the correspondence between operators and states. This is usually done using a worldsheet having the geometry of a hemisphere.

Another motivation comes from string theory. A string amplitude is given by the sum over all topologies and conformal classes of Riemann surfaces, and a starting point is to formulate the amplitude for a fixed genus surface of arbitrary geometry. There is no obstruction to formulating a supersymmetric theory on a curved Riemann surface Σ. We must, however, take care to choose a spin structure so that one can put spinors on the surface. Once this is done, the action is well defined. However the action is not necessarily supersymmetric. Consider the supersymmetry variation of the action, which would be given by

$$\delta S = \int_\Sigma \left(\nabla_\mu\epsilon_+ G_-^\mu - \nabla_\mu\epsilon_- G_+^\mu - \nabla_\mu\overline{\epsilon}_+\overline{G}_-^\mu + \nabla_\mu\overline{\epsilon}_-\overline{G}_+^\mu\right)\sqrt{h}d^2x. \tag{16.13}$$

Here $\epsilon_\pm$ and $\overline{\epsilon}_\pm$ are the variational parameters that are spinors on Σ. In the case where Σ is flat and ϵ_+ et al. can be chosen to be constant, the above equation means that the action is invariant under the supersymmetry transformations. On a general Riemann surface, the variation δS vanishes for covariantly constant parameters $\nabla_\mu \epsilon_\pm = \nabla_\mu \bar{\epsilon}_\pm = 0$. However, if Σ is a curved Riemann surface, there is no covariantly constant spinor on Σ. Thus, supersymmetry is lost on a curved surface. However, it would be interesting to find a theory with a fermionic symmetry where one can make full use of the localization principle and deformation invariance. This motivates the *twisting* of the theory. It modifies the theory so that there is a conserved fermionic symmetry even on a curved Riemann surface. This twisting agrees with the original theory on regions of the surface where the metric is flat. In particular the Hilbert space of the physical theory can also be used for the topological theory, because the two theories are the same on a flat cylinder (however, the interesting states, from the perspective of the topological theory turn out to be the ground states of the physical theory, as we will discuss later).

16.2.1. The Definition. We start with the Euclidean version of the theory obtained from the Minkowski theory by Wick rotation $x^0 = -ix^2$. (We choose the orientation of the worldsheet so that $z = x^1 + ix^2$ is the complex coordinate.) The theory still has supersymmetry with the algebra in Eqs. (12.70)–(12.76) as before, and the same Hermiticity condition. In particular, the R-symmetry generators (if both are conserved) act on the supercharges as

$$[F_V, Q_\pm] = -Q_\pm, \ [F_V, \overline{Q}_\pm] = \overline{Q}_\pm,$$
$$[F_A, Q_\pm] = \mp Q_\pm, \ [F_A, \overline{Q}_\pm] = \pm \overline{Q}_\pm.$$

The $SO(1,1)$ Lorentz group is now the Euclidean rotation group $SO(2)_E = U(1)_E$ with the (anti-Hermitian) generator

$$M_E = iM \tag{16.14}$$

that acts on the supercharges as

$$[M_E, Q_\pm] = \mp Q_\pm, \ [M_E, \overline{Q}_\pm] = \mp \overline{Q}_\pm. \tag{16.15}$$

We consider a theory that possesses either one of $U(1)_V$ or $U(1)_A$ R-symmetries under which the R-charges are all integral. *Twisting* is done with respect to such an R-symmetry, which we call $U(1)_R$ and denote the generator by R. Its effect is to replace the group $U(1)_E$ by the diagonal subgroup $U(1)'_E$ of $U(1)_E \times U(1)_R$ with the generator

$$M'_E = M_E + R, \tag{16.16}$$

i.e., if we consider $U(1)'_E$ as the new rotation group. We call this the *A-twist* or *B-twist* depending on which R-symmetry we use;

$$\text{A-twist}: \ R = F_V, \ U(1)_R = U(1)_V,$$
$$\text{B-twist}: \ R = F_A, \ U(1)_R = U(1)_A.$$

The twisted theory on a curved worldsheet is obtained by gauging the new rotation group $U(1)'_E$ (instead of the one $U(1)_E$) by the spin connection. What this means is that the fields will now be sections of different bundles over the Riemann surface, i.e., the "spins" of the fields are modified. For instance, consider a chiral superfield of trivial R-charges $q_V = q_A = 0$

$$\Phi = \phi + \theta^+ \psi_+ + \theta^- \psi_- + \cdots. \tag{16.17}$$

For the lowest component ϕ, the M_E charge, vector R-charge and axial R-charge are all zero. Therefore ϕ has M'_E charge zero and remains as the scalar field after twisting. For ψ_- it has M_E charge 1 and R-charge $q_V = -1$, $q_A = 1$. That the M_E charge of ψ_- is 1 means that ψ_- is a spinor field, or a section of the spinor bundle $\sqrt{K}$ (K is the canonical bundle on Σ) before twisting. After A-twist, it has M'_E charge $1 + q_V = 0$ and it becomes a scalar field. After B-twist, it has M'_E charge $1 + q_A = 2$ and it becomes a vector or a one-form field which is a section of K. This result and the result for other component fields are summarized in the following table.

	Before twisting				A-twist		B-twist	
	$U(1)_V$	$U(1)_A$	$U(1)_E$	$\mathcal{L}$	$U(1)'_E$	$\mathcal{L}$	$U(1)'_E$	$\mathcal{L}$
ϕ	0	0	0	$\mathbb{C}$	0	$\mathbb{C}$	0	$\mathbb{C}$
ψ_-	-1	1	1	$\sqrt{K}$	0	$\mathbb{C}$	2	K
$\overline{\psi}_+$	1	1	-1	$\sqrt{\overline{K}}$	0	$\mathbb{C}$	0	$\mathbb{C}$
$\overline{\psi}_-$	1	-1	1	$\sqrt{K}$	2	K	0	$\mathbb{C}$
ψ_+	-1	-1	-1	$\sqrt{\overline{K}}$	-2	$\overline{K}$	-2	$\overline{K}$

In this table, $\mathcal{L}$ is the complex line bundle on Σ in which the field takes values. ($\mathbb{C}$ is the trivial bundle, $\overline{K}$ is the complex conjugate (or the dual) of the canonical bundle K.)

16.2.2. Some Consequences. There are several important consequences of twisting that can be derived immediately.

First of all, twisting has no effect on a Riemann surface with a flat metric. Everything is the same before and after the twisting, since $K = \overline{K} = \mathbb{C}$. However, as we vary the metric away from the flat metric, the topological theory differs from the original theory. In particular, even for the flat metric, the energy-momentum tensor, which by definition is the variation of the action with respect to change of metric, will be different for the twisted and the untwisted theories. More specifically, the energy-momentum tensor $T_{\mu\nu}$ which is defined classically by $\delta S = -\frac{1}{4\pi}\int \sqrt{h} d^2x\, \delta h^{\mu\nu} T_{\mu\nu}$, and quantum mechanically by

$$\delta_h \langle \mathcal{O} \rangle = \left\langle \mathcal{O} \frac{1}{4\pi} \int \sqrt{h} d^2x\, \delta h^{\mu\nu} T_{\mu\nu} \right\rangle, \tag{16.18}$$

is modified in the twisted theory in the following way:

$$T^{\text{twisted}}_{\mu\nu} = T_{\mu\nu} + \tfrac{1}{4}(\epsilon^{\lambda}_{\mu} \partial_\lambda J^R_\nu + \epsilon^{\lambda}_{\nu} \partial_\lambda J^R_\mu). \tag{16.19}$$

Here J^R_μ is the $U(1)_R$ current that is defined with respect to the variation of the R-symmetry gauge field A^R by

$$\delta_{A^R} \langle \mathcal{O} \rangle = \left\langle \mathcal{O} \frac{1}{2\pi i} \int *J^R \wedge \delta A^R \right\rangle. \tag{16.20}$$

Twisting affects the spin of the supercharges. It is easy to see from the commutation relations of M_E, F_V, F_A and the supercharges that the changes are as in the table.

	Before twisting				A-twist		B-twist	
	$U(1)_V$	$U(1)_A$	$U(1)_E$	$\mathcal{L}$	$U(1)'_E$	$\mathcal{L}$	$U(1)'_E$	$\mathcal{L}$
Q_-	-1	1	1	$\sqrt{K}$	0	$\mathbb{C}$	2	K
$\overline{Q}_+$	1	1	-1	$\sqrt{\overline{K}}$	0	$\mathbb{C}$	0	$\mathbb{C}$
$\overline{Q}_-$	1	-1	1	$\sqrt{K}$	2	K	0	$\mathbb{C}$
Q_+	-1	-1	-1	$\sqrt{\overline{K}}$	-2	$\overline{K}$	-2	$\overline{K}$

Here $\mathcal{L}$ is the line bundle in which the supercurrent G takes values, $G \in \Omega^1(\Sigma, \mathcal{L})$. Thus in the A-twisted theory $\overline{Q}_+$ and Q_- have spin zero, whereas in the B-twisted theory $\overline{Q}_+$ and $\overline{Q}_-$ are the spin-zero charges. Since these supercharges are now scalars, they make sense without reference to the choice

of worldsheet coordinates. We notice that the combination $Q_A = \overline{Q}_+ + Q_-$ (or $Q_B = \overline{Q}_+ + \overline{Q}_-$), which we used to define the twisted chiral ring (or chiral ring), is a scalar in the twisted theory.

16.2.3. Physical Observables of the Topological Theories. The Hilbert spaces of a topologically-twisted theory and the ordinary theory do not differ. The same is true of the operators in the two theories. However, what does change in considering topological theories is the set of operators and states in the Hilbert space which we consider "physical". In particular, in the topologically twisted theory we define physical operators to be operators that commute with $Q = Q_A$ or Q_B. Moreover, the physical states are labelled by Q-cohomology elements, which are in turn in one-to-one correspondence with the ground states of the supersymmetric theory.

EXERCISE 16.2.1. *Just as in the case of our discussion of zero-dimensional LG QFT, show that this notion of physical is compatible with the path-integral definition of the topological theory.*

Since the physical operators of the topological theory are defined to be the Q-cohomology group elements, the physical operators of A-twisted (B-twisted) theory are the twisted chiral ring elements (chiral ring elements).

We also note that the other supercharges become one-form operators. A consequence is that the relations Eqs. (16.7)–(16.8) and a generalization can be written in a covariant form. If $\mathcal{O}^{(0)} = \mathcal{O}$ is a Q-closed operator, then one can find a one-form operator $\mathcal{O}^{(1)}$ and a two-form operator $\mathcal{O}^{(2)}$ such that

$$0 = [Q, \mathcal{O}^{(0)}], \tag{16.21}$$
$$d\mathcal{O}^{(0)} = \{Q, \mathcal{O}^{(1)}\}, \tag{16.22}$$
$$d\mathcal{O}^{(1)} = [Q, \mathcal{O}^{(2)}], \tag{16.23}$$
$$d\mathcal{O}^{(2)} = 0. \tag{16.24}$$

The required operators are

$$\text{B-twist}: \left\{ \begin{array}{l} \mathcal{O}^{(1)} = i dz [Q_-, \mathcal{O}] - i d\bar{z} [Q_+, \mathcal{O}], \\ \mathcal{O}^{(2)} = dz\, d\bar{z}\, \{Q_+, [Q_-, \mathcal{O}]\}, \end{array} \right\} \tag{16.25}$$

($Q_- dz$ and $Q_+ d\bar{z}$ are covariant combinations), and for A-twist they are obtained from these expressions with the replacement $Q_- \to \overline{Q}_-$. Note that the supercharges have the right spin for $\mathcal{O}^{(1)}$ and $\mathcal{O}^{(2)}$ to be a one-form and a two-form. The Eqs. (16.21)–(16.24) are called *descent relations*. The first Eq. (16.21) simply says that $\mathcal{O}^{(0)}$ is Q-closed and the last Eq. (16.24) is a triviality that there is no three-form in two dimensions. The second Eq. (16.22) is the Euclidean version of Eq. (16.7) and Eq. (16.8) (for B-twist with $Q = Q_B$) and the third Eq. (16.23) can be derived in a similar way. The construction shown in Eq. (16.25) is the topological version of the construction shown in Eq. (12.82) of the full chiral multiplet from a chiral field.

From Eqs. (16.22)–(16.23), we see that

$$\int_\gamma \mathcal{O}^{(1)} \text{ and } \int_\Sigma \mathcal{O}^{(2)} \tag{16.26}$$

are Q-invariant operators, where γ is a closed one-cycle and Σ is the worldsheet (assumed to have no boundary). We will see below that the second type of operator, namely $\int_\Sigma \mathcal{O}^{(2)}$, effects a deformation of the theory.

In many classes of $(2,2)$ theories, including those considered in this chapter, the energy-momentum tensor of the twisted theory is actually Q-exact,

$$T^{\text{twisted}}_{\mu\nu} = \{Q, G_{\mu\nu}\}, \tag{16.27}$$

where $G_{\mu\nu}$ is a certain fermionic symmetric tensor. This relation requires some explanation: The energy-momentum tensor can be viewed as a density field for the energy and momentum vector. Thus the integrated version of the above equation is already familiar, namely H and P can be written as Q anti-commutators with certain fields. The above formula is a refinement of this relation, extended to the twisted theories. This is a very important relation. In particular, if we consider the variation of the correlation functions as we change the worldsheet metric h, we have

$$\delta_h\langle \mathcal{O}_1\cdots\mathcal{O}_s\rangle = \left\langle \frac{1}{4\pi}\int \sqrt{h}d^2x\,\delta h^{\mu\nu}\{Q, G_{\mu\nu}\}\mathcal{O}_1\cdots\mathcal{O}_s\right\rangle. \tag{16.28}$$

This vanishes if all of $\mathcal{O}_1,\ldots,\mathcal{O}_s$ are physical operators (Q-closed). Thus, the correlation functions of the physical operators are independent of the choice of the worldsheet metric. In this sense the twisted theory is called a topological field theory.

16.3. Topological Correlation Functions and Chiral Rings

Here we study some general properties of topological correlation functions. To be specific, we consider the B-twisted theory where $Q = Q_B$.

As we will mention, the statements apply to the A-twisted theory with an obvious exchange of terminology.)

16.3.1. Dependence on the Parameters. As we have seen above, topological correlation functions are independent of the choice of the worldsheet metric. We now see how they depend on the parameters of the theory. We have seen that there are three classes of parameters — parameters that enter in D-terms, complex parameters that enter in F-terms (and their complex conjugates which enter in conjugate F-terms), and complex parameters that enter in twisted F-terms (and their complex conjugates which enter in the conjugate twisted F-terms). We consider them separately.

• A topological correlation function is independent of deformations of the D-terms. The variation of a D-term inserts in the path-integral an operator of the form

$$\int d^4\theta\,\Delta K = \int d\overline{\theta}^+ d\overline{\theta}^- d\theta^- d\theta^+\,\Delta K. \tag{16.29}$$

This is proportional to

$$\left\{\overline{Q}_+, \left[\overline{Q}_-, \int d\theta^+ d\theta^-\,\Delta K\Big|_{\overline{\theta}^\pm=0}\right]\right\} = \left\{Q, \left[\overline{Q}_-, \int d\theta^+ d\theta^-\,\Delta K\Big|_{\overline{\theta}^\pm=0}\right]\right\}, \tag{16.30}$$

where we have used the nilpotency of $\overline{Q}_-$. Thus, the inserted operator is Q-exact and the correlation function vanishes. Thus the variation of a D-term does not affect the topological correlation functions.

• It is independent of the twisted chiral and anti-twisted chiral parameters. To see this, we note that the twisted chiral deformation corresponds to the insertion of the operator of the form

$$\int \sqrt{h}d^2x \int d^2\widetilde{\theta}\,\Delta\widetilde{W}(\widetilde{\Phi}) \propto \int \sqrt{h}d^2x\left\{Q_+, \left[\overline{Q}_-, \Delta\widetilde{W}(\widetilde{\phi})\right]\right\}, \tag{16.31}$$

where $\Delta\widetilde{W}(\widetilde{\phi})$ is a twisted chiral operator annihilated by both $\overline{Q}_+$ and Q_-. By the fact that $\Delta\widetilde{W}(\widetilde{\phi})$ is annihilated by $\overline{Q}_+$ we have

$$\begin{aligned}\left\{Q_+, \left[Q_-, \Delta\widetilde{W}(\widetilde{\phi})\right]\right\} &= \left\{Q_+, \left[Q_- + \overline{Q}_+, \Delta\widetilde{W}(\widetilde{\phi})\right]\right\}\\ &= -\left\{Q, \left[Q_+, \Delta\widetilde{W}(\widetilde{\phi})\right]\right\} + \text{total derivative}\end{aligned} \tag{16.32}$$

where in the last step we have used the anti-commutation relation of the supercharges. Thus, the inserted operator is Q-exact and therefore annihilates any topological correlation function. So, topological correlation functions are independent of the twisted chiral parameters. In a similar way, one can show that they are independent of the anti-twisted chiral parameters.

- It is independent of anti-chiral parameters. The variation of an anti-chiral parameter corresponds to the insertion of the type of operator

$$\int \sqrt{h} d^2x \int d^2\overline{\theta}\, \Delta\overline{W}(\overline{\Phi}) \propto \int \sqrt{h} d^2x\, \{\overline{Q}_+, [\overline{Q}_-, \Delta\overline{W}(\overline{\phi})]\}\,. \tag{16.33}$$

Using the nilpotency of $\overline{Q}_-$ we have

$$\begin{aligned} \{\overline{Q}_+, [\overline{Q}_-, \Delta\overline{W}(\overline{\phi})]\}\,. &= \{\overline{Q}_+ + \overline{Q}_-, [\overline{Q}_-, \Delta\overline{W}(\overline{\phi})]\} \\ &= \{Q, [\overline{Q}_-, \Delta\overline{W}(\overline{\phi})]\}\,. \end{aligned} \tag{16.34}$$

Thus, the inserted operator is Q-exact and annihilates correlation functions.

- It can depend on the chiral parameters. The variation of the chiral parameters corresponds to the insertion of the operator

$$\begin{aligned} \int \sqrt{h} d^2x \int d^2\theta\, \Delta W(\Phi) &\propto \int \sqrt{h} d^2x\, \{Q_+, [Q_-, \Delta W(\phi)]\} \\ &\propto \int \Delta W(\phi)^{(2)}. \end{aligned} \tag{16.35}$$

This is a non-trivial Q-invariant operator that is the second descendant of the chiral operator $\Delta W(\phi)$ (see Eq. (16.25)).

To summarize, we have seen that, in the B-twisted theory, the topological correlation functions are independent of D-term variations, twisted chiral and anti-twisted chiral parameters and anti-chiral parameters. This means that they depend only on the chiral parameters, and the dependence is holomorphic. (Similarly, in the A-twisted theory, topological correlation functions depend holomorphically on twisted chiral parameters only.) Since typically we consider only a finite number of variations of the F-terms, we thus see that in these cases the topological correlation functions depend on only a finite-dimensional subspace of the infinite-dimensional parameter space of the underlying QFT.

16.3.2. Chiral Ring from Three-Point Functions. Let us consider a genus 0 three-point function, namely, a correlation function for the spherical worldsheet $\Sigma = S^2$ where three physical operators – ϕ_i, ϕ_j, ϕ_k – are inserted at three distinct points. It does not matter which metric one puts on S^2, nor where the operators are inserted. We denote this as

$$C_{ijk} = \langle \phi_i \phi_j \phi_k \rangle_0. \tag{16.36}$$

A special role is played by the three-point function with one operator being the identity operator 1. We denote this by

$$\eta_{ij} = C_{ij0}. \tag{16.37}$$

Since the insertion of the identity operator has no effect, this is really a two-point function, $\eta_{ij} = \langle \phi_i \phi_j \rangle_0$. For the classes of theories we are dealing with, the matrix η_{ij} is invertible. We consider this matrix η_{ij} as determining a metric on the parameter space and we call it the *topological metric*. We denote the inverse matrix by η^{ij}, so that

$$\eta^{ij}\eta_{jk} = \delta^i_k. \tag{16.38}$$

Now, let us consider again the general three-point function shown in Eq. (16.36). Since it does not depend on the position of the insertion points, we can consider making the insertion points for ϕ_j and ϕ_k approach each other. One can replace the product of ϕ_j and ϕ_k by

$$\phi_j \phi_k = \phi_l C^l_{jk} + [Q, \Lambda], \tag{16.39}$$

where C^l_{jk} are the chiral ring structure constants as in Sec. 16.1. Then we see that

$$\langle \phi_i \phi_j \phi_k \rangle_0 = \langle \phi_i (\phi_l C^l_{jk} + [Q, \Lambda]) \rangle_0 = \langle \phi_i \phi_l \rangle_0 C^l_{jk}, \tag{16.40}$$

where we use the vanishing of Q-trivial operators. We thus see that

$$C_{ijk} = \eta_{il} C^l_{jk}. \tag{16.41}$$

Using the invertibility of the matrix η_{il}, we find

$$C^i_{jk} = \eta^{il} C_{ljk}. \tag{16.42}$$

Thus, we have shown that the chiral ring is determined by computing three-point (B-twisted) topological correlation functions at genus 0. From the property of the topological correlation functions studied above, the chiral ring structure constants depend holomorphically on the chiral parameters only. (Similarly, the twisted chiral ring is determined by three-point correlators of the A-twisted theory. The structure constants depend holomorphically on the twisted chiral parameters only.)

16.4. Examples

We work out twisting in some detail for several classes of theories. The classes of theories we consider here are non-linear sigma models and Landau–Ginzburg models. For the non-linear sigma model on a Kähler manifold X, $U(1)_V$ is always unbroken but $U(1)_A$ R-symmetry is anomalously broken if the first Chern class $c_1(X)$ is non-vanishing. Thus, for a manifold with $c_1(X) \neq 0$, only the A-twist is possible, while for a Calabi–Yau manifold, the B-twist is also possible since $c_1(X) = 0$. For the linear sigma model, the condition for unbroken $U(1)_A$ (B-twistability) is whether the sums of charges are all zero or not. For Landau–Ginzburg models, $U(1)_V$ is broken by the superpotential unless the superpotential is quasi-homogeneous, while the condition for unbroken $U(1)_A$ is the same as in the non-linear sigma models. In what follows, when we say LG model we mean that the target space is Calabi–Yau. Then the LG model is always B-twistable.

We consider A-twist of non-linear sigma models on general target spaces, B-twist of LG models, and B-twist of Calabi–Yau sigma models. The structure of the theory for the A-twist of linear sigma models is similar to that of the non-linear sigma model and we leave it to the reader to work it out in detail. We explicitly compute the chiral ring in several examples.

16.4.1. A-Twist of Non-linear Sigma Models. We first consider the A-twist of the non-linear sigma model on a Kähler manifold X of dimension n. Before twisting, it is described by the n chiral multiplet fields Φ^i. The lowest components ϕ^i represent the complex coordinates of the map of the worldsheet to the target space

$$\phi : \Sigma \to M. \tag{16.43}$$

The fermions $\psi^i_\pm$ are considered as the components of the spinors $\psi_\pm$ with values in $\phi^*TM^{(1,0)}$. The Lagrangian is given in Eq. (13.15) with $W = 0$, and the supersymmetry variations of the component fields are given in Eq. (13.16).

Now let us perform the A-twist. This is done simply by changing the spin of the fermions $\psi_\pm$ and $\overline{\psi}_\pm$. The changes are as in the table shown before; ψ_- and $\overline{\psi}_+$ are now scalars while ψ_+ and $\overline{\psi}_-$ are anti-holomorphic and holomorphic one-forms respectively (where all these forms are valued in the pull-back of suitable tangent bundles of M). In order to make the new spin manifest, we rename these fields as

$$\begin{aligned} \chi^i &:= \psi^i_-, \quad \chi^{\bar\imath} := \overline{\psi}^{\bar\imath}_+, \\ \rho^{\bar\imath}_z &:= \overline{\psi}^{\bar\imath}_-, \quad \rho^i_{\bar z} := \psi^i_+. \end{aligned} \tag{16.44}$$

The action is then written as

$$\begin{aligned} S = \int d^2z\Big(& g_{i\bar\jmath}h^{\mu\nu}\partial_\mu\phi^i\partial_\nu\overline{\phi}^{\bar\jmath}\sqrt{h} - ig_{i\bar\jmath}\rho^{\bar\jmath}_z D_{\bar z}\chi^i \\ &+ ig_{i\bar\jmath}\rho^i_{\bar z}D_z\chi^{\bar\jmath} - \frac{1}{2}R_{i\bar k j\bar l}\rho^i_{\bar z}\chi^j\rho^{\bar k}_z\chi^{\bar l}\Big). \end{aligned} \tag{16.45}$$

We now look at the action of the scalar supercharge $\overline{Q}_+$ and Q_-. For this we simply have to set $\overline{\epsilon}_+ = \epsilon_- = 0$ in the supersymmetry transformation given by Eq. (13.16). This gives us the variation of the fields under $\delta = \overline{\epsilon}_-\overline{Q}_+ + \epsilon_+Q_-$:

$$\begin{aligned} &\delta\phi^i = \epsilon_+\chi^i, && \delta\overline{\phi}^{\bar\imath} = \overline{\epsilon}_-\chi^{\bar\imath}, \\ &\delta\rho^i_{\bar z} = 2i\overline{\epsilon}_-\partial_{\bar z}\phi^i + \epsilon_+\Gamma^i_{jk}\rho^j_{\bar z}\chi^k, && \delta\chi^{\bar\imath} = 0, \\ &\delta\chi^i = 0, && \delta\rho^{\bar\imath}_z = -2i\epsilon_+\partial_z\overline{\phi}^{\bar\imath} + \overline{\epsilon}_-\Gamma^{\bar\imath}_{\bar\jmath\bar k}\rho^{\bar k}_z\chi^{\bar\jmath}. \end{aligned} \tag{16.46}$$

The variation under the Q operator $Q_A = \overline{Q}_+ + Q_-$ is obtained by setting $\epsilon_+ = \overline{\epsilon}_-$ in these formulae.

Physical Operators. Let us analyze the $Q = Q_A$-cohomology classes of operators. Let us focus on operators associated to points on the manifold (i.e., operators of type $\mathcal{O}^{(0)}$ and not $\int_\gamma \mathcal{O}^{(1)}$ or $\int_\Sigma \mathcal{O}^{(2)}$). To have a covariant zero-form operator, we can only use the scalar fields ϕ and χ. We cannot use their derivatives, nor ρ which is a one-form on the worldsheet, because the only way to construct a zero-form out of them is to use the worldsheet metric and thus the operator we get would be Q-exact.

EXERCISE 16.4.1. *Show this by using the fact that the variation of the worldsheet metric is a Q-trivial operation and so it should not change the topological correlation functions.*

We thus consider operators made up only of ϕ and χ. We can associate such operators to the differential forms on X according to the rule

$$\chi^i \leftrightarrow dz^i, \; \chi^{\bar\imath} \leftrightarrow d\overline{z}^{\bar\imath}, \tag{16.47}$$

or more generally

$$\begin{aligned} &\omega_{i_1i_2\cdots i_p\bar\jmath_1\bar\jmath_2\cdots\bar\jmath_q}(\phi)\chi^{i_1}\chi^{i_2}\cdots\chi^{i_p}\chi^{\bar\jmath_1}\chi^{\bar\jmath_2}\cdots\chi^{\bar\jmath_q} \\ \longleftrightarrow \quad &\omega_{i_1i_2\cdots i_p\bar\jmath_1\bar\jmath_2\cdots\bar\jmath_q}(z)dz^{i_1}dz^{i_2}\cdots dz^{i_p}d\overline{z}^{\bar\jmath_1}d\overline{z}^{\bar\jmath_2}\cdots d\overline{z}^{\bar\jmath_q}. \end{aligned} \tag{16.48}$$

It is easy to see that the operators Q_- and $\overline{Q}_+$ are identified as the Dolbeault operators

$$Q_- \leftrightarrow \partial,\ \overline{Q}_+ \leftrightarrow \overline{\partial}, \tag{16.49}$$

and Q_A corresponds to the de Rham operator (exterior derivative) $d = \partial + \overline{\partial}$. Then the Q_A-cohomology classes are identified as the d-cohomology classes of differential forms. Namely, they are the de Rham cohomology classes. Thus, the physical operators of this class are in one-to-one correspondence with the de Rham cohomology classes,

$$\Big\{\text{physical operator}\Big\} \cong H^*_{DR}(X). \tag{16.50}$$

A particularly useful representative of the Q-cohomology class is in terms of the dual homology cycles. Let D be a homology cycle of real codimension r. Then its Poincaré dual $[D]$ is a cohomology class in $H^r(X)$ which is represented as the delta function r-form supported on D. Let us denote the corresponding operator by $\mathcal{O}_D$. This operator, inserted at $x \in \Sigma$, $\mathcal{O}_D(x)$, vanishes for a configuration in which the map ϕ sends x outside the cycle D.

Correlation Functions. Let us analyze the correlation function of physical operators $\mathcal{O}_i$,

$$\langle \mathcal{O}_1 \cdots \mathcal{O}_s \rangle = \int \mathcal{D}\phi \mathcal{D}\chi \mathcal{D}\rho\ \mathrm{e}^{-S} \mathcal{O}_1 \cdots \mathcal{O}_s. \tag{16.51}$$

The path-integral is over all possible configurations. In particular, it contains the integral over all possible maps ϕ from Σ to X. We classify the space of maps by the homology class of the image:

$$\beta = \phi_*[\Sigma] \in H_2(X, \mathbb{Z}). \tag{16.52}$$

Accordingly, the path-integral is decomposed into the sum over these homology classes,

$$\langle \mathcal{O}_1 \mathcal{O}_2 \cdots \mathcal{O}_s \rangle = \sum_{\beta \in H_2(M,\mathbb{Z})} \langle \mathcal{O}_1 \mathcal{O}_2 \cdots \mathcal{O}_s \rangle_\beta, \tag{16.53}$$

where

$$\langle \mathcal{O}_1 \mathcal{O}_2 \cdots \mathcal{O}_s \rangle_\beta := \int_{\phi_*[\Sigma]=\beta} \mathcal{D}\phi \mathcal{D}\chi \mathcal{D}\rho\ \mathrm{e}^{-S} \mathcal{O}_1 \mathcal{O}_2 \cdots \mathcal{O}_s. \tag{16.54}$$

Selection Rule. The classical $U(1)_V$ and $U(1)_A$ R-symmetries of the theory remain after twisting as the symmetry of the classical Lagrangian. As a consequence, the correlation function given by Eq. (16.54) obeys certain

selection rules. Let us assume that each operator $\mathcal{O}_i$ corresponds to a differential form of definite Hodge degrees:

$$\mathcal{O}_i \longleftrightarrow \omega_i \in H^{p_i,q_i}(X). \tag{16.55}$$

This operator has vector R-charge $q_V = -p_i + q_i$ and axial R-charge $q_A = p_i + q_i$.

First, the vector R-symmetry is not anomalous, as in the theory before twisting. Thus, the correlation function given by Eq. (16.54) is non-vanishing only if $\sum_{i=1}^s p_i = \sum_{i=1}^s q_i$. Second, the axial R-symmetry is generically anomalous. The anomaly manifests itself as the mismatch of the fermion zero modes. For a fixed map $\phi : \Sigma \to X$, the relevant mismatch is

$$\#(\chi \text{ zero modes}) - \#(\rho \text{ zero modes}).$$

This is the index of the differential operators that appear in the fermion kinetic terms. They are the Dolbeault operators in this case (rather than the Dirac operator for the untwisted theory) and the index is $2k$, where

$$\begin{aligned} k = &\int_\Sigma \mathrm{ch}(\phi^* TX^{(1,0)})\mathrm{td}(\Sigma) = \int_\Sigma \phi^* c_1(X) + \dim X(1-g) \\ = &c_1(X)\cdot\beta + \dim X(1-g). \end{aligned} \tag{16.56}$$

Thus, the axial rotation $\mathrm{e}^{iF_A\alpha}$ rotates the path-integral measure by $\mathrm{e}^{i2k\alpha}$. Note that it depends only on the homology class $\beta = \phi_*[\Sigma]$ of the map. Therefore, the correlation function given by Eq. (16.54) for the fixed degree β must obey the selection rule $\sum_{i=1}^s (p_i + q_i) = 2k$. Combining this with the selection rule $\sum_i p_i = \sum_i q_i$ from the vector R-symmetry, we see that the correlation function in Eq. (16.54) is non-vanishing only when

$$\sum_{i=1}^s p_i = \sum_{i=1}^s q_i = c_1(M)\cdot\beta + \dim X(1-g). \tag{16.57}$$

Localization to Q-fixed Points (Generic Case). We now make use of the localization principle. Since there is a fermionic symmetry Q under which all inserted operators are invariant, the path-integral in Eq. (16.54) picks up contributions only from the loci where the Q-variation of the fermions vanishes. By looking at $\delta\rho^i_{\bar{z}}$ and $\delta\rho^{\bar{i}}_z$ in Eq. (16.46), we see that a Q-fixed point obeys

$$\partial_{\bar{z}}\phi^i = 0. \tag{16.58}$$

This simply says that the map $\phi : \Sigma \to X$ is a holomorphic map (for a fixed metric of Σ, namely for a fixed complex structure of Σ). The path-integral localizes on such configurations. The bosonic part of the action S_b is given by

$$\begin{aligned} S_b &= \int_\Sigma g_{i\bar{j}} \left(\partial_z \phi^i \partial_{\bar{z}} \overline{\phi^j} + \partial_{\bar{z}} \phi^i \partial_z \overline{\phi^j}\right) d^2 z \\ &= 2 \int_\Sigma g_{i\bar{j}} \partial_{\bar{z}} \phi \partial_z \overline{\phi^j}\, d^2 z + \int_\Sigma \phi^* \omega \geq \int_\Sigma \phi^* \omega = \omega \cdot \beta \end{aligned} \tag{16.59}$$

where ω is the Kähler form. It is bounded from below, and the minimum $S_b = \omega \cdot \beta$ is attained for a holomorphic map. (This gives another reason why the path-integral receives the dominant contribution at holomorphic maps; however, by deformation invariance these are not just dominant contributions but actually the only relevant ones.) If we had a non-trivial B-field, the action for a holomorphic map would be

$$S_b = \int_\Sigma \phi^*(\omega - iB) = (\omega - iB) \cdot \beta. \tag{16.60}$$

Let us make some technical assumptions; We assume that the number k in Eq. (16.56) is non-negative and there is no ρ zero mode. We further assume that the moduli space of homolorphic maps of degree β,

$$\mathcal{M}_\Sigma(X, \beta) = \left\{ \phi : \Sigma \to X \;\middle|\; \begin{array}{c} \text{holomorphic} \\ \phi_*[\Sigma] = \beta \end{array} \right\}, \tag{16.61}$$

is a smooth manifold. Then the tangent space of the moduli space $\mathcal{M}_\Sigma(X, \beta)$ is identified as the space of χ zero modes and hence

$$\dim_{\mathbb{C}} \mathcal{M}_\Sigma(X, \beta) = k. \tag{16.62}$$

The path-integral from Eq. (16.54) reduces to the integral over the finite-dimensional space $\mathcal{M}_\Sigma(X, \beta)$, where the measure is given as a result of the integration of the infinitely many nonzero modes. The latter integration gives 1 due to the cancellation of bosonic and fermionic determinants. Therefore the measure on $\mathcal{M}_\Sigma(X, \beta)$ simply comes from the inserted operators (and the weight $e^{-(\omega - iB)\cdot\beta}$). The operator $\mathcal{O}_i$ (inserted at $x_i \in \Sigma$) can be identified as the pull-back of $\omega_i \in H^*(X)$ by the evaluation map at x_i

$$\begin{aligned} \mathrm{ev}_i : \mathcal{M}_\Sigma(X, \beta) &\longrightarrow X, \\ \phi &\longmapsto \phi(x_i). \end{aligned} \tag{16.63}$$

Then the correlation function is given by

$$\langle \mathcal{O}_1 \cdots \mathcal{O}_s \rangle_\beta = e^{-(\omega - iB)\cdot\beta} \int_{\mathcal{M}_\Sigma(X,\beta)} \mathrm{ev}_1^* \omega_1 \wedge \cdots \wedge \mathrm{ev}_s^* \omega_s. \tag{16.64}$$

The total form degree indeed agrees with the dimension of the moduli space if and only if the selection rule Eq. (16.57) is satisfied.

If $[\omega_i]$ are the Poincaré duals of the cycles D_i in X, the integral on the right-hand side of Eq. (16.64) has a simple geometric meaning. In such a case, ω_i can be chosen to be the delta function form supported on D_i and the integral can be identified as the number of holomorphic maps (of degree β) where x_i is mapped into D_i:

$$n_{\beta, D_1, \ldots, D_s} = \# \left\{ \phi : \Sigma \to X \;\middle|\; \begin{array}{c} \text{holomorphic} \\ \phi(x_i) \in D_i \ \forall i \\ \phi_*[\Sigma] = \beta \end{array} \right\}. \tag{16.65}$$

Thus, the total correlation function is

$$\langle \mathcal{O}_1 \cdots \mathcal{O}_s \rangle = \sum_{\beta \in H_2(X, \mathbb{Z})} e^{-(\omega - iB)\cdot\beta}\, n_{\beta, D_1, \ldots, D_s}. \tag{16.66}$$

For a homology class β that contributes to the sum in Eq. (16.66), we have the bound

$$\omega \cdot \beta \geq 0. \tag{16.67}$$

This is because the Kähler form ω restricted to a holomorphic curve $\phi(\Sigma)$ is positive semi-definite. It vanishes only if $\phi(\Sigma)$ is a point; namely, for the case where $\beta = 0$. Thus, in the large volume limit where ω becomes large, the sum in Eq. (16.66) is dominated by the $\beta = 0$ contribution. The moduli space of $\beta = 0$ holomorphic maps is the moduli space of a point in X, namely X itself

$$\mathcal{M}_\Sigma(X, 0) \cong X. \tag{16.68}$$

The evaluation map ev_i is the identity map for all i; $\mathrm{ev}_i = \mathrm{id}_X$. Here we have been assuming that there are no ρ zero modes. This implies, from Eq. (16.57), since $\beta = 0$ and assuming the left-hand side of that equation is positive (which is always the case except for no insertions) that we are dealing with a genus $g = 0$ Riemann surface, i.e., a sphere. This selection rule is the same as in the case of classical intersection theory. In fact we have for the degree 0 contribution on the sphere

$$\langle \mathcal{O}_1 \cdots \mathcal{O}_s \rangle_0 = \int_X \omega_1 \wedge \cdots \wedge \omega_s = \#(D_1 \cap \cdots \cap D_s). \tag{16.69}$$

Thus, the correlation function from Eq. (16.66) can be interpreted as a quantum deformation of the classical intersection numbers.

We finally note that the topological metric is given by

$$\eta_{ij} = \langle 1 \mathcal{O}_i \mathcal{O}_j \rangle_0 = \int_X \omega_i \wedge \omega_j. \tag{16.70}$$

Namely, it is the intersection pairing on X. However unlike C_{ijk} it turns out that η does not receive quantum corrections from non-trivial holomorphic maps, and continues to be given by the classical intersection pairing for the full quantum theory.

Non-generic Case. We next consider the cases where some of the assumptions made above are relaxed. In particular, we consider the cases where there are some ρ zero modes. We recall that a ρ zero mode is a solution to the equation

$$\partial_{\bar{z}} \rho_{zi} = 0, \tag{16.71}$$

where we have used the variable $\rho_{zi} = g_{i\bar{j}} \rho_z^{\bar{j}}$. Thus the space of ρ-zero modes is the space of holomorphic sections

$$H^0(\Sigma, K \otimes \phi^* T_M^*). \tag{16.72}$$

We assume that the dimension of the space Eq. (16.72) (or the number of ρ zero modes) is a constant, ℓ, along the moduli space of holomorphic maps. Then the dimension of the moduli space $\mathcal{M}_\Sigma(X, \beta)$ is $k + \ell$, and the family of vector spaces shown in Eq. (16.72) defines a vector bundle $\mathcal{V}$ of rank ℓ over $\mathcal{M}_\Sigma(X, \beta)$.

The path-integral starts with the integration over infinitely many nonzero modes in the quadratic approximation. The bosonic and fermionic determinants almost cancel with each other and we are left with the action for the zero modes which reads as

$$S_0 = \int_\Sigma d^2 z \left(\frac{1}{2} \rho_{\bar{z}}^i R_{i\bar{k}j\bar{l}} \chi^j \chi^{\bar{l}} \rho_z^{\bar{k}} - \frac{1}{4} \rho_{\overline{z j}} \chi^{\bar{k}} \partial_z \phi^l R^{\bar{j}i}{}_{\bar{k}l} G^{z\bar{z}}{}_i{}^j \chi^k \partial_{\bar{z}} \overline{\phi}^{\bar{l}} R^m{}_{jk\bar{l}} \rho_{zm} \right). \tag{16.73}$$

The first term is the four-fermi interaction in the classical action given by Eq. (16.45) and the second term comes from completing the square of the bosonic nonzero modes. ($G^{z\bar{z}}{}_i{}^j$ in the second term represents the inverse of the Laplacian $D_{\bar{z}} D_z$ that appears in the bosonic kinetic term.) One can write this action as

$$S_0 = (\rho, F_\mathcal{V} \rho) \tag{16.74}$$

where $(,)$ is a Hermitian inner product on the bundle $\mathcal{V}$ and $F_\mathcal{V}$ is an expression bilinear in χ. One can show that this $F_\mathcal{V}$ is proportional to the curvature of a Hermitian connection of $\mathcal{V}$ if we identify the χ's as the one-forms on $\mathcal{M}_\Sigma(X, \beta)$. As we have seen in the zero-dimensional quantum field theory, integration of $e^{-(\rho, F_\mathcal{V} \rho)}$ over ρ yields the Pfaffian of $F_\mathcal{V}$. Up to a constant, this is equal to the Euler class of $\mathcal{V}$:

$$\mathrm{Pf}(F_\mathcal{V}) \propto e(\mathcal{V}). \tag{16.75}$$

Then the correlation function can be written as

$$\langle \mathcal{O}_1 \cdots \mathcal{O}_s \rangle_\beta = \int_{\mathcal{M}_\Sigma(X,\beta)} e(\mathcal{V}) \wedge \mathrm{ev}_1^* \omega_1 \wedge \cdots \wedge \mathrm{ev}_s^* \omega_s. \tag{16.76}$$

Note that $e(\mathcal{V})$ is represented by an (ℓ, ℓ)-form and the integrand has exactly the right form-degree to be integrated over the $(k + \ell)$-dimensional space $\mathcal{M}_\Sigma(X, \beta)$, provided the selection rule $\sum_i p_i = \sum_i q_i = k$ is obeyed.

EXAMPLE 16.4.1 ($X = \mathbb{CP}^1$). *As an example, consider the case when the target space is $\mathbb{CP}^1$. We determine the twisted chiral ring of the $\mathbb{CP}^1$ sigma model by computing the three-point topological correlators. The cohomology group of $\mathbb{CP}^1$ is non-trivial for $H^0(\mathbb{CP}^1)$ and $H^2(\mathbb{CP}^1)$, where $H^2(\mathbb{CP}^1)$ is generated by the class H which is Poincaré dual to a point. It therefore integrates to 1;*

$$\int_{\mathbb{CP}^1} H = 1. \tag{16.77}$$

We denote by P and Q the operators corresponding to the cohomology class $1 \in H^0(\mathbb{CP}^1)$ and $H \in H^2(\mathbb{CP}^1)$ respectively. Since Eq. (16.77) is the only non-trivial integral, we have

$$\langle P \mathcal{O}_\alpha \mathcal{O}_\beta \rangle = \eta_{\alpha\beta} = \begin{cases} 1 & \eta_{PQ} \textit{ and } \eta_{QP}, \\ 0 & \textit{otherwise.} \end{cases} \tag{16.78}$$

The remaining correlator to consider is $\langle QQQ \rangle$. *It is expanded as*

$$\langle QQQ \rangle = \sum_{n \in \mathbb{Z}} \langle QQQ \rangle_n, \tag{16.79}$$

where $\langle QQQ \rangle_n$ *is the contribution from the degree* n *maps* ($\beta = n[\mathbb{CP}^1]$). *Since the first Chern class is* $c_1(\mathbb{CP}^1) = 2H$, *the axial anomaly for degree* n *maps is* $2k$ *with* $k = c_1(\mathbb{CP}^1) \cdot \beta + \dim \mathbb{CP}^1 (1-0) = 2n+1$. *Since* Q *has axial charge* 2, *only the degree* $n = 1$ *maps contribute to this correlation function. The correlation function can be computed using Eq. (16.65). Since* Q *corresponds to the class* H *which is Poincaré dual to a point* $y \in X = \mathbb{CP}^1$, *the correlator is the number of maps where three distinct insertion points* $x_1, x_2, x_3 \in \Sigma = \mathbb{CP}^1$ *are mapped to arbitrarily chosen distinct points* $y_1, y_2, y_3 \in X = \mathbb{CP}^1$. *It is obvious that there is only one such map;*

$$n_{1, y_1, y_2, y_3} = 1. \tag{16.80}$$

Thus, we have shown

$$\langle QQQ \rangle = \langle QQQ \rangle_1 = \mathrm{e}^{-t}, \tag{16.81}$$

where we have abbreviated $(\omega - iB) \cdot [\mathbb{CP}^1]$ *by* t. *What we have computed determines the twisted chiral ring as*

$$\begin{aligned} PP &= P, \\ PQ &= QP = Q, \\ QQ &= \mathrm{e}^{-t} P. \end{aligned} \tag{16.82}$$

In the classical cohomology ring, the last equation would be $QQ = 0$. *Note that the chiral ring given above reduces to the classical ring as we let* $t \to \infty$, *as expected. The above ring is thus the quantum deformation of the cohomology ring of* $\mathbb{CP}^1$, *and is sometimes called its quantum cohomology ring.*

16.4.2. B-Twist of Landau–Ginzburg Models. We next consider the B-twist of the Landau–Ginzburg model on a (non-compact) Calabi–Yau manifold M with a superpotential

$$W : M \to \mathbb{C}. \tag{16.83}$$

We will soon focus our attention on the case where M is flat, but for later use we will be general for the moment. The Lagrangian and the supersymmetry transformations before twisting are as in Eqs. (13.15)–(13.16). The B-twist

is done by changing the spin of the fermions $\psi_\pm$ and $\overline{\psi}_\pm$. The changes are as in the table: $\overline{\psi}_\pm$ are now scalars while ψ_+ and ψ_- are anti-holomorphic and holomorphic one-forms respectively (all with values in the pull-back of the holomorphic or anti-holomorphic tangent bundle of M). We rename these fields as

$$\begin{aligned} \psi^{\bar{\imath}} := \overline{\psi}^{\bar{\imath}}_-, \quad \overline{\psi}^{\bar{\imath}} := \overline{\psi}^{\bar{\imath}}_+, \\ \rho^i_z := \psi^i_-, \quad \rho^i_{\bar{z}} := \psi^i_+. \end{aligned} \tag{16.84}$$

The action is written

$$\begin{aligned} S = \int d^2z \Big(& g_{i\bar{\jmath}} h^{\mu\nu} \partial_\mu \phi^i \partial_\nu \overline{\phi}^{\bar{\jmath}} \sqrt{h} - i g_{i\bar{\jmath}} \psi^{\bar{\jmath}} D_{\bar{z}} \rho^i_z \\ & + i g_{i\bar{\jmath}} \overline{\psi}^{\bar{\jmath}} D_z \rho^i_{\bar{z}} - \frac{1}{2} R_{i\bar{k}j\bar{l}} \rho^i_{\bar{z}} \rho^j_z \psi^{\bar{k}} \overline{\psi}^{\bar{l}} \\ & + \frac{1}{8} g^{\bar{\jmath}i} \partial_{\bar{\jmath}} \overline{W} \partial_i W + \frac{1}{4} (D_i \partial_j W) \rho^i_{\bar{z}} \rho^j_z + \frac{1}{4} (D_{\bar{\imath}} \partial_{\bar{\jmath}} \overline{W}) \psi^{\bar{\imath}} \overline{\psi}^{\bar{\jmath}} \Big). \end{aligned} \tag{16.85}$$

The scalar supercharges in the B-twisted model are $\overline{Q}_+$ and $\overline{Q}_-$. Their action on the fields can be seen by setting $\epsilon_\pm = 0$ in the supersymmetry transformations as shown by Eq. (13.16). This gives us the variation of the fields under $\delta = \overline{\epsilon}_- \overline{Q}_+ - \overline{\epsilon}_+ \overline{Q}_-$;

$$\begin{aligned} \delta\phi^i &= 0, & \delta\overline{\phi}^{\bar{\imath}} &= -\overline{\epsilon}_+ \psi^{\bar{\imath}} + \overline{\epsilon}_- \overline{\psi}^{\bar{\imath}}, \\ \delta\rho^i_{\bar{z}} &= 2i\overline{\epsilon}_- \partial_{\bar{z}} \phi^i, & \delta\overline{\psi}^{\bar{\imath}} &= \overline{\epsilon}_+ (-\tfrac{1}{2} g^{\bar{\imath}j} \partial_j W + \Gamma^{\bar{\imath}}_{\bar{\jmath}\bar{k}} \bar{\Psi}^{\bar{\jmath}}_- \bar{\Psi}^{\bar{k}}_+), \\ \delta\rho^i_z &= 2i\overline{\epsilon}_+ \partial_z \phi^i, & \delta\psi^{\bar{\imath}} &= \overline{\epsilon}_- (-\tfrac{1}{2} g^{\bar{\imath}j} \partial_j W + \Gamma^{\bar{\imath}}_{\bar{\jmath}\bar{k}} \bar{\Psi}^{\bar{\jmath}}_- \bar{\Psi}^{\bar{k}}_+). \end{aligned} \tag{16.86}$$

The variation under the Q operator $Q_B = \overline{Q}_+ + \overline{Q}_-$ is obtained by setting $\overline{\epsilon}_+ = -\overline{\epsilon}_-$ in these formulae.

Now let us focus on the LG model on a flat manifold, M.

Physical Operators. Let us first determine the physical operators of the model. For this purpose it is convenient to rewrite Eq. (16.86) for the Q_B variation $\overline{\epsilon}_+ = -\overline{\epsilon}_- =: \overline{\epsilon}$ in the following form:

$$\begin{aligned} \delta\phi^i &= 0, & & \\ \delta(\psi^{\bar{\imath}} - \overline{\psi}^{\bar{\imath}}) &= \overline{\epsilon} g^{\bar{\imath}j} \partial_j W, & \delta\overline{\phi}^{\bar{\imath}} &= -\overline{\epsilon}(\psi^{\bar{\imath}} + \overline{\psi}^{\bar{\imath}}), \\ \delta\rho^i_\mu &= -2\overline{\epsilon} J^\nu_\mu \partial_\nu \Phi^i. & \delta(\psi^{\bar{\imath}} + \overline{\psi}^{\bar{\imath}}) &= 0. \end{aligned} \tag{16.87}$$

Here we have used (local) flat coordinates on M. From this formulation it is obvious that the physical operators are holomorphic combinations of ϕ^i, namely holomorphic functions on M. A function is Q_B-trivial (Q_B-exact) if and only if it can be written as

$$vW = v^i \partial_i W, \tag{16.88}$$

where $v := v^i\partial/\partial z^i$ is a holomorphic vector field on M. Thus, the space of physical operators is the space of holomorphic functions of M modulo the subspace spanned by functions of the form Eq. (16.88). We denote the operator corresponding to a function f by O_f. If $M = \mathbb{C}^n$, the physical operators are polynomials in $\phi^1, \dots, \phi^n$ modulo $\partial_1 W, \dots, \partial_n W$. The chiral ring is simply the ring of functions on M modulo the ideal of functions of the form Eq. (16.88):

$$O_f O_g = O_{fg}. \tag{16.89}$$

In the case of $X = \mathbb{C}^n$ it is the ring of polynomials

$$\text{chiral ring} = \mathbb{C}[\phi^1, \dots, \phi^n]/(\partial_i W). \tag{16.90}$$

We will reproduce this also from the point of view of topological correlation functions.

Correlation Functions. Let us analyze the topological correlation functions

$$\langle O_{f_1} \cdots O_{f_s} \rangle = \int \mathcal{D}\phi \mathcal{D}\psi \mathcal{D}\rho \, \mathrm{e}^{-S} O_{f_1} \cdots O_{f_s}. \tag{16.91}$$

Again we make use of the localization principle to evaluate this function. By looking at Eq. (16.86) or Eq. (16.87) we see that a Q-fixed point obeys

$$\partial_\mu \phi^i = 0, \tag{16.92}$$

$$\partial_i W = 0. \tag{16.93}$$

Namely, it is a constant map into a critical point of W. To simplify the analysis we assume that there are only isolated, non-degenerate critical points $y_1, \dots, y_N$. The path-integral in Eq. (16.91) decomposes into the sum over critical points

$$\langle O_{f_1} \cdots O_{f_s} \rangle = \sum_{i=1}^{N} \langle O_{f_1} \cdots O_{f_s} \rangle |_{y_i}. \tag{16.94}$$

Each summand can be computed by the quadratic approximation around $\phi(x) \equiv y_i$. The integration variables are classified into non-constant modes, where the kinetic terms in Eq. (16.85) are non-trivial, and the "constant modes" that annihilate the kinetic terms. The bosonic and fermionic non-constant modes are paired as usual, and the determinants from their integrals cancel against each other. On the other hand, the constant modes are not paired. Each of ϕ^i, $\overline{\phi^i}$, ψ^i and $\overline{\psi^i}$ has a single constant mode. Each of ρ^i_z and $\rho^i_{\bar z}$ has g "constant modes" where g is the genus of the worldsheet Σ. By a standard computation, the integrals over these constant modes are

$$\int d^{2n}\phi \, \mathrm{e}^{-\frac{1}{4} g^{i\bar{j}} \partial_i W \partial_{\bar{j}} \overline{W}} = \frac{1}{|\det \partial_i \partial_j W|^2 (y_i)}, \tag{16.95}$$

$$\int d^n\overline{\psi} d^n\psi \, \mathrm{e}^{-\frac{1}{2} \partial_{\bar{i}} \partial_{\bar{j}} \overline{W} \psi^{\bar{i}} \psi^{\bar{j}}} = \det \partial_{\bar{i}} \partial_{\bar{j}} \overline{W} (y_i), \tag{16.96}$$

$$\int d^{ng}\rho d^{ng}\overline{\rho} \, \mathrm{e}^{-\frac{1}{2} \partial_i \partial_j W \rho^i_{\bar{z}} \rho^j_z} = (\det \partial_i \partial_j W)^g (y_i), \tag{16.97}$$

and the product is simply

$$(\det \partial_i \partial_j W)^{g-1} (y_i). \tag{16.98}$$

Therefore, the correlation function is given by

$$\langle O_{f_1} \cdots O_{f_s} \rangle_g = \sum_{i=1}^{N} f_1(y_i) \cdots f_s(y_i) (\det \partial_i \partial_j W)^{g-1} (y_i). \tag{16.99}$$

Note in particular the independence from $\overline{W}$, as was expected by general arguments discussed earlier. The result for $g = 1$ is the same as the correlation functions of the zero-dimensional and one-dimensional LG QFT discussed before. This is not an accident. If we consider the LG theory on a T^2 with periodic boundary conditions for fermions, then as we deform the metric on T^2 we can obtain a reduction to S^1 or to a point, i.e., reduction to one-dimensional or zero-dimensional QFT. Since the topological correlation functions are independent of the metric of T^2 this implies that the answer should have agreed with those obtained in the one-dimensional and zero-dimensional cases discussed earlier.

From Eq. (16.99), we see that the genus 0 three-point functions and the topological metric are given by

$$C_{ijk} = \sum_{dW=0} \frac{f_i f_j f_k}{\det \partial_i \partial_j W}, \tag{16.100}$$

$$\eta_{ij} = \sum_{dW=0} \frac{f_i f_j}{\det \partial_i \partial_j W}. \tag{16.101}$$

EXERCISE 16.4.2. *Show that the three-point function is compatible with the chiral ring of the topological LG theory given by Eqs. (16.89)–(16.90).*

EXAMPLE 16.4.2 (Sine-Gordon model). *We take M to be the flat cylinder $\mathbb{C}^\times$ with the global coordinate z so that $\mathbb{C}^\times = \{z \neq 0\}$. The flat coordinate on the cylinder is given by $\log z$. We consider the LG model on $\mathbb{C}^\times$ with the superpotential*

$$W = z + \mathrm{e}^{-t} z^{-1}. \tag{16.102}$$

The chiral ring is generated by $1, z$ with the ring relation

$$z^2 = \mathrm{e}^{-t}. \tag{16.103}$$

The critical points of the superpotential are given by $z = \pm \mathrm{e}^{-\frac{t}{2}}$. The Hessians at these points are $z\partial(z\partial W) = \pm 2\,\mathrm{e}^{-\frac{t}{2}}$. Thus the correlation functions can easily be determined using Eq. (16.99):

$$\begin{aligned}
\langle 1\,1\,1\rangle_{g=0} &= \frac{1\cdot 1\cdot 1}{2\,\mathrm{e}^{-\frac{t}{2}}} + \frac{1\cdot 1\cdot 1}{-2\,\mathrm{e}^{-\frac{t}{2}}} = 0,\\
\langle 1\,z\,z\rangle_{g=0} &= \frac{1\cdot \mathrm{e}^{-\frac{t}{2}}\cdot \mathrm{e}^{-\frac{t}{2}}}{2\,\mathrm{e}^{-\frac{t}{2}}} + \frac{1\cdot(-\mathrm{e}^{-\frac{t}{2}})\cdot(-\mathrm{e}^{-\frac{t}{2}})}{-2\,\mathrm{e}^{-\frac{t}{2}}} = 0,\\
\langle 1\,1\,z\rangle_{g=0} &= \frac{1\cdot 1\cdot \mathrm{e}^{-\frac{t}{2}}}{2\,\mathrm{e}^{-\frac{t}{2}}} + \frac{1\cdot 1\cdot(-\mathrm{e}^{-\frac{t}{2}})}{-2\,\mathrm{e}^{-\frac{t}{2}}} = \frac{1}{2} + \frac{1}{2} = 1,\\
\langle z\,z\,z\rangle_{g=0} &= \frac{(\mathrm{e}^{-\frac{t}{2}})^3}{2\,\mathrm{e}^{-\frac{t}{2}}} + \frac{(-\mathrm{e}^{-\frac{t}{2}})^3}{-2\,\mathrm{e}^{-\frac{t}{2}}} = \frac{\mathrm{e}^{-t}}{2} + \frac{\mathrm{e}^{-t}}{2} = \mathrm{e}^{-t}.
\end{aligned} \tag{16.104}$$

Note that the ring relation and the correlation functions agree with those for the A-twisted $\mathbb{CP}^1$ sigma model. As we will see later this is not a coincidence; this is actually a consequence of mirror symmetry.

16.4.3. B-Twist of Calabi–Yau Sigma Models.

As our final example, we consider the B-twist of the sigma model on a compact Calabi–Yau manifold M. For the Lagrangians and supersymmetry transformations of the twisted model, we can use the ones written above — Eqs. (16.85)–(16.86) — where we set $W = 0$. It is actually more convenient to change the variables as

$$\psi^{\bar{\imath}} + \overline{\psi}^{\bar{\imath}} = -\eta^{\bar{\imath}},\ \psi^{\bar{\imath}} - \overline{\psi}^{\bar{\imath}} = g^{\bar{\imath}j}\theta_j. \tag{16.105}$$

Then the action of the Q_B transformation simplifies to

$$\begin{aligned}
&\delta\phi^i = 0,\ \delta\theta_i = 0,\\
&\delta\overline{\phi}^{\bar{\imath}} = \bar{\epsilon}\eta^{\bar{\imath}},\ \delta\eta^{\bar{\imath}} = 0,\\
&\delta\rho^i_\mu = \pm 2i\bar{\epsilon}\partial_\mu\phi^i.
\end{aligned} \tag{16.106}$$

Physical operators. The space of physical operators can be read off by looking at Eq. (16.106). These are constructed from ϕ^i, $\overline{\phi}^{\bar{\imath}}$, $\eta^{\bar{\imath}}$ and θ_i. It is useful here to make the correspondence

$$\begin{aligned}
\eta^{\bar{\imath}} &\longleftrightarrow d\bar{z}^{\bar{\imath}},\\
\theta_i &\longleftrightarrow \frac{\partial}{\partial z^i}.
\end{aligned} \tag{16.107}$$

A general expression in ϕ^i, $\overline{\phi}^{\bar{\imath}}$, $\eta^{\bar{\imath}}$ and θ_i corresponds to

$$\omega^{j_1\cdots j_q}_{\bar{\imath}_1\cdots\bar{\imath}_p}\eta^{\bar{\imath}_1}\cdots\eta^{\bar{\imath}_p}\theta_{j_1}\cdots\theta_{j_q} \longleftrightarrow \omega^{j_1\cdots j_q}_{\bar{\imath}_1\cdots\bar{\imath}_p} d\bar{z}^{\bar{\imath}_1}\cdots d\bar{z}^{\bar{\imath}_p}\frac{\partial}{\partial z^{j_1}}\cdots\frac{\partial}{\partial z^{j_q}}, \tag{16.108}$$

which is identified as an anti-holomorphic p-form with values in the q-th exterior power of the holomorphic tangent bundle T_M – an element of $\Omega^{0,p}(M, \wedge^q T_M)$. The operator Q_B is identified as the Dolbeault operator $\overline{\partial}$ acting on the Dolbeault complex

(16.109)
$$0 \to \Omega^{0,0}(M,\wedge^q T_M) \xrightarrow{\overline{\partial}} \Omega^{0,1}(M,\wedge^q T_M) \xrightarrow{\overline{\partial}} \cdots \xrightarrow{\overline{\partial}} \Omega^{0,n}(M,\wedge^q T_M) \to 0.$$

Thus, the Q_B-cohomology is identified as the Dolbeault cohomology groups

$$\bigoplus_{p,q=0}^{n} H^{0,p}(M, \wedge^q T_M). \tag{16.110}$$

Correlation Functions. Let us consider the correlation function

$$\langle O_1\cdots O_s\rangle = \int \mathcal{D}\phi\mathcal{D}\eta\mathcal{D}\theta\ \mathrm{e}^{-S} O_1\cdots O_s, \tag{16.111}$$

where the O_i correspond to $\omega_i \in H^{0,p_i}(M, \wedge^{q_i} T_M)$. The $U(1)_V$ symmetry implies that this is non-vanishing only if $\sum_{i=1}^s p_i = \sum_{i=1}^s q_i$, whereas the $U(1)_A$ symmetry has an anomaly after twisting and tells us that this is non-vanishing only if $\sum_{i=1}^s (p_i + q_i) = 2\dim M(1-g) = 2n(1-g)$. Thus, the selection rule at genus $g = 0$ is

$$\sum_{i=1}^s p_i = \sum_{i=1}^s q_i = n. \tag{16.112}$$

On the other hand, at $g = 1$ the condition is $\sum_i p_i = \sum_i q_i = 0$. At higher genus, the condition can never be satisfied.

Let us evaluate the correlation function in Eq. (16.111) at $g = 0$ using the localization principle. As follows immediately from Eq. (16.106), a Q-fixed point obeys

$$\partial_\mu\phi^i = 0, \tag{16.113}$$

i.e., it is a constant map. The space of constant maps is the same as M itself. Thus, the path-integral reduces to an integral over M. It may appear that we only have to integrate $\omega := \omega_1 \wedge \cdots \wedge \omega_s$ over M. However, that is not an ordinary differential form but a $(0,p)$-form with values in $\wedge^q T_M$ where $p = \sum_{i=1}^s p_i$ and $q = \sum_{i=1}^s q_i$. We note here that $p = q = n$ when the selection rule Eq. (16.112) is satisfied. Then it is natural to expect that the product of the θ zero modes sends it to an (n,n)-form via

$$\omega \mapsto \langle \omega, \Omega \rangle \wedge \Omega := \omega^{i_1 \ldots i_n}_{\bar{\jmath}_1 \cdots \bar{\jmath}_n} d\bar{z}^{\bar{\jmath}_1} \cdots d\bar{z}^{\bar{\jmath}_n} \Omega_{i_1 \ldots i_n} \wedge \Omega, \tag{16.114}$$

where Ω is the holomorphic n-form of the Calabi–Yau manifold M. This indeed follows from the definition of the path-integral. Note that the integration over fermions requires a choice of a section for the holomorphic n-form Ω. This in particular means that the topological correlation functions are not really functions, but sections of a suitable bundle on themoduli space of complex structures of the Calabi–Yau, related to the choice of this section. Aspects of this will be important for our later discussions and a better global understanding of what topological partition functions are.

Thus, the correlation function is given by

$$\langle O_1 \cdots O_s \rangle = \int_M \langle \omega_1 \wedge \cdots \wedge \omega_s, \Omega \rangle \wedge \Omega. \tag{16.115}$$

In the case of a Calabi–Yau threefold, the three-point function of operators corresponding to the Beltrami differentials $\mu_1, \mu_2, \mu_3 \in H^1(M, T_M)$ is

$$\langle O_1 O_2 O_3 \rangle = \int_M \mu_1^i \wedge \mu_2^j \wedge \mu_3^k \Omega_{ijk} \wedge \Omega, \tag{16.116}$$

which is precisely the third-order derivative of the prepotential

$$\partial_1 \partial_2 \partial_3 \mathcal{G}, \tag{16.117}$$

as explained in Sec. 6.4.

CHAPTER 17

Chiral Rings and the Geometry of the Vacuum Bundle

We have seen two important aspects of $(2,2)$ supersymmetric theories. One of them is the structure of the vacuum states and the other is the structure of chiral fields and the ring that they form. The operator/state correspondence in QFT (which we will review in the present context below) relates the two: For each chiral field operator there is a vacuum state. However at first sight it appears that there is more information in the chiral rings than in the vacuum states. In particular the chiral ring gets deformed as we change the (relevant) superpotential term, whereas the number of ground states in the $(2,2)$ theories do not change. The question therefore is whether there is any further information in the structure of the ground states that encodes the structure of the chiral ring. The answer is yes. This information is encoded in how the vacuum states vary in the full Hilbert space of the theory as we change the superpotential parameter. The connection and metric on this vacuum bundle, and their relation to chiral rings, are described by the tt^* equations, which we now derive. It turns out that the chiral ring is an ingredient in formulating certain differential equations that lead to computation of the connection and the metric on the vacuum bundle.

17.1. tt^* Equations

Let $\mathcal{H}$ be the Hilbert space of any QFT. Suppose the Hilbert space $\mathcal{H}$ has a distinguished subspace V of fixed dimension. For example, in the context of $(2,2)$ theories we would be considering the space of ground states,

$$Q|\alpha\rangle = Q^\dagger|\alpha\rangle = 0\ , \quad |\alpha\rangle \in V, \tag{17.1}$$

which has a fixed dimension. We will study the effect of a change in the parameters of the physical theory on V. We denote the parameters by $m \in \mathcal{M}$. In the case of $(2,2)$ theories, it will turn out that the relevant parameters are the ones appearing in the relevant superpotential and its conjugate: $m = (t, \bar{t})$. Let us denote the corresponding subspace by $V(m)$. States, operators and the correlation functions will be continuous functions of the parameters. The space of parameters $\mathcal{M}$ is the moduli space of the theory, and is naturally a complex manifold. The family of subspaces $V(m)$ defines a natural bundle over this moduli space. In the context of $(2,2)$ theories, this is called the "vacuum bundle."

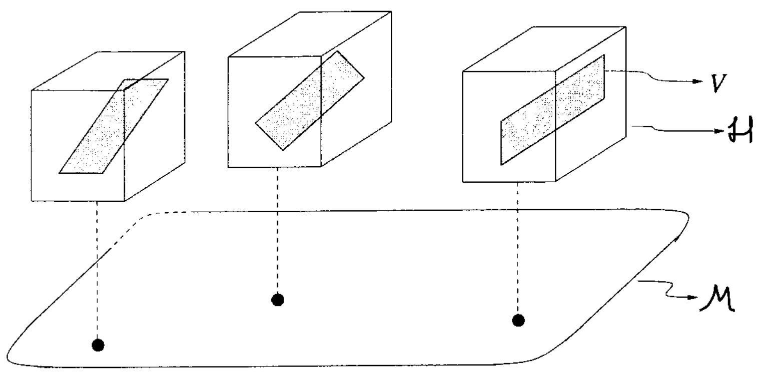


FIGURE 1. The V-bundle corresponds to a sub-bundle of the Hilbert space $\mathcal{H}$ over the moduli space of parameters $\mathcal{M}$ of the physical theory

Let $|\alpha(m_i)_j\rangle$ be an orthonormal basis of $V(m)$,

$$\langle\alpha(m)_k|\alpha(m)_j\rangle = \delta_{jk}. \tag{17.2}$$

These basis states are sections of the V-bundle over the moduli space. Note that as we change the parameters, the full Hilbert space of physical theories does not change, i.e., the Hilbert space $\mathcal{H}$ forms a trivial bundle over $\mathcal{M}$. The triviality of the Hilbert space bundle over $\mathcal{M}$ naturally defines a connection on the V-bundle,

$$(A_i)_j^k = \langle\alpha(m)_k|\frac{\partial}{\partial m_i}|\alpha(m)_j\rangle. \tag{17.3}$$

To see that the above equation defines a connection consider a change of the basis states,

$$\begin{aligned} |\widetilde{\alpha(m)}_j\rangle &= g_{ij}(m)|\alpha(m)_i\rangle, \\ A &\mapsto g^{-1}Ag + g^{-1}dg, \end{aligned} \tag{17.4}$$

consistent with the transformation property of the connection by a change of section.

One of the basic axioms of a QFT is that for each state in the Hilbert space there corresponds an operator that creates it from the vacuum state. Here we would like to study this correspondence in the context of the relation between the ground states of $(2,2)$ theories and chiral fields in the theory.

Suppose $|a\rangle$ denotes a ground state in a $(2,2)$ theory. Let ϕ be a chiral field. Then viewing ϕ as an operator,

$$\phi|a\rangle,$$

defines a state in the Hilbert space. If we consider the projection of this state on the ground state subspace, it will not depend on the position of the ϕ field. Nor does it change if we choose a different representative of the chiral field in the same Q class. Namely, let

$$\phi' = \phi + [Q, \rho].$$

Then the projections of $\phi|a\rangle$ and $\phi'|a\rangle$ on the vacuum subspace are the same.

EXERCISE 17.1.1. *Verify the above statements.*

This equality of projections implies that the chiral fields can be used to relate different ground states. In fact it turns out that more is true: All the ground states can be obtained from the operation of chiral fields on some canonical vacuum state that we will now define.

Consider the path-integral on the hemisphere. The boundary of the hemisphere is a circle on which our Hilbert space is based. The path-integral will give us a number, and so defines a functional from boundary field configurations to numbers — equivalently, a state in the Hilbert space. But the standard path-integral (by an argument we do not supply here), gives us a state with anti-periodic boundary conditions i.e., a state in what is called the NS sector. To obtain a state in the Ramond sector, where fermions have periodic boundary conditions, we consider the topologically twisted version of the theory (which amounts to introducing a background gauge field that couples to the R-charge and is equal to the spin connection; the field strength coupling to the fermions in the interior is equivalent to changing the boundary condition of the fermions by a sign). To obtain a ground state at the boundary we consider the "neck" of the hemisphere to be infinitely stretched. In other words we imagine connecting the hemisphere to a semi-infinite flat tube. Note that on the flat tube the twisted and untwisted theories are equivalent. Thus the effect of introducing a factor of an infinitely flat tube on the boundary state $|\psi\rangle$ is to evolve it to $e^{-TH}|\psi\rangle$ as $T \to \infty$, which is equivalent to projecting it to a ground state with $H = 0$.

The state we obtain in this way does not depend on the choice of the precise metric on the hemisphere. The reason for this is that any variation of the metric corresponds to insertion of a Q-trivial operator (recall that this is why the theory is called topological) and this implies that the state $|\psi\rangle$ changes by

$$|\psi\rangle \to |\psi\rangle + Q|\chi\rangle$$

by such a change in metric. Thus e^{-tH} acting on ψ, in the limit of $t \to \infty$, does not change, because $e^{-tH}Q = 0$ as $t \to \infty$ (as the image of Q, which corresponds to states with positive eigenspace for H, is annihiliated by e^{-tH} as $t \to \infty$). The ground state that we obtain in this way, when we insert no operators on the hemisphere, will be denoted by $|0\rangle$. The path-integral thus picks a distinguished element of the Hilbert space.

Similarly, if we consider the topological path-integral together with the insertion of the corresponding chiral fields (i.e., chiral fields for the B-twisting and twisted chiral fields for the A-twisting) we obtain a correspondence between chiral fields and the ground state. For each chiral field ϕ_i we get a ground state $|i\rangle$. In the path-integral language this state is obtained by doing a path-integral on the hemisphere with the chiral operator ϕ_i inserted as shown in Fig. 2.

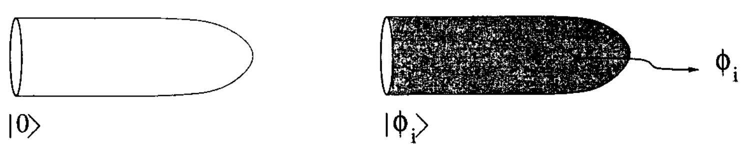


FIGURE 2. The topological path-integral on a hemisphere attached to a semi-infinite tube results in a ground state at the boundary. For each chiral operator ϕ_i we obtain a corresponding ground state $|i\rangle$ given by the path-integral

The state $|i\rangle$ and the state $|0\rangle$ are related. In fact, since changing the position of ϕ_i does not modify the state we obtain (due to the fact that topological observables are position independent) we can consider moving it

to the boundary, which by definition becomes equivalent to the action

$$|i\rangle = \phi_i|0\rangle. \tag{17.5}$$

Note also that this relation implies that the vacuum states provide a realization of the chiral ring:

$$\phi_i|j\rangle = C_{ij}^k|k\rangle,$$

where the above equality holds up to Q-trivial deformations in states and operators.

It is a natural question to ask whether in this way we get a one-to-one correspondence between chiral fields and vacuum states. This is not generally the case for arbitrary $(2,2)$ theories (for example consider topological LG theories with the A-twist), but is the case for the theories we will be mainly considering, for example the LG models with a B-twist, or sigma models with an A-twist. For such cases a non-degenerate pairing, η_{ij}, between the states is defined by the path-integral over the sphere with the corresponding operators inserted on the sphere as shown in Fig. 3.

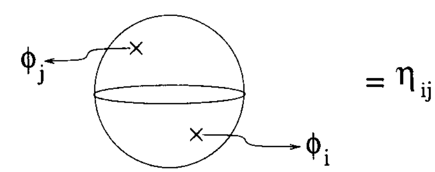


FIGURE 3. An insertion of two chiral fields leads to the definition of the topological metric η_{ij}

Note that this pairing can also be viewed as

$$\eta_{ij} = \langle i|j\rangle,$$

where the state $\langle i|$ corresponds to the state we obtain by applying the topological theory on the hemisphere with the ϕ_i insertion, and $|j\rangle$ is the state we obtain on the boundary of the hemisphere with the ϕ_j insertion. The non-degeneracy of η_{ij} follows from the assumption of the one-to-one correspondence between ground states and chiral fields.

We can also consider the complex conjugate topological twisting, sometimes known as the *anti-topological theory*. In this way we will also obtain a correspondence between anti-chiral fields $\overline{\phi}_i$ and the ground states $|\bar{i}\rangle$. Note that these fields do not correspond to different ground states. Thus there exists an invertible matrix M relating them.

$$|i\rangle = M_i^{\bar{j}}|\bar{j}\rangle. \tag{17.6}$$

Moreover the CPT symmetry of the QFT, i.e., the statement that complex conjugation of all quantities in the path-integral sends the state $|i\rangle$ to the state $|\bar{i}\rangle$, implies that

$$MM^* = 1. \tag{17.7}$$

We can also define a Hermitian matrix ($g_{j\bar{i}} = g^*_{i\bar{j}} = g_{\bar{i}j}$) using the topological and the anti-topological basis,

$$g_{\bar{i}j} := \langle \bar{i}|j\rangle. \tag{17.8}$$

Note that g, η, M are related by

$$M = g^{-1}\eta$$

The moduli space $\mathcal{M}$ in the case of $(2,2)$ theories, corresponding to deformations of the superpotential term and its conjugate, has a natural complex structure. In particular the superpotential parameters t are holomorphic and the ones in the conjugate superpotential $\bar{t}$ are viewed as complex conjugate parameters. Thus the space $\mathcal{M}$ has a natural complex structure. It is natural to ask whether the vacuum bundle V is a holomorphic bundle with a connection compatible with it. We will now demonstrate this by viewing the topological path-integral as defining holomorphic sections of this bundle, and the anti-topological path-integral as defining anti-holomorphic sections. In particular we will now show that $(A_{\bar{i}})^k_j = 0$, which shows that the connection is compatible with the holomorphic structure.

Consider in this basis the components of the connection

$$(A_{\bar{i}})^k_j = \langle k|\overline{\partial}_i|j\rangle. \tag{17.9}$$

We need to show $(A_{\bar{i}})^k_j = 0$. The state $|j\rangle$ is represented by a path-integral with the insertion of the operator ϕ_j. The derivative $\overline{\partial}_i$ acting on $|j\rangle$ is represented by the state obtained by acting with $\overline{\partial}_i$ on the path-integral corresponding to the state $|j\rangle$, which brings down from the action the field

$\int \overline{\Phi}_i d^2\overline{\theta}$):

$$\begin{aligned} \overline{\partial}_i|j\rangle &= \int D\phi \,\left(\int \overline{\Phi}_i d^2\overline{\theta}\right)\, \phi_j\, e^{-S}, \\ &= \int D\phi \,(Q\Lambda)\, \phi_j\, e^{-S}. \end{aligned} \tag{17.10}$$

Here we have used the fact that the $d^2\theta$ integral corresponds to a Q-trivial deformation. As shown in Fig. 4, the overlap of the above state with the state $\langle k|$ is zero, using the Q-symmetry of the right hemisphere path-integral (and noting that ϕ_j is Q-invariant). Since Q can be brought to act on the boundary of the right-hemisphere, it can also be viewed as acting on the state coming from the left hemisphere path-integral. Since the left state is $\langle k|$ and is a ground state, it is annihilated by Q. So we obtain that in the topological basis, $A_{\bar{i}} = 0$. We can thus view the topological basis as the holomorphic basis. The conjugate statement holds for the anti-holomorphic basis and the anti-topological path-integral.

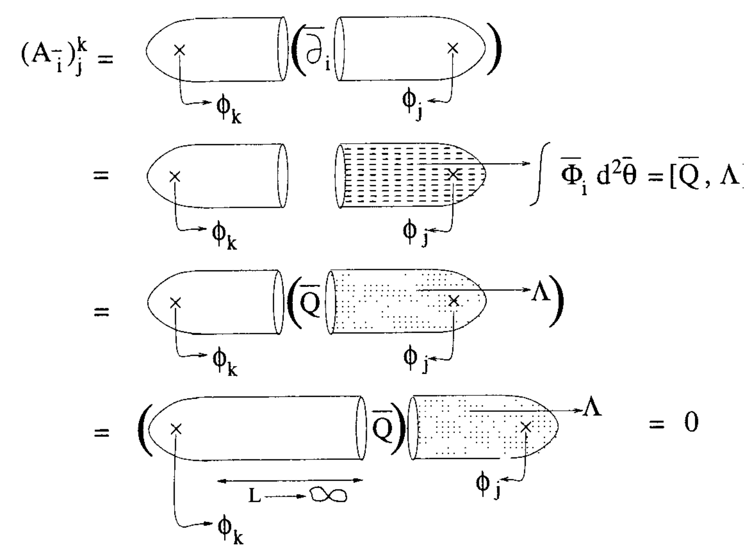


FIGURE 4. The path-integral formulation of the topological theory can be used to show that sections of V defined by chiral ring operators give holomorphic sections of the V-bundle

Note that the three-point function, the chiral ring coefficients, and the two-point function are related as shown in Fig. 5,

$$C_{ijk} = C^l_{ij}\eta_{lk}, \tag{17.11}$$

and are holomorphic in the sense defined above, i.e.,

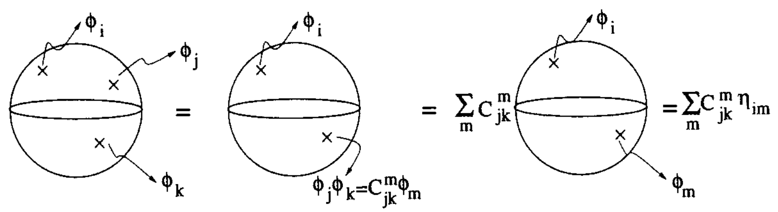


FIGURE 5

$$\overline{\partial}_k\eta_{ij} = 0, \quad \overline{\partial}_l C^k_{ij} = 0. \tag{17.12}$$

This result follows from our discussion of topological field theory amplitudes in the previous chapter.

In the topological basis, the connection and the matrix C_i satisfy the following equations, called the tt^* equations:

$$\begin{gathered} [D_i, D_j] = 0, \\ [\overline{D}_i, \overline{D}_j] = 0, \\ [D_i, C_j] = [D_j, C_i] \qquad [\overline{D}_i, \overline{C}_j] = [\overline{D}_j, \overline{C}_i], \\ [D_i, \overline{D}_j] = -[C_i, \overline{C}_j], \end{gathered} \tag{17.13}$$

where

$$(D_i)^k_j = \delta^k_j\partial_i - (A_i)^k_j, \tag{17.14}$$

$$[D_i, \overline{D}_j] = \partial_i\overline{A}_j - \overline{\partial}_j A_i + [A_i, A_j], \tag{17.15}$$

and C_i denotes the action of the ϕ_i chiral field on the ground states (and similarly for $\overline{C}_i$). The tt^* equations are equivalent to the existence of improved flat connections (which in some geometrical cases are known as the "Gauss–Manin connection")

$$\nabla^\alpha_i = D_i + \alpha C_i, \tag{17.16}$$

$$\overline{\nabla}^\alpha_j = \overline{D}_j + \alpha^{-1}\overline{C}_j, \tag{17.17}$$

where α is an arbitrary constant. The tt^* equations imply that the above *improved* connection is flat,

$$[\nabla^\alpha_i, \nabla^\alpha_j] = [\nabla^\alpha_i, \overline{\nabla}^\alpha_j] = [\overline{\nabla}^\alpha_i, \overline{\nabla}^\alpha_j] = 0. \tag{17.18}$$

Using the fact that the Hermitian metric $g_{i\bar{j}}$ is covariantly constant, i.e., that

$$\partial_k g_{i\bar{j}} = (\partial_k\langle i|)|\bar{j}\rangle + \langle i|(\partial_k|\bar{j}\rangle) \tag{17.19}$$

and the fact that topological and anti-topological theories define holomorphic and anti-holomorphic sections for the V-bundle we see that in the topological (i.e., holomorphic) basis,

$$\partial_k g_{i\bar{j}} = (\partial_k\langle i|)|\bar{j}\rangle \quad \Rightarrow \quad A_i = g^{-1}\partial_i g. \tag{17.20}$$

17.1.1. Proof of tt^* equations. Here we present the proof of most of the tt^* equations (we leave the proof of $[D_i, C_j] = [D_j, C_i] = 0$ and its complex conjugate as an exercise). One can prove these equations in any basis. We will choose the holomorphic basis, i.e., the topological basis, to prove these equations.

$[\overline{D}_i, \overline{D}_j] = 0$**:** We saw earlier that in the topological basis $A_{\bar{i}} = 0$; from which it immediately follows that

$$[\overline{D}_i, \overline{D}_j] = 0. \tag{17.21}$$

The complex conjugate equation follows by considering the anti-topological (i.e., anti-holomorphic) basis.

$[D_i, \overline{D_j}] = -[C_i, \overline{C_j}]$**:** This is one of the most important parts of the tt^* equation, and it relates the curvature of the vacuum bundle with the structure of chiral/anti-chiral rings. We now establish this equation in the holomorphic basis. Since the anti-holomorphic components of the connection vanish, we have

$$\begin{aligned} [D_i, \overline{D}_{\bar{j}}] &= -\overline{\partial}_{\bar{j}} A_i, \\ &= \partial_i A_{\bar{j}} - \overline{\partial}_{\bar{j}} A_i, \end{aligned} \tag{17.22}$$

where we have added the first term on the right-hand side, which is a vanishing quantity in the holomorphic basis, for later convenience. Thus to compute the curvature of the vacuum bundle we need to compute

$$\begin{aligned} \overline{\partial}_{\bar{j}}(A_i)^l_k - \partial_i(A_{\bar{j}})^l_k &= \overline{\partial}_{\bar{j}}\langle k|\partial_i|l\rangle - \partial_i\langle k|\overline{\partial}_{\bar{j}}|l\rangle, \\ &= (\overline{\partial}_{\bar{j}}\langle k|)\partial_i|l\rangle - (\partial_i\langle k|)\overline{\partial}_{\bar{j}}|l\rangle. \end{aligned} \tag{17.23}$$

We continue to use the notation that a state

$$|\Lambda\rangle \tag{17.24}$$

corresponds to the path-integral on the right hemisphere with the operator Λ inserted. Similarly a state

$$\langle\Lambda| \tag{17.25}$$

corresponds to the path-integral on the left hemisphere with the operator Λ inserted. Then

$$\begin{aligned}[D_i, \overline{D}_{\bar{j}}]_k^l &= -(\overline{\partial}_{\bar{j}}\langle k|)\partial_i|l\rangle + (\partial_i\langle k|)\overline{\partial}_{\bar{j}}|l\rangle, \\ &= \langle \phi_k[\overline{Q}_+, [\overline{Q}_-, \Phi_i]] \mid [Q_+, [Q_-, \overline{\Phi}_j]\phi_l\rangle \\ &\quad - \langle \phi_k[Q_+, [Q_-, \overline{\Phi}_j]] \mid [\overline{Q}_+, [\overline{Q}_-, \Phi_i]]\phi_l\rangle,\end{aligned} \tag{17.26}$$

where integration over the positions of $\overline{\Phi}_j$ and Φ_i is implicit. We can move Q_+ and Q_- in either of these two terms to the boundary of the two hemispheres, because we are doing computations in the topological theory and both Q_+ and Q_- are symmetries. Next, we can take them to act on the other hemisphere. In this case it is a symmetry, except where it acts on $[\overline{Q}_+, [\overline{Q}_-, \Phi_i]]$, in which case, by using the SUSY algebra we obtain $\partial\overline{\partial}\Phi_i$. For example, from the second term above we get

$$I_2 = -\langle \phi_k \, \overline{\Phi}_j | \partial\overline{\partial}\Phi_i \; \phi_l\rangle. \tag{17.27}$$

The integral of $\partial\overline{\partial}\Phi_i$ on the right hemisphere (which is implicit in the above formula) is equal to the integral of $\partial_n\Phi_i$ (normal derivative) over the boundary circle, C, of the right hemisphere,

$$I_2 = -\langle \phi_k\overline{\Phi}_j| \oint_C \partial_n\Phi_i \; \phi_l\rangle. \tag{17.28}$$

Since the derivative in the normal direction to the circle C is the generator of time translation we have

$$\partial_n\Phi_i = [H, \Phi_i]. \tag{17.29}$$

Since $|\phi_l\rangle$ is killed by the Hamiltonian H, therefore

$$I_2 = -\langle \phi_k \; \overline{\Phi}_j \mid H \oint_C \Phi_i \; \phi_l\rangle. \tag{17.30}$$

Now divide the left hemisphere into two parts each of which is infinitely long. One part includes the insertion of the field $\overline{\Phi}_j$ over the curved half-sphere. The other part consists of the insertion only in the infinitely long

cylinder. The integral on the first part does not contribute, since the state one propagates infinitely on the second part and therefore is projected to the ground state and is killed by the Hamiltonian H on the circle. Thus only the integral on the infinite cylinder contributes. Let t parametrize the length of the cylinder, with t going from zero to $T \gg 1$. Since the contribution from the first part was to convert the insertion of ϕ_k into a ground state, we get

$$I_2 = -\langle k| \int dt \oint \overline{\Phi}_j(t) \, H \oint \Phi_i|l\rangle, \tag{17.31}$$

where $\oint$ denotes integration along the circle of the cylinder. Since H annihilates $\langle k|$, we can replace H with its commutator with $\oint\overline{\Phi}_j$ to obtain $-\partial_\tau \oint\overline{\Phi}_i$. Thus we can integrate over t and only get contributions from the boundaries $t = 0, T$. The contribution from $t = L$ is cancelled by a similar term, which we get from identical manipulations on the first term of Eq. (17.26). Thus we get, including the contribution from both terms in Eq. (17.26),

$$-\langle k| \oint \Phi_i e^{-TH} \oint \overline{\Phi_j}|l\rangle + \langle k| \oint \overline{\Phi}_j e^{-TH} \oint \Phi_i|l\rangle. \tag{17.32}$$

If we send $T \mapsto \infty$ we will project to the intermediate ground states and we obtain

$$-(C_i\overline{C}_j - \overline{C}_jC_i)_k^l, \tag{17.33}$$

where we have assumed the circumference of the cylinder is 1 — otherwise there would be an extra factor of β^2 in the above equation, where β is the circumference of the cylinder. Thus we get

$$[D_i, \overline{D}_j] = -[C_i, \overline{C}_j], \tag{17.34}$$

which implies that in the topological (holomorphic) basis

$$\begin{aligned}\overline{\partial}_{\bar{j}}A_i &= [C_i, \overline{C}_{\bar{j}}] \\ &= [C_i, g^{-1}C_j^\dagger g]\end{aligned} \tag{17.35}$$

We leave the derivation of the other tt^* equations as an exercise.

EXERCISE 17.1.2. *Consider an LG theory with a single chiral superfield X and with superpotential $W = X^n - \lambda X$. Write equations that determine the Hermitian ground state metric g as a function of λ. (Hint: Use the discrete $\mathbb{Z}_{n-1}$ R-symmetry of this theory to argue for the vanishing of the off-diagonal components of the metric, g. Begin with $n = 3$.)*

In the derivation of tt^* geometry we have only considered the variation with respect to the relevant superpotential terms. It is possible to show that the geometry of the vacuum bundle is independent of the D-terms.

17.1.2. Special Geometry and tt^* Equations. Consider a $(2,2)$ theory corresponding to a sigma model on a Calabi–Yau manifold. We wish to consider the geometry of the vacuum bundle as a function of complex or Kähler deformations. For definiteness let us say we consider the variation with respect to complex structure. What do the above tt^* equations tell us in this case?

In the case of CY, both the axial and vector R-charges are conserved and take integral values. Thus the chiral ring respects a $\mathbb{Z}$-grading. Moreover, the deformations corresponding to complex structure deformations come from fields with left/right R-charge equal to 1. The lowest R-charge state $|0\rangle$, with R-charge 0, is the unique vacuum state corresponding to the identity operator. Note in particular that $g_{0\bar{i}} = 0$ if $i \neq 0$, due to the R-symmetry.

Let us consider the components of Eq. (17.35) in the vacuum-vacuum direction, i.e., the $\overline{0}0$ direction, where $i, \bar{j}$ correspond to moduli of the Calabi–Yau. Using the expression $A_i = g^{-1}\partial_i g$ we get

$$\begin{aligned} \overline{\partial}_{\bar{j}}(g^{-1}_{0\bar{k}}\partial_i g_{k\bar{0}}) &= [C_i, g^{-1}(C_j)^\dagger g]_{0\bar{0}}, \\ \overline{\partial}_{\bar{j}}\partial_i \ln g_{0\bar{0}} &= -\, g^{-1}_{0\bar{0}}(C_j)^{\dagger\, j}_{0}\, g_{i\bar{j}} C^0_{i0}, \\ &= -\,\frac{g_{i\bar{j}}}{g_{0\bar{0}}}, \end{aligned} \tag{17.36}$$

where we used the $\mathbb{Z}$-grading symmetry of R-charge and the fact that multiplication by C_i raises the R-charge by 1.

From the discussion of B-twisted topological theory we know that the identity operator gets mapped to the holomorphic n-form Ω on the CY. Therefore $g_{0\bar{0}}$ is given by

$$\langle \bar{0}|0\rangle = \langle \overline{\Omega} \mid \Omega \rangle = \int \overline{\Omega} \wedge \Omega = e^{-K}, \tag{17.37}$$

where

$$\partial_i \overline{\partial}_j K = \frac{\int \omega_i \wedge \overline{\omega}_j}{\int \Omega \wedge \overline{\Omega}} = \frac{g_{i\bar{j}}}{g_{0\bar{0}}} = G_{i\bar{j}}. \tag{17.38}$$

The metric $G_{i\bar{j}}$ is a Kähler metric on the moduli space corresponding to the Kähler potential K. It is also known as the Weil–Petersson metric on the moduli space of complex deformations of the CY manifold.

CHAPTER 18

BPS Solitons in $\mathcal{N}$=2 Landau–Ginzburg Theories

In the study of Landau–Ginzburg models in two dimensions, we found quantities that depend only on the superpotential term and are independent of the choice of the D-term. For example, we saw that the chiral ring is completely determined by the superpotential terms.

In this section, we will see that the spectrum of ("BPS") solitons is another example of this. This is a beautiful subject in its own right and connects the study of Landau–Ginzburg theories to a branch of mathematics called the Picard–Lefschetz theory of vanishing cycles. The BPS solitons will also turn out to be related to the interpretation of the tt^* geometry discussed before.

The action for a Landau–Ginzburg model of n chiral superfields Φ_i $(i = 1, \ldots, n)$ with superpotential $W(\Phi)$ is given by

$$S = \int d^2x \left[\int d^4\theta\, K(\Phi_i, \overline{\Phi}_i) + \tfrac{1}{2}\left(\int d^2\theta\, W(\Phi_i) + \int d^2\bar{\theta}\, \overline{W}(\overline{\Phi}_i) \right) \right]. \tag{18.1}$$

Here $K(\Phi_i, \bar{\Phi}_i)$ is the Kähler potential that defines the Kähler metric $g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K(\Phi_i, \bar{\Phi}_i)$. If the superpotential $W(\Phi)$ is a quasi-homogeneous function with an isolated critical point (which means $dW = 0$ can only occur at $\Phi_i = 0$) then, as discussed in previous sections, the above action for a particular choice of $K(\Phi, \bar{\Phi})$ is believed to define a superconformal theory. For a general superpotential the vacua are labeled by critical points ϕ_* of W, i.e., where

$$\phi^i(x) = \phi^i_*, \quad \partial_i W|_{\phi_*} = 0 \ \ \forall i. \tag{18.2}$$

The theory is purely massive if all the critical points are isolated and non-degenerate, which means that near the critical points W is quadratic in the fields. We assume this is the case, and label the non-degenerate critical points as $\{\phi_a \mid a = 1, \cdots, N\}$. In such a case, as discussed before, the number of vacua of the theory is equal to the dimension of the local ring of $W(\Phi)$,

$$\mathcal{R} = \frac{\mathbb{C}[\Phi]}{\partial_{\phi^i} W}.$$

Let us take our target space to be $\mathbb{R}$. When we have more than one vacuum, we can have *solitonic* states: at left spatial infinity, $x^1 = -\infty$, the field values are at one vacuum; at right infinity, $x^1 = +\infty$, they are in another vacuum. The topology of the solutions guarantees that they cannot totally disappear (as long as the left and the right infinities are distinct) and so one can look at minimal energy configurations in each topological sector.

Consider a massive Landau–Ginzburg theory with superpotential $W(\Phi_i)$. *Solitons* are static (time-independent) solutions, $\phi^i(x^1)$, of the equations of motion interpolating between *different* vacua i.e., $\phi^i(-\infty) = \phi^i_a$ and $\phi^i(+\infty) = \phi^i_b$, $a \neq b$. The energy of a static field configuration interpolating between two vacua is given by

$$E_{ab} = \int_{-\infty}^{+\infty} dx^1 \left\{ g_{i\bar{j}} \frac{d\phi^i}{dx^1}\frac{d\bar{\phi}^{\bar{i}}}{dx^1} + \frac{1}{4} g^{i\bar{j}} \partial_i W \partial_{\bar{j}} \overline{W} \right\} \tag{18.3}$$

$$= \int_{-\infty}^{+\infty} dx^1 \left| \frac{d\phi^i}{dx^1} - \frac{\alpha}{2} g^{i\bar{j}} \partial_{\bar{j}} \overline{W} \right|^2 + \mathrm{Re}(\bar{\alpha}(W(b) - W(a)) \tag{18.4}$$

where $g_{i\bar{j}} = \partial_i \partial_{\bar{j}} K$ is the Kähler metric and α is an arbitrary phase. The full integrand is independent of α, yet by choosing an appropriate α we can maximize the second term. Since α is a phase, it is clear that the second term is maximal when the phase of $W(b) - W(a)$ is equal to α. Since the first term is non-negative, this implies a lower bound on the energy of the configuration,

$$E_{ab} \geq |W(b) - W(a)|. \tag{18.5}$$

In fact the central charge in the supersymmetry algebra (recall Eqs. (12.78)–(12.79)) in this sector is $(W(b) - W(a))$. "BPS solitons" are solitonic solutions that saturate this bound, and therefore satisfy the equation

$$\frac{d\phi^i}{dx^1} = \frac{\alpha}{2} g^{i\bar{j}} \partial_{\bar{j}} \bar{W}, \quad \alpha = \frac{W(b) - W(a)}{|W(b) - W(a)|}. \tag{18.6}$$

An important consequence of the above equation of motion of a BPS soliton is that along the trajectory of the soliton the superpotential satisfies the equation

$$\partial_{x^1} W = \frac{\alpha}{2} g^{i\bar{j}} \partial_i W \partial_{\bar{j}} \bar{W}. \tag{18.7}$$

Now since the metric $g^{i\bar{j}}$ is positive definite, we know $g^{i\bar{j}} \partial_i W \partial_{\bar{j}} \overline{W}$ is real, and therefore the image under ϕ of the BPS soliton in the W-plane is a straight line connecting the corresponding critical values $W(a)$ and $W(b)$.

EXERCISE 18.0.3. *Consider the $\mathcal{N} = 2$ algebra given in Sec. 12.3. Show that the central term of the algebra can be interepreted as the ΔW in the LG theory. Further, classify the representations of the supersymmetry algebra and show that they are either one-dimensional (corresponding to a vacuum with $E = 0$, two-dimensional, corresponding to BPS solitons where two combinations of supercharges annihilate the state, or otherwise four-dimensional. The two-dimensional represenations are also known as "short multiplets" or "BPS multiplets".*

The number of solitons between two vacua is equal to the number of solutions of Eq. (18.6) satisfying the appropriate boundary conditions. The general way to count the number of solitons will be reviewed in the next subsection. Here we note that for the case of a single chiral superfield the number of solitons between two vacua can also be determined using Eq. (18.7). Since the image of the soliton trajectory is a straight line in the W-plane, by looking at the pre-image of the straight line connecting the corresponding critical values in the W-plane, we can determine the number of solitons between the two vacua. But since the map to the W-plane is many-to-one, not every pre-image of a straight line in the W-plane is a soliton. It is possible for the trajectory to start at a critical point, follow a path whose image is a straight line in the W-plane, and end on a point which is not a critical point but whose image in the W-plane is a critical value. The BPS solitons are those pre-images of the straight line in the W-plane which start and end on the critical points.

18.1. Vanishing Cycles

The soliton numbers also have a topological description in terms of intersection numbers of vanishing cycles. The basic idea is to solve the soliton equation, Eq. (18.6), along all possible directions emanating from one of the critical points. In other words, we study the "wave-front" of all possible solutions to Eq. (18.6).

With no loss of generality we may assume $\alpha = 1$. Near a critical point ϕ_a^i we can choose coordinates u_a^i such that,

$$W(\phi) = W(\phi_a) + \sum_{i=1}^{n}(u_a^i)^2 . \tag{18.8}$$

In this case it is easy to see that the solutions to Eq. (18.6) will have an image in the W-plane which is on a positive real line starting from $W(\phi_a)$. Consider a point w on this line. Then the space of solutions to Eq. (18.6) emanating from $u_a^i = 0$ over this w is a real $(n-1)$-dimensional sphere defined by

$$\sum_{i=1}^{n}(\mathrm{Re}(u_a^i))^2 = w - w_a\,, \quad \mathrm{Im}(u_a^i) = 0 \tag{18.9}$$

where $w_a = W(\phi_a)$. Note that as we take $w \mapsto w_a$ the sphere vanishes. This is the reason for calling these spheres "vanishing cycles". As we move away, the wave-front will no longer be as simple as near the critical point, but nevertheless, over each point w on the positive real line emanating from $w_a = W(\phi_a)$, the pre-image is a real $(n-1)$-dimensional homology cycle Δ_a in the $(n-1)$-dimensional complex manifold defined by $W^{-1}(w)$. Similarly, as we move from w_b towards w_a, there is a cycle Δ_b evolving according to the soliton equation Eq. (18.6) (this would correspond to $\alpha = -1$). Over a common value of w we can compare Δ_a and Δ_b. Solitons originating from ϕ_a and traveling all the way to ϕ_b correspond to the points in the intersection $\Delta_a \cap \Delta_b$. This number, counted with appropriate signs, is the intersection number of the cycles, $\Delta_a \circ \Delta_b$. It turns out that the intersection number counts the number of solitons weighted with $(-1)^F$ for the lowest component of each soliton multiplet, where F is the fermion number. This is independent of deformation of the D-terms. In particular this measures the net number of solitons that cannot disappear by deformations of the D-terms. We will denote this number by A_{ab}, and sometimes loosely refer to it as the number of solitons between a and b. We thus have

$$A_{ab} = \Delta_a \circ \Delta_b . \tag{18.10}$$

Note that to calculate the intersection numbers we have to consider the two cycles Δ_a and Δ_b in the same manifold $W^{-1}(w)$. Since intersection numbers are topological, a continuous deformation does not change them, and hence we can actually calculate them using some deformed path in the W-plane (rather than the straight line) — as long as the path we are choosing is homotopic to the straight line. We are free to vary the path, keeping fixed the homotopy class in the W-plane with the critical values deleted. One way, but not the only way, to transport vanishing cycles along arbitrary paths is to use the soliton equation, Eq. (18.6), but instead of having a fixed α, as would be the case for a straight line, choose α to be $e^{i\theta}$ where θ denotes the varying slope of the path.

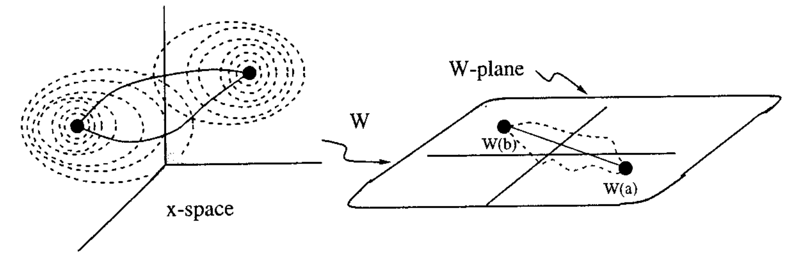


FIGURE 1. BPS soliton map to straight line in the W-plane. Soliton solutions exist for each intersection point of vanishing cycles. Lines in the W-plane that are homotopic to the straight line (dotted lines) can also be used to calculate soliton numbers

Let us fix a point w in the W-plane. For each critical point a of W, we choose an arbitrary path in the W-plane emanating from $W(a)$ and ending on w, but not passing through other critical values. This yields N cycles Δ_a over $W^{-1}(w)$ and it is known that these cycles form a complete basis for the middle-dimensional homology cycles of $W^{-1}(w)$. Hence, if we choose different paths the vanishing cycle we get is a linear combination of the above, and the relation between them is known through the Picard–Lefschetz theory, as we will now review.

18.2. Picard–Lefschetz Monodromy

The basis for the vanishing cycles over each point w in the W-plane depends on the choice of paths connecting it to the critical point. Picard–Lefschetz monodromy relates how the basis changes if we change paths connecting w to the critical values. This is quite important for the study of solitons, and leads to a jump in the soliton numbers. To explain the physical motivation for the question, consider three critical values $W(a), W(b)$

and $W(c)$ depicted in Fig. 2(a), with no other critical values nearby. Suppose we wish to compute the number of solitons between them. According to our discussion above we need to connect the critical values by straight lines in the W-plane and ask about the intersection numbers of the corresponding cycles. As discussed above, due to invariance of intersection numbers under deformation, this is the same as the intersection numbers of the vanishing cycles over the point w connecting to the three critical values as shown in Fig. 2(a). Thus the soliton number is $A_{ab} = \Delta_a \circ \Delta_b$. However, suppose now that we change the superpotential W so that the critical values change according to what is depicted in Fig. 2(b), and that $W(b)$ passes through the straight line connecting $W(a)$ and $W(c)$. In this case, to find the soliton numbers between the a vacuum and the c vacuum, we have to change the homotopy class of the path connecting w to the critical value $W(a)$ as depicted by Fig. 2(b).

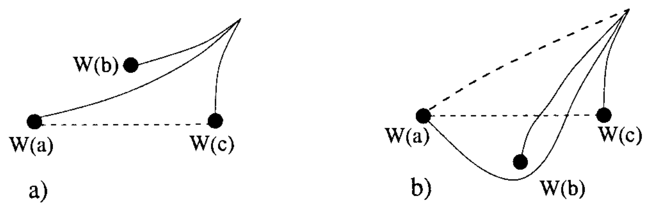


FIGURE 2. As the positions of critical values change in the W-plane, the choice of the vanishing cycles relevant for computing the soliton numbers change

In particular the homology element corresponding to vanishing cycle a changes, $\Delta_a \to \Delta_a'$, and we need to find out how it changes. Picard–Lefschetz theory gives a simple formula for this change. In particular it states that

$$\Delta_a' = \Delta_a \pm (\Delta_a \circ \Delta_b)\Delta_b \,. \tag{18.11}$$

The sign in the above formula is determined once the orientations of the cycles are fixed and will depend on the handedness of the crossing geometry. This is perhaps most familiar in the context of the moduli space of Riemann surfaces, where if we consider a point on the moduli space of Riemann surfaces where a one-cycle shrinks to zero, as we go around this point all the other cycles intersecting it will pick up a monodromy in the class of the

vanishing cycle (the case of the torus and $\tau \to \tau + 1$ is the most familiar case, where the b-cycle undergoes a monodromy $b \to b + a$).

As a consequence of the above formula we can now find how the number of solitons between the a and the c vacuum changes. We simply have to take the inner product $\Delta_a' \circ \Delta_c$ and we find

$$A_{ac}' = A_{ac} \pm A_{ab}A_{bc} \,.$$

18.3. Non-compact n-Cycles

An equivalent description which will be important for later discussion involves defining soliton numbers in terms of the intersection numbers of n-real-dimensional, non-compact cycles, which are closely related to the $(n-1)$-dimensional vanishing cycles we have discussed. The idea is to consider the basis for the vanishing cycles in the limit where the point $w \to e^{i\theta}\infty$. Let us consider the case where $\theta = 0$. In this case we are taking w to go to infinity along the positive real axis. Let us assume that the imaginary parts of the critical values are all distinct. In this case a canonical choice of paths to connect the critical points to w is along straight lines starting from the critical values $W(a)$ stretched parallel to the positive real axis. We denote the corresponding non-compact n-dimensional cycles by γ_a. Then we have

$$W(\gamma_a) = I_a \,, \qquad \text{and} \quad \partial\gamma_a \cong \Delta_a\Big|_{w \mapsto +\infty} , \tag{18.12}$$

where

$$I_a \equiv \{w_a + t \mid t \in [0, \infty)\} \,. \tag{18.13}$$

Two such cycles are shown in Fig. 3.

Let B be the region of $\mathbb{C}^n$ where $\mathrm{Re}\, W$ is larger than a fixed value which is chosen sufficiently large. The non-compact cycles γ_a can be viewed as elements of the homology group $H_n(\mathbb{C}^n, B)$ corresponding to n-cycles with boundary in B, and again it can be shown that they provide a complete basis for such cycles.

For a pair of distinct critical points, a and b, the non-compact cycles γ_a and γ_b do not intersect each other, since their images in the W-plane are parallel to each other (and are separate from each other in the present situation). In this situation we consider deforming the second cycle γ_b so that its image in the W-plane is rotated by an infinitesimally small positive

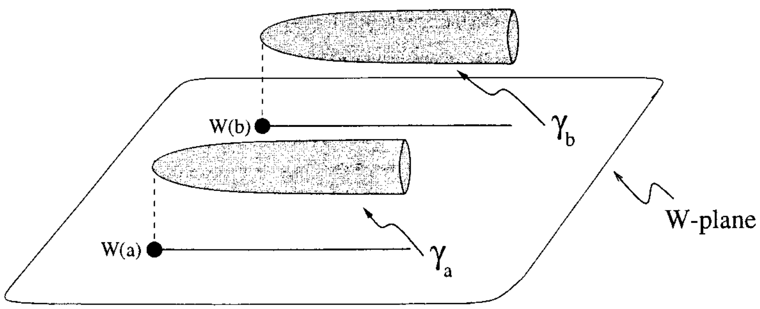


FIGURE 3. The cycles emanating from the critical points. The images in the W-plane are the straight lines emanating from the critical values and extending to infinity in the real positive direction

angle ϵ from the real axis. We denote this deformed cycle by γ_b'. We define the "intersection number" of γ_a and γ_b as the geometric intersection number of γ_a and γ_b'. Depending on whether $\operatorname{Im} W(a)$ is smaller or larger than $\operatorname{Im} W(b)$, the images of γ_a and γ_b' in the W-plane either do or do not intersect each other. In the former case the "intersection number" is zero. In the latter

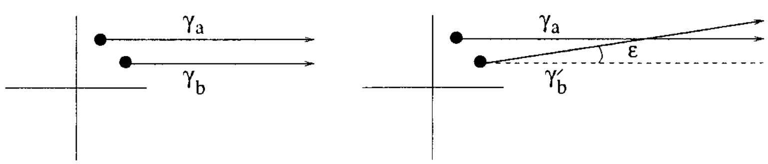


FIGURE 4. The images in the W-plane of γ_a and γ_b (left); and γ_a and γ_b' (right). The second will give rise to an "intersection number." As we will see in the next chapter, this contains certain information on D-branes in the LG model

case, as shown in Fig. 4, the intersection number $\gamma_a \circ \gamma_b'$ is counted by going to the point on the W-plane where their images intersect and asking what is the intersection of the corresponding vanishing cycles $\Delta_a \circ \Delta_b$. Thus the intersection of these n-dimensional cycles contains the information of the soliton numbers. In particular, if there are no extra critical values between the I_a and I_b we will have

$$\gamma_a \circ \gamma_b' = A_{ab}, \quad a \neq b\,. \tag{18.14}$$

If there are extra critical values between I_a and I_b, then these intersection numbers are related to the soliton numbers by the Picard-Lefschetz action as discussed before.

We will see in later chapters that the cycles γ_a defined through parallel transport by the soliton equation, Eq. (18.6), can be viewed as D-branes for LG models that preserve half of the supersymmetries on the worldsheet. There we will also see that the "intersection number" of γ_a and γ_b as defined above can be interpreted as the supersymmetric index for the worldsheet theory of open strings stretched between these cycles.

18.4. Examples

In this section we are going to discuss some examples of soliton numbers in the case of LG models. We will concentrate on LG models representing a class of theories known as $\mathcal{N}=2$ minimal models, as well as the LG models mirror to $\mathbb{P}^N$ sigma models.

Deformed $\mathcal{N}=2$ Minimal models: The k-th minimal model is described by an LG theory with one chiral superfield X with superpotential

$$W(X) = \tfrac{1}{k+2} X^{k+2}\,. \tag{18.15}$$

If we add generic relevant operators to the superpotential, we can deform this theory to a purely massive theory. In this case we will get $k+1$ vacua and we can ask how many solitons we get between each pair. For example, if we consider the (integrable) deformation,

$$W(X) = \tfrac{1}{k+2} X^{k+2} - X, \tag{18.16}$$

then there are $k+1$ vacua that are solutions of $dW = 0$ given by $X = e^{\frac{2\pi i n}{k+1}}, n = 0, \cdots, k$. In this case one can count the preimage of the straight lines in the W-plane and ask which ones connect critical points and in this way compute the number of solitons. It turns out that in this case there is exactly one soliton connecting each pair of critical points.

EXERCISE 18.4.1. *Demonstrate this claim.*

If we deform W, the number of solitons will in general change as discussed above. In this case one can show (by taking proper care of the relevant signs in the soliton number jump) that there is always at most one soliton between vacua. The precise number can be determined starting from the above symmetric configuration. The analogue of the non-compact one-cycles γ_i

in this case will be discussed in more detail later on, after we discuss their relevance as D-branes. They are cycles in the X-plane, asymptotic to a $(k+2)$-th root of unity as $X \to \infty$. That there are $k+1$ inequivalent such homology classes for $H_1(\mathbb{C}, \mathrm{Re}\, W = \infty)$ is related to the fact that there are $k+1$ such classes defined by γ's up to linear combinations.

$\mathbb{P}^{N-1}$: We next consider the $\mathbb{P}^{N-1}$ sigma model. We will use an equivalent LG description of it. That there is an equivalent LG description will be demonstrated later, when we prove mirror symmetry. The soliton matrix of the non-linear sigma model with target space $\mathbb{P}^{N-1}$ can be computed directly by studying the tt^* equations. The mirror LG theory provides a simple way of calculating the soliton matrix. We start with the case $N=2$, where we can present explicit solutions to the soliton equation.

The Landau–Ginzburg theory, which is mirror to the non-linear sigma model with $\mathbb{P}^1$ target space, is the so-called $\mathcal{N} = 2$ sine-Gordon model defined by the superpotential

$$W(x) = x + \frac{\lambda}{x}. \tag{18.17}$$

Here $x = e^{-y}$ is a single-valued coordinate of the cylinder $\mathbb{C}^\times$ and $-\log\lambda$ corresponds to the Kähler parameter of $\mathbb{P}^1$. The critical points are $x_*^\pm = \pm\sqrt{\lambda}$ with critical values $w_*^\pm = \pm 2\sqrt{\lambda}$. As mentioned in the previous section the BPS solitons are trajectories, $x(t)$, starting and ending on the critical points such that their image in the W-plane is a straight line,

$$x(t) + \frac{\lambda}{x(t)} = 2\sqrt{\lambda}(2t-1)\,, \quad t \in [0,1]\,. \tag{18.18}$$

This is a quadratic equation with two solutions given by,

$$x(t)_\pm = \sqrt{\lambda}(2t-1) \pm 2i\sqrt{\lambda}\sqrt{t-t^2} = \sqrt{\lambda}e^{\pm i \tan^{-1}\frac{2\sqrt{t-t^2}}{2t-1}}. \tag{18.19}$$

Since $x_+(t) = x_-(t)^*$ and $|x_+(t)| = |\sqrt{\lambda}|$, there are two solitons between the two vacua such that their trajectories in the x-plane lie on two half-circles, as shown in Fig. 5(a). Since x is a $\mathbb{C}^\times$ coordinate we can consider the x-plane as a cylinder. Soliton trajectories on the cylinder are shown in Fig. 5(b). This description is useful in determining the intersection numbers of middle-dimensional cycles. As described in the previous section the number of solitons between two critical points is given by the intersection number of middle-dimensional cycles starting from the critical points. In our case there are two such cycles that are the preimages of two semi-infinite lines in the

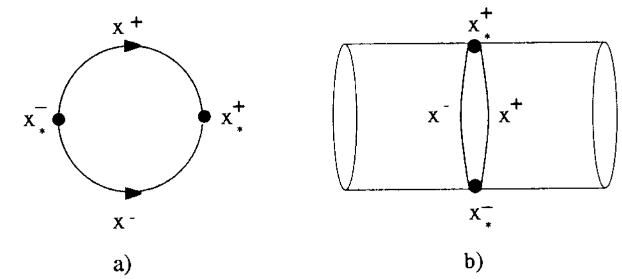


FIGURE 5. The two solitons of the $\mathbb{P}^1$ model

W-plane starting at the critical values as shown in Fig. 6(a). The preimage of these cycles on the cylinder is shown in Fig. 6(b). The cycles in the x-space intersect only if the lines in the W-plane intersect each other and the intersection number in this case is 2.

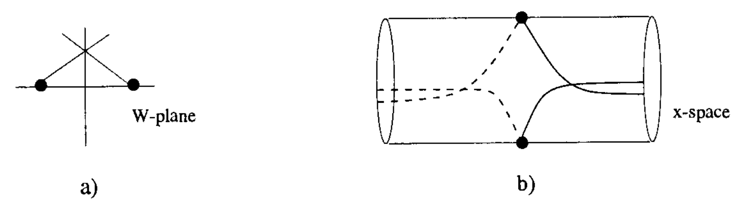


FIGURE 6. Intersecting lines in the W-plane and the corresponding intersecting cycles in the x-space

We now turn to the study of solitons of the $\mathbb{P}^{N-1}$ sigma model. The LG theory mirror to the non-linear sigma model with $\mathbb{P}^{N-1}$ target space has superpotential

$$W(X) = \sum_{k=1}^{N-1} X_k + \frac{\lambda}{X_1 \cdots X_{N-1}}\,. \tag{18.20}$$

This superpotential has N critical points given by

$$X_i^{(a)} = e^{\frac{2\pi i a}{N}} \quad i = 1, \cdots, N-1; \quad a = 0, \cdots, N-1, \tag{18.21}$$

with the critical values (when $\lambda = 1$)

$$w_a \equiv W(\vec{X}^{(a)}) = N e^{\frac{2\pi i a}{N}}\,. \tag{18.22}$$

Here, unlike the previous case of $\mathbb{P}^1$, to be able to solve for the preimage of a straight line, we will make an assumption about the soliton solution.

We assume that the soliton trajectory is determined by a function $f(t)$ such that

(18.23)
$$X_1 = X_2 = \cdots = X_k = f(t)^{N-k}\,,\ X_{k+1} = X_{k+2} = \cdots = X_N = f(t)^{-k}\,.$$

This parametrization of the solution satisfies the constraint $\prod_{i=1}^N X_i = 1$ by construction. With this ansatz, the straight line equation in the W-plane becomes (for λ=1)

$$P(f) := kf^{N-k} + (N-k)f^{-k} = N(1 - t + te^{\frac{2\pi ik}{N}})\,, \tag{18.24}$$

where the right-hand side is the straight line $w(t)$ starting from $w(0) = N$ and ending on $w(1) = Ne^{\frac{2\pi ik}{N}}$. Here we have chosen the parameter t running in the range $[0,1]$ that is linear in the W-plane. We are interested in the solutions that start at $t = 0$ from $X_i^{(0)}$ and end at $t = 1$ on $X_i^{(k)}$. This implies that $f(0)^{N-k} = f(0)^{-k} = 1$ and $f(1)^{N-k} = f(1)^{-k} = e^{\frac{2\pi ik}{N}}$. Thus the number of solitons that satisfy Eq. (18.23) is given by the number of solutions to Eq. (18.24) such that $f(0) = 1$ and $f(1) = e^{-\frac{2\pi i}{N}}$. We will show that there is only a single solution that satisfies these conditions.

Since $P'(1) = 0$ and $P''(1) \neq 0$, where prime denotes a differentiation with respect to f, only two trajectories start from $f = 1$. Thus it follows that the number of solutions is less than or equal to 2. From Eq. (18.24) it is clear that f can be real only at $t = 0$. Thus a trajectory cannot cross the real axis for $t > 0$. For t very close to zero one of the trajectories moves into the upper half-plane. Since the trajectory in the upper half-plane cannot cross the real axis it cannot end on $e^{-\frac{2\pi ik}{N}}$. Thus there can be at most one solution.

To show that there actually exists a solution we will construct a solution whose image in the W-plane is homotopic to the straight line $w(t)$. Consider the function $f_*(t) = e^{-\frac{2\pi i}{N}t}$ where $t \in [0,1]$. Since

$$|P(f_*(t))| = |ke^{-2\pi it} + (N-k)| \leq |ke^{-2\pi it}| + (N-k) = N\,, \tag{18.25}$$

the image of $f_*(t)$ in the W-plane always lies inside the circle of radius N and only intersects the circle for $t = 0$ and $t = 1$ at $w = w_0$ and $w = w_k$, respectively. Thus the image is homotopic to the straight line $w(t)$ and therefore there exists a solution $f_0(t)$ homotopic to $f_*(t)$ with the required properties.

Since permuting the N coordinates among themselves does not change the superpotential, it follows that we can choose any k coordinates to be equal to f^{N-k} and the remaining $(N-k)$ coordinates equal to f^{-k}. Thus we see that there are $\binom{N}{k}$ solitons between the critical points $X_i^{(0)}$ and $X_i^{(k)}$ consistent with the ansatz of Eq. (18.23). The case for $k = 1$ when $N \gg 1$ was already discussed in the section on linear sigma models, and the above result is consistent with and generalizes this discussion. Note that if the $\mathbb{P}^{N-1}$ has a round metric having $SU(N)$ symmetry, then the solitons should form representations of this group. In fact, the permutations of X_i can be viewed as the Weyl group of the $SU(N)$. It thus follows, given how the permutations act on the solutions, that in this case the solitons connecting vacua k units apart correspond to the k-fold anti-symmetric tensor product of the fundamental representation of $SU(N)$.

18.5. Relation Between tt^* Geometry and BPS Solitons

The two objects we have defined for LG theories — the solutions to tt^* equations and the spectrum of BPS solitons — are not unrelated. The relation turns out to be the following. Let the worldsheet be given by an infinite cylinder with circumference β. Consider the operator formulation of the LG theory on the real line, viewing the circumference of the cylinder as the Euclidean time direction. Consider periodic boundary conditions for fermions around the circumference. Define

$$Q = \frac{\beta}{L}\mathrm{Tr}(-1)^F F\exp(-\beta H)$$

where L is the length of the "real line" (what we mean by this is that F/L is simply the local density of the fermion number inserted at any point along the real line). Note that Q is a matrix with indices describing which vacua one ends up with at left and right infinity. One can show that only BPS configurations can contribute to the above expression.

EXERCISE 18.5.1. *Show that, at least formally, the non-reduced multiplets of the $\mathcal{N} = 2$ algebra (i.e., the four-dimensional ones) do not contribute to the above trace, and only the BPS, or "reduced multiplets" (i.e., the two-dimensional ones) can.*

The above statement is essentially true, but it turns out that (due to an anomaly) a combination of BPS solitons can also contribute to the above trace, and thus the computation turns out to be rather non-trivial. It is thus very interesting that the same quantity can be captured by solutions to tt^* equations, as we will now discuss.

Consider the tt^* connection along the one-parameter deformation of the superpotential W, given by

$$W \to e^{-\tau} W$$

where one considers tt^* geometry on a circle of circumference β (which effectively is the same as considering $W \to \beta W$ on a circle of circumference 1). Let us denote the corresponding connection by A_τ. Then it turns out (by a canonical choice of gauge) that one has

$$Q = A_\tau.$$

We will not present the proof of this statement here. It is quite satisfying to see a relation between a Hilbert space computation and objects appearing in tt^* geometry. $Q = \mathrm{Tr}\,(-1)^F F e^{-\beta H}$ is a kind of an index generalizing the $\mathrm{Tr}\,(-1)^F e^{-\beta H}$ index for general supersymmetric theories, which captures the BPS content of the supersymmetric theory, just as $\mathrm{Tr}\,(-1)^F$ captures the ground state content of the supersymmetric theory.

CHAPTER 19

D-branes

One important piece of the mirror symmetry story that we have not discussed yet is D-branes. D-branes not only deepen our understanding of mirror symmetry, they also help us grasp the meaning of topological string amplitudes from the viewpoint of target space physics. In this section, we develop some basic aspects of D-branes. More details, especially in the context of fermionic fields, will appear in Ch. 39.

19.1. What are D-branes?

We have considered (bosonic) sigma models of maps from Riemann surfaces without boundaries to target spaces. It is natural in this context to ask: what if we have Riemannn surfaces with boundaries, with some natural boundary conditions?

Consider a sigma model of maps from the cylinder $\Sigma = S^1 \times \mathbb{R}$ to $\mathbb{R}$ with (Euclidean) action

$$S = \int \partial_\mu \phi\, \partial^\mu \phi d^2 x.$$

The classical equation of motion, which is obtained by setting to zero the variation of the action ($\delta S = 0$) with respect to arbitrary variations of the field ϕ, is

$$\partial_\mu \partial^\mu \phi = 0.$$

However this assumes there are no boundary terms generated by varying the field. The contribution of the boundary to the variation is given by

$$\delta\phi\, \partial_n \phi|_{boundary} = 0,$$

where $\partial_n \phi$ is the normal derivative of ϕ at the boundary.

EXERCISE 19.1.1. *Verify that the variation of the action gives rise to the above boundary term.*

We would like to set this boundary term to zero. There are two natural ways of doing this:

$$\text{Neumann (N)} : \partial_n\phi|_{\partial\Sigma} = 0, \tag{19.1}$$

$$\text{Dirichlet (D)} : \delta\phi|_{\partial\Sigma} = 0. \tag{19.2}$$

In the Dirichlet case, the image of the boundary $\partial\Sigma$ is a point in the target space ($\mathbb{R}$ in this case) — we will call this a D0-brane. In the case of Neumann boundary conditions, the worldsheet boundary can be at any point in the target — we will say in this case that there is a D1-brane stretched along the real line $\mathbb{R}$.

We can write the Dirichlet (D) and Neumann (N) boundary conditions more symmetrically as

$$\mathrm{N} : \partial_n\phi|_{\partial\Sigma} = (\partial\phi - \overline{\partial}\phi)|_{\partial\Sigma} = *d\phi_{\partial\Sigma} = 0,$$

and

$$\mathrm{D} : d\phi|_{\partial\Sigma} = (\partial\phi + \overline{\partial}\phi)|_{\partial\Sigma} = 0.$$

The terminology in general is as follows: Consider a p-dimensional subspace N^p of the target space and restrict the boundary of the Riemann surface to map to it. Moreover we require Neumann boundary conditions for directions *normal* to the space N^p. In such a situation, we say that we have a "Dp-brane wrapping the subspace N^p of the target space." In general we may have many different D-branes and we can consider Riemann surfaces with more than one boundary, where different boundaries are mapped to different D-branes.

Let us now consider the target space being a circle S^1 of radius R. We recall from Sec. 11.2 the T-duality symmetry which relates $R \to 1/R$ symmetry, and ask how the D-branes, i.e., the D0- and D1-branes, get identified under this symmetry.

Recall from our discussion of the $R \to 1/R$ duality that this has the effect

$$\partial\phi \to \partial\widetilde{\phi}, \tag{19.3}$$

$$\overline{\partial}\phi \to -\overline{\partial}\widetilde{\phi}, \tag{19.4}$$

where $\widetilde{\phi}$ is a coordinate on the dual circle. We can see, therefore, that when the worldsheet has boundaries, this symmetry interchanges Neumann and Dirichlet boundary conditions. In other words, the $R \to \frac{1}{R}$ symmetry induces an action on D-branes exchanging D0-branes with D1-branes.

So far we have talked about bosonic sigma models. A similar story repeats for the fermionic sigma model, and the worldsheet supersymmetry will dictate what the appropriate boundary conditions on the fermions are. We can then ask if the D-brane boundary conditions preserve all the supersymmetries of the world-sheet theory, and the answer is no: the D-brane can preserve only half of the supersymmetries. We saw, in our discussion of (2,2) supersymmetry, that there were four combinations of supercharges: $Q_A = Q_- + \overline{Q_+}$, $Q_B = \overline{Q_-} + \overline{Q_+}$, and their complex conjugates $\overline{Q}_A$, $\overline{Q}_B$. The A-model supercharges $Q_A, \overline{Q}_A$, are preserved when the D-brane is a Lagrangian submanifold of the Kähler target space. The B-model supercharges are preserved when the D-brane is a holomorphic submanifold to preserve the corresponding supercharges. (Note that we are talking about world-sheet supersymmetry, not supersymmetry in space-time). This will be discussed in Ch. 37 in detail.

Let us now recall our first example of mirror symmetry, which is the supersymmetric sigma model with target space the flat torus

$$T = S^1_{R_1} \times S^1_{R_2}.$$

The mirror is the torus

$$T' = S^1_{1/R_1} \times S^1_{R_2}.$$

Now a D0-brane at a point on T corresponds, in the dual theory, to a D1-brane wrapping the first S^1 in T' (see Fig. 1, (a)).

If we had started with a D2-brane wrapping T, we would have ended up with a D1-brane on the mirror, this time wrapped on the second S^1 in T' (see Fig. 1, (b)). In general, we expect mirror symmetry at the level of cohomology elements realized by chiral fields act by the reflection $h^{p,q} \leftrightarrow h^{d-p,q}$. The action of mirror symmetry in this example is providing a concrete integral homology realization of this map ($d = 1$ here), realized through the D-branes. More generally, for a Calabi–Yau d-fold, one expects that D-branes represented by Lagrangian real d-dimensional spaces will be mapped to holomorphic objects of all possible complex dimensions by the mirror map.

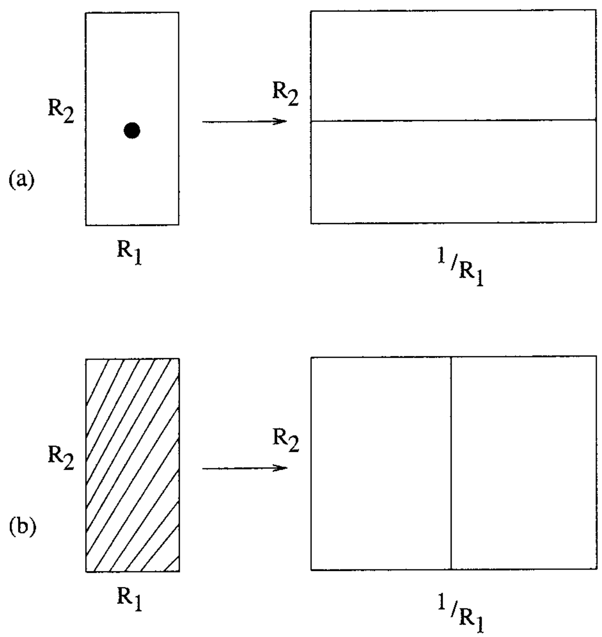


FIGURE 1. Mirror symmetry for sigma model on flat torus

19.2. Connections Supported on D-branes

We saw that $R \to 1/R$ interchanged D0- and D1-branes on the circle. So if we start with a D1-brane wrapping the circle, we end up in the dual description with a D0-brane localized at a point on the dual circle. There seems to be a contradiction. We can change the position of the D0-brane, so there is a one-dimensional moduli space of choices for the D0-brane. What about the D1-brane? It seems to have no moduli! So how could the two objects become equivalent under T-duality? The answer turns out to be that on the D-brane there lives a rank 1 bundle with connection, and that turns out to have moduli in the case of a D1-brane.

Recall that we modified the sigma model by introducing an integral two-form $B \in H^2(M, \mathbb{Z})$ in the target space and modifying the path-integral by the phase

$$\exp(2\pi i \int_\Sigma \phi^* B).$$

This makes sense on worldsheets without boundary, but we can see that there is going to be a subtlety when we allow worldsheets with boundary: Under $B \to B + d\Lambda$,

$$\int_\Sigma \phi^* B \to \int_\Sigma \phi^*(B + d\Lambda) = \int_\Sigma \phi^* B + \int_{\partial\Sigma} \phi^* \Lambda.$$

So this pairing will not be well defined. We can compensate for this shift by introducing a one-form A (connection) on the D-brane and modifying the action by

$$S \to S - 2\pi i \int_{\partial\Sigma} \phi^* A.$$

We see then that under the combined transformation

$$(B, A) \to (B + d\Lambda, A + \Lambda),$$

the path-integral is invariant. In other words, the data of a D-brane includes a $U(1)$ connection (not necessarily flat) on the D-brane. More generally, we could put several (say n) D-branes on top of each other. In this case, the n $U(1)$ bundles get enhanced to a $U(n)$ bundle.[1] The path-integral modification in this case is

$$\operatorname{Tr} P\exp(-2\pi i \oint A)$$

(i.e., the path-ordered exponentiation of the connection, which gives the holonomy) and the B field mixes only with the diagonal $U(1)$ subgroup of the $U(n)$. Note that in case $A = 0$ this corresponds to putting an extra factor of n for each hole. In other words n identical D-branes with no connection turned on affects the worldsheet theory by associating a factor of n for each hole mapped to it.

Now we come back to the question of where on the dual circle the D0-brane sits. The D1-brane wraps the circle and, as we have just seen, has a $U(1)$ connection on it. The moduli space of flat connections on the D1-brane is, in this example, the dual S^1, so specification of the connection on the D1-brane is equivalent to the specification of a point on the dual circle, which is the point where the D0-brane sits! This restores the symmetry between the corresponding moduli spaces of D-branes expected from mirror symmetry considerations. Our discussion of D-branes in this case suggests that the study of moduli spaces of D-branes should be very relevant for the

[1] Massive strings stretching between the two branes become massless, thus filling out the off-diagonal parts of a $U(n)$ connection.

study of mirror symmetry. Aspects of this will be discussed in more detail later in Ch. 37.

19.3. D-branes, States and Periods

From the path-integral point of view, a D-brane is a specification of boundary conditions on the fields. The path-integral viewpoint then suggests that it can also be viewed as a *state*. Consider the worldsheet with the topology of a semi-infinite cylinder, and put appropriate D-brane boundary conditions on the boundary circle. Viewing the time evolution along the semi-infinite length of the cylinder, the D-brane boundary condition can also be viewed in the operator formulation as specifying a (generally non-normalizable) state in the Hilbert space on the circle. If we consider propagating this state along the semi-infinite length of the cylinder, we get a projection of this D-brane state to a ground state of the supersymmetric QFT. In fact, the labeling of the ground states via D-brane states is topological as it does not refer to any moduli (such as Kähler or complex structure) of the target space.

Previously we have noted a labeling of the ground states of the QFT by chiral (and twisted chiral) ring elements, as in our discussion of tt^* equations. It is natural to ask how this new way of labeling ground states, via D-branes, fits with the previous constructions and how this is related to geometric aspects of Calabi–Yau.[2] More specifically we consider the case where the D-brane is given by a Lagrangian middle-dimensional cycle in the Calabi–Yau and ask about the pairing of the corresponding boundary state with the ground states labelled by chiral fields. For example, we would like to know the overlap between the distinguished ground state $|0\rangle$, corresponding to the identity operator, which is the state with the lowest R-charge, and the D-brane boundary state $\langle\gamma|$, corresponding to a Lagrangian D-brane, i.e.,

$$\langle\gamma|0\rangle = ?$$

EXERCISE 19.3.1. *Show that this pairing will only depend on the homology class of the D-brane.*

[2]This relation can also be considered in the more general context of Kähler manifolds, but for simplicity here we restrict our attention to the Calabi–Yau case.

On the mathematical side, we discussed in Ch. 6 the pairing between homology and cohomology given by the periods $\int_{A_i}\Omega$ and $\int_{B^i}\Omega$, where A_i and B^i are middle-dimensional (integral) homology cycle and Ω is the holomorphic n-form. One would naturally expect that the above question for the overlap of the Lagrangian D-branes and the periods of the holomorphic n-form are related. We will now argue this is indeed the case.

How is this pairing realized physically? From the physical viewpoint, we saw in Ch. 17 that Ω corresponds to the identity operator in the chiral ring. Now we will see that the D-branes provide the integral structure corresponding to the cycles A_i, B^i, and that the pairing is realized in terms of overlaps of the D-brane boundary states corresponding to the middle-dimensional homology cycles and the field (state) corresponding to the identity operator of the chiral ring.

To formulate this as a path-integral computation we consider the topological B-model on a Calabi–Yau n-fold on a semi-infinite cigar. Recall that the observables of the topological theory are labelled by $H^p(\wedge^q TM)$, which can be identified with $H^{p,n-q}(M)$ through contraction of indices with Ω. The path-integral representation of $\int_{A_i}\Omega$ is given by the semi-infinite cigar with the state corresponding to the A_i D-brane at the boundary, as we will now argue.

We noted earlier that a D-brane preserves the supersymmetries of the B-model when it is a holomorphic submanifold. Now we are finding that the D-brane needs to be a middle-dimensional cycle, in order for the pairing with Ω to make sense, and in particular preserves the A-model supersymmetries. This means that the B-model supercharges are not preserved by this D-brane, i.e, that, after twisting, the amplitude on the disk with no insertion and boundary conditions corresponding to the D-brane on Lagrangian D-branes is *not* a purely topological amplitude. However, we are interested in the overlap of the (canonical) ground state with the D-brane boundary state. The path-integral on the disk with no operator insertions gives a state at the boundary which is Q-cohomologically equivalent to this ground state. To get the actual ground state, we evolve this state along an infinite tube to project out all components except the ground state. Then take the overlap with the D-brane boundary state. The result is a path-integral on a semi-infinite cigar with no operator insertions and with the boundary conditions on the terminal circle corresponding to the D-brane.

Since the circumference of the circle direction of the cigar is *irrelevant* in the case of the sigma model on the Calabi–Yau manifold (since the two-dimensional theory is conformal), we can take an infinitesimally thin, infinitely long cigar — this allows us to dimensionally reduce the problem to quantum mechanics and compute the overlap as the integral of the ground state wave-function over the delta-function constraint specified by the D-brane boundary conidition. Since the state corresponding to the identity operator in the B-model is realized by the holomorphic n-form, this overlap integral is thus realized through the integral

$$Z_i = \langle A_i|0\rangle = \int_{A_i} \Omega,$$

$$F_i = \langle B_i|0\rangle = \int_{B_i} \Omega,$$

which is what we wished to establish.

Recall that the topological B-model could be defined by twisting either a Calabi–Yau σ model or a Landau–Ginzburg theory. We also saw that some Calabi–Yau σ models (for some values of moduli) admit a Landau-Ginzburg description. It is then natural to ask what the analogue of D-branes, D-brane states, and their overlap, with topological ground states are in the Landau–Ginzburg case.

The analogue of D-branes for Landau–Ginzburg theories turns out to be the non-compact middle-dimensional Lagrangian cycles we defined in the context of studying solitons of Landau–Ginzburg theories in Ch. 18. They are the lifts of the straight line images in the W plane to the field space, emanating from critical points, using the soliton equations. In fact they naturally pair up with the chiral ring elements. Recall that Landau–Ginzburg theories have a chiral ring given by $\mathcal{R} = \mathbb{C}[\phi_i]/(\partial W)$ which is the Landau–Ginzburg analogue of the ring $H^p(\wedge^q(TM))$ in the case of the Calabi–Yau. Assuming that W has isolated singularities, the (relative) homology group $H_n(\mathbb{C}^n, \{\mathrm{Re}\, W \to \infty\})$ has the same dimension as the chiral ring $\mathcal{R}$, and there is a natural pairing between them. The overlap integral of the D-brane states with the vacua are a natural analogue of the pairing between the holomorphic n-form and the Lagrangian cycles of the Calabi–Yau. Namely

$$\langle\gamma|0\rangle = \Pi_\gamma = \int_\gamma e^{-W} d\phi_1 \cdots d\phi_n. \tag{19.5}$$

Note that this is a well-defined integral because the cycles γ are defined by the condition that in the non-compact direction $\mathrm{Re}\, W \to \infty$. Note also that in the case of vanishing superpotential, i.e., the sigma model on the n-dimensional complex plane, the result reduces to the period pairing we have discussed in the context of Calabi–Yau manifolds. One can also consider periods of the form

$$\langle\gamma|\varphi_\alpha\rangle = \Pi_\gamma^\alpha = \int_\gamma \varphi_\alpha e^{-W} d\phi_1 \cdots d\phi_n$$

where φ_α are chiral fields. Actually, the above identity holds only for a special representative of chiral fields φ_α corresponding to topological "flat coordinates". Even though this identity can be derived with some work (by going over to supersymmetric quantum mechanics), we will limit ourselves here to providing evidence for this formula.

Note that the right-hand side of the above identity, which is a weighted period integral, is given by derivatives of the fundamental period Π_γ with respect to the moduli (as in the Calabi–Yau case):

$$\Pi_\gamma^\alpha = \frac{\partial \Pi_\gamma}{\partial t_i}$$

where the superpotential involves couplings $W = \int d^2\theta t_\alpha \varphi_\alpha$. ($\varphi_\alpha$ will be functions of the ϕ_i.)

EXERCISE 19.3.2. *In the non-conformal case (Landau–Ginzburg superpotential that is not quasi-homogeneous) the overlaps of D-brane states with ground states characterized by chiral ring elements do depend on the circumference of the circle. Show (assuming we have a complete basis for ground states specified by some collection of D-branes) that the solution to the tt^* equations can be written in terms of such overlaps (which in general will be very non-trivial functions of the moduli of the theory). Note that in the limit of infinitesimal circumference the overlap will agree with the above truncation to a finite-dimensional integral. (Hint: $g_{i\bar{j}} = \langle\overline{\phi_j}|\phi_i\rangle = \langle\overline{\phi_j}|\gamma\rangle C^{\gamma\gamma'}\langle\gamma'|\phi_i\rangle$ for some suitable intersection matrix C)*

EXERCISE 19.3.3. *For the conformal case of the Calabi–Yau, rederive from the above result the relation between the Kähler potential on moduli space $e^{-K} = \langle\overline{0}|0\rangle$ and the period integral.*

We noted before that when the superpotential W is quasi-homogeneous, in some cases the (orbifold of) Landau–Ginzburg theory corresponds to a

Calabi–Yau sigma model — so our expression for the periods should reduce to the periods on a Calabi–Yau. Let us see how this works in an example.

Consider the Landau–Ginzburg theory given by the quasi-homogeneous

$$W = x_1^n + \cdots + x_n^n + \psi x_1 \cdots x_n$$

mod $\mathbb{Z}_n$ acting as diagonal phase multiplication on all fields. This Landau–Ginzburg theory corresponds to the Calabi-Yau $(n-2)$-fold given by the equation $W = 0$ in $\mathbb{CP}^{n-1}$ (for Kähler moduli $r \to -\infty$). Consider the period

(19.6)

$$\begin{aligned}\Pi &= \int_\gamma \exp[-(x_1^n + \cdots + x_n^n + \psi x_1 \cdots x_n)] dx_1 \cdots dx_n \\ &= \int_\gamma \exp[-x_1^n(1 + (\frac{x_2}{x_1})^n + \cdots + (\frac{x_n}{x_1})^n + \psi \frac{x_2}{x_1} \cdots \frac{x_n}{x_1})] dx_1 \cdots dx_n.\end{aligned}$$

Defining $\zeta_i = x_i/x_1$ for $i \neq 1$ and $\zeta_1 = x_1^n$, we have

$$\begin{aligned}\Pi &= \int_\gamma \exp[-\zeta_1(1 + \zeta_2^n + \cdots + \zeta_n^n + \psi\zeta_2 \cdots \zeta_n)] d\zeta_1 \cdots d\zeta_n \\ &= \int_\gamma \delta(1 + \zeta_2^n + \cdots + \zeta_n^n + \psi \zeta_2 \cdots \zeta_n) d\zeta_2 \cdots d\zeta_n \\ &= \int_\gamma \frac{d\zeta_2 \cdots d\zeta_n}{\frac{\partial W}{\partial \zeta_2}|_{f=0}},\end{aligned} \tag{19.7}$$

where $f = 1 + \zeta_2^n + \cdots + \zeta_n^n + \psi\zeta_2 \cdots \zeta_n$. This is exactly the integral of the holomorphic $(n-2)$-form on the Calabi–Yau $(n-2)$-fold in the patch $\zeta_1 \neq 0$. This connects the Landau–Ginzburg computation to that expected for the Calabi–Yau case. Even the fact that we have to consider the orbifold theory is needed for this correspondence: Note that when we made the change of variables $\zeta_1 = x_1^n$, $\zeta_2 = \frac{x_2}{x_1}$, $\cdots$, $\zeta_n = \frac{x_n}{x_1}$, this change of variables has a Jacobian of unity (precisely when W is quasi-homogeneous). Furthermore the change of variables is n-to-1, which corresponds to modding out by the $\mathbb{Z}_n$ action on x_i identifying them with an overall $\mathbb{Z}_n$ phase rotation. Alternatively, one could derive Picard-Fuchs equations from the Landau–Ginzburg expression for the periods and check that they are the same equations as arise in the corresponding Calabi–Yau.

EXERCISE 19.3.4. *Verify this for the case of the quintic threefold corresponding to setting* $n = 5$ *for the above Landau–Ginzburg theory.*

The difference between these two different expressions for the periods is that the Landau–Ginzburg periods are expressed as integrals over non-compact cycles γ in flat space ($\mathbb{C}^n$), which makes them potentially easier to compute. We can see that these period integrals are independent of the Kähler moduli, because the B-model parameters (complex structure moduli) are decoupled from the A-model parameters (Kähler moduli).

The metric on the Calabi–Yau moduli space is given by (recalling our discussion of tt^* geometry) $G_{i\bar{j}} = g_{i\bar{j}}/g_{0\bar{0}}$ where, as follows from the exercise,

$$g_{i\bar{j}} = \Pi_\gamma^{\bar{j}} C^{\gamma\gamma'} \Pi_{\gamma'}^i, \tag{19.8}$$

where $C^{\gamma\gamma'}$ is the (inverse of the) intersection matrix for cycles γ, γ'. The periods Π_γ^i are holomorphic in their dependence on the moduli only when W is (quasi-) homogeneous. We will learn more about D-branes in Landau–Ginzburg theories in Ch. 39.

Part 3

Mirror Symmetry: Physics Proof

CHAPTER 20

Proof of Mirror Symmetry

We are now ready to present a physical proof of mirror symmetry. First we have to clarify what we mean by a proof of mirror symmetry. Next we divide the proof into a few steps. The basic ingredient in the proof is a formulation of the sigma model in the context of the gauged linear sigma model and application of $R \to 1/R$ duality to the charged fields of the gauged linear sigma model.

20.1. What is Meant by the Proof of Mirror Symmetry

As discussed in detail in the context of $(2,2)$ supersymmetric field theories in two dimensions, in the action there are F-terms and D-terms. Moreover, many interesting aspects of the theory, including correlation functions of topological field theories, are completely captured by the F-terms. And in the context of conformal theories, the D-terms are believed to be fixed by F-terms if one wishes to have a two-dimensional superconformal theory, as in the case of sigma models on Calabi–Yau manifolds.

What we mean by "proving" mirror symmetry is establishing the equivalence, up to D-term variations, of two different theories: a gauged linear sigma model, which has a low-energy description as a non-linear sigma model, and a Landau–Ginzburg theory with a certain superpotential, W. Moreover, the A-ring (and all the other topological data) of the gauged linear sigma model, maps to the B-ring (and the corresponding topological amplitudes) of the Landau–Ginzburg model.

As we have seen in previous chapters, some Calabi–Yau sigma models have a Landau–Ginzburg description in a certain regime of parameters. Put differently, certain Landau-Ginzburg theories can be viewed as Calabi–Yau sigma models, where the B-ring of the Landau-Ginzburg theory maps to the B-model topological ring of the Calabi–Yau. Some of the mirror Landau–Ginzburg theories that we obtain are of this type and can thus be related to a Calabi–Yau sigma model. In such a case, mirror symmetry maps the A-model topological amplitudes in one Calabi–Yau, M, to the B-model topological amplitudes of another Calabi–Yau, $\widetilde{M}$. Then one should have a relation between the Hodge numbers of the Calabi–Yau:

$$h^{p,q}(M) = h^{d-p,q}(\widetilde{M}),$$

where d is the complex dimension of M and $\widetilde{M}$. This is the original form in which mirror symmetry was posited. But the proof we present is more general, and the Landau–Ginzburg theories we obtain do not always correspond to sigma models on Calabi–Yau manifolds (or in general any manifold). For example, we will uncover the mirrors of general Fano varieties, and we will find that in general the mirror is a Landau-Ginzburg theory that does not admit a sigma model description on a compact manifold. Even for Calabi–Yau manifolds, the notion that mirror symmetry should not be thought of as simply an equivalence between two CY manifolds has long been understood — for example, there are examples of rigid Calabi–Yau threefolds where $h^{2,1} = 0$, which cannot have geometric mirrors as Calabi–Yau sigma models (as that would require $h^{1,1} = 0$, which is not possible for a Kähler manifold[1])

20.2. Outline of the Proof

The proof of mirror symmetry will be completed in three steps. In Step 1, we consider a (2,2) supersymmetric $U(1)$ gauge theory coupled to a single matter (chiral) field of charge Q. We will find that there is a dual description of this theory in which the charged chiral field Φ is replaced by a neutral, twisted chiral field, Y. This dualization amounts to a dualization of the phase of the complex-valued field Φ, which is just T-duality (i.e., $R \to 1/R$ duality) on the cylinder. In Step 2, we shall generalize to a $U(1)$ gauge theory with n chiral fields Φ_i with charges Q_i. To this end, we will consider a $U(1)^n$ gauge theory with this content and argue that deforming this to a $U(1)$ theory does not affect the superpotential (F-terms are all that we are interested in). But the $U(1)^n$ theory with n chiral fields reduces to Step 1, so at the end of Step 2 we will have constructed the mirrors of toric varieties. To generalize this to hypersurfaces (or complete intersections) in toric varieties, we need one additional ingredient, which we consider as Step

[1]An example of this is the $\mathbb{Z}_3 \times \mathbb{Z}_3$ orbifold of $T^2 \times T^2 \times T^2$ where the $\mathbb{Z}_3$'s act on hexagonal tori, with determinant 1. The mirror is a Landau–Ginzburg theory with six fields with a homogeneous superpotential of degree 3, modded out by $\mathbb{Z}_3$.

3. For simplicity of presentation we consider the mirror of a hypersurface of degree d in $\mathbb{CP}^n$. To this end we will study the $U(1)$ gauge theory with $(n+2)$ matter fields with charges $(-d, 1, \cdots, 1)$. This is a special case of the theories considered in Step 2 and describes the (non-compact) total space of the bundle $\mathcal{O}(-d)$ over $\mathbb{CP}^n$. We will connect this to the hypersurface of degree d in $\mathbb{CP}^n$ by a deformation which does not affect the relevant rings under consideration but does affect the topology of field space, and in particular leads to a change of field spaces from products of cylinders to products of complex planes.

The most important step in the proof is Step 1, and already one sees that the basic idea of mirror symmetry is simply T-duality.

20.3. Step 1: T-Duality on a Charged Field

Consider a $U(1)$ gauge theory coupled to a single chiral field Φ of charge Q. The gauge invariant field strength is in a twisted chiral superfield Σ. Recall that there is a linear superpotential in the twisted chiral sector given by $\widetilde{W} = -t\Sigma$ where $t = r - i\theta$ is the Fayet–Iliopoulos parameter. There is no superpotential for Φ, basically because there is no conceivable term that could be added consistent with gauge invariance. Since Σ and Φ live in different superspaces, they do not mix with each other as far as F-terms are concerned. We will dualize the phase of the field Φ and the dual description will be in terms of a twisted chiral field Y. Since Y and Σ are both twisted chiral, they can mix in the twisted superpotential. We will compute this superpotential $\widetilde{W}$ by studying vortices in the original theory.

The action for the gauged linear sigma model is given by

$$\mathcal{L} = \int d^4\theta(\overline{\Phi}e^{2V}\Phi - \frac{1}{2}\overline{\Sigma}\Sigma) + \frac{1}{2}(\int d^2\tilde{\theta}(-t\Sigma) + \text{complex conjugate}). \tag{20.1}$$

Recall that V is a real superfield and Σ is the gauge invariant field strength superfield given by $\Sigma = \overline{D}_+ D_- V$. $t = r - i\theta$ is the Fayet–Iliopoulos parameter. In the language of the non-linear sigma model, t typically parametrizes the complexified Kähler class: r corresponds to the Kähler form and θ to the two-form field B. Φ is chiral: $\overline{D}_+\Phi = 0 = \overline{D}_-\Phi$.

The vacuum manifold is the space of gauge-inequivalent minima of the bosonic potential (see Ch. 17), and for this theory is given by

$$\mathcal{M} = \{\phi \in \mathbb{C} : |\phi|^2 - r = 0\}/U(1),$$

which is just a point. Even though this might seem to be a rather uninteresting example, we will find that our analysis of the general case will reduce to this rather trivial looking case.

We wish to dualize the phase of the field $\phi = \rho e^{i\psi}$. This is just $R \leftrightarrow 1/R$ on the circle-valued variable ψ, but we will set the dualization in the superfield language, where we get a twisted chiral field Y whose real part $Y+\overline{Y} = 2\overline{\Phi}e^{2QV}\Phi$, and whose imaginary part $\theta = \mathrm{Im}\, Y$ is given by $d\theta = *d\psi$, as discussed in Sec. 13.4.2. Furthermore, the (twisted) superpotential (at the level of classical equivalence) will turn out to be $\widetilde{W}(\Sigma, Y) = -t\Sigma + Q\Sigma Y$. There is a quick heuristic way to see why the second term must be generated: the original (gauge invariant) kinetic energy of ϕ is given by

$$\int (D_\mu\phi)^2 = Q\int \rho^2 A\wedge *d\psi + \cdots \to Q\int A\wedge d\theta + \cdots = -Q\int F\wedge\theta + \cdots,$$

where F is the field strength of the gauge field A. The last term arises from the superspace term

$$\int d^2\theta Q\Sigma Y$$

upon integration over the superspace. Let us look more carefully at how this superpotential comes about.

Consider the Lagrangian for a vector superfield V, a real superfield B, and a twisted chiral superfield, Y whose imaginary part is periodic with period 2π :

$$\mathcal{L}_0 = \int d^4\theta (e^{2QV+B} - \frac{1}{2}(Y+\overline{Y})B), \tag{20.2}$$

where Q is an integer. Our strategy will be to integrate out one of the fields — Y or B. The different resulting theories will look different but be the same (at the level of equations of motion — quantum corrections will be considered shortly).

First we integrate over Y. This yields the constraint $\overline{D}_+D_-B = 0 = D_+\overline{D}_-B$, solved by $B = \Psi\overline{\Psi}$. Inserting this into the original Lagrangian, and setting $\Phi = e^{\Psi}$, we obtain

$$\mathcal{L}_1 = \int d^4\theta\ \overline{\Phi}e^{2QV}\Phi. \tag{20.3}$$

Now reverse the order of integration and integrate B out of $\mathcal{L}_0$, using its equation of motion, $B = -2QV + \log((Y+\overline{Y})/2)$, to obtain

$$\mathcal{L}_2 = \int d^4\theta [QV(Y+\overline{Y}) - \frac{1}{2}(Y+\overline{Y})\log(Y+\overline{Y})] \tag{20.4}$$

(we used the fact that $\int d^4\theta(Y+\overline{Y}) = 0$). Since Y is twisted chiral ($\overline{D}_+Y = 0 = D_-Y$), we can rewrite

$$\int d^4\theta VY = -\frac{1}{4}\int d\theta^+ d\bar{\theta}^- \overline{D}_+ D_- VY = \frac{1}{2}\int d^2\tilde{\theta}\Sigma Y.$$

Adding back the gauge kinetic term and the linear twisted superpotential for Σ (neither of which depend on Y or B, so are unaffected by the above manipulations), we find that the Lagrangian

$$\mathcal{L}' = \int d^4\theta(-\frac{1}{2e^2}\overline{\Sigma}\Sigma - \frac{1}{2}(Y+\overline{Y})\log(Y+\overline{Y})) + \frac{1}{2}(\int d^2\tilde{\theta}\Sigma(QY - t) + \text{c.c}) \tag{20.5}$$

is dual to

$$\mathcal{L} = \int d^4\theta(-\frac{1}{2e^2}\overline{\Sigma}\Sigma + \overline{\Phi}e^{2QV}\Phi) + \int d^2\theta(-t\Sigma). \tag{20.6}$$

Notice that the gauge superfield V (hence the gauge symmetry) has disappeared in $\mathcal{L}'$ and that Y is neutral.[2]

Comparing the two different expressions for B we see that $Y + \overline{Y} = 2\overline{\Phi}e^{2QV}\Phi$. The imaginary part of Y is related to the phase of Φ, something that is not easy to see in the superfield language, but evident in component form.

There is a puzzle here: The real part of Y looks manifestly positive, which means that the field space is the half-plane. One might be worried about potential singularities coming from the boundary in field space. However, the field Y that we have been talking about is really the bare field Y_0. We saw in our discussion of renormalization effects in Ch. 14 that in general fields get renormalized. Indeed the field Y undergoes a renormalization given by $Y_0 = \log(\Lambda_{UV}/\mu) + Y$ where μ is the renormalization scale. The condition $\mathrm{Re}\,(Y_0) \geq 0$ translates into $\mathrm{Re}\,(Y) \geq -\log(\Lambda_{UV}/\mu)$ for the renormalized field. In the continuum limit ($\Lambda_{UV} \to \infty$), the bound on Y disappears. In other words, the boundary at $Y_0 = 0$ is not physically relevant.

What we have done so far is to establish an equivalence between two Lagrangians by introducing and integrating out auxiliary fields using their equations of motion. All this has been at a classical level. The next step is to ask what happens when we take quantum effects into consideration.

[2]On worldsheets with boundaries, there would also be boundary terms that we have neglected in the above manipulations. These can be dealt with straightforwardly.

The D-terms will receive corrections, but as discussed before, these are not important for us. How about corrections to the (twisted) superpotential induced by quantum effects? One may naively think that there are no further corrections due to the non-renormalization theorems we have discussed for the superpotential terms. However, what we are doing here is integrating out fields at the quantum level. We saw in the examples we discussed in Ch. 14, in the context of non-renormalization theorems, that superpotential terms can be generated in the process without contradicting the non-renormalization theorem.

We have taken into account classical configurations in the original theory and now we want to see what quantum configurations could contribute to $\widetilde{W}$. We know that, by supersymmetry, the path-integral for this computation will localize to field configurations whose fermionic variation is zero.

EXERCISE 20.3.1. *Prove this statement by using the fact that the superpotential terms are invariant under suitable supersymmetry.*

These configurations will turn out to be instantons or vortices (gauge configurations with non-vanishing c_1) in the original formulation of the theory, and they will result in a correction of the form $\delta\widetilde{W} = e^{-Y}$.

For the fermionic variations, we have (where $\delta = \epsilon_+ Q_- + \bar{\epsilon}_- \overline{Q}_+$, since we are interested in variations with respect to Q_- and $\overline{Q}_+$)

$$\delta\lambda_+ = \sqrt{2} i \epsilon_+ \left(D + iF_{12} + \frac{i}{2}[\sigma, \bar{\sigma}] \right), \tag{20.7}$$

$$\delta\lambda_- = \sqrt{2}\epsilon_+ D_- \sigma, \tag{20.8}$$

$$\delta\overline{\lambda}_+ = \sqrt{2}\bar{\epsilon}_- D_+ \sigma, \tag{20.9}$$

$$\delta\overline{\lambda}_- = -\sqrt{2} i \bar{\epsilon}_- \left(D + iF_{12} + \frac{i}{2}[\sigma, \bar{\sigma}] \right), \tag{20.10}$$

$$\delta\psi_+ = \sqrt{2} i \bar{\epsilon}_- D_+ \phi + \sqrt{2}\epsilon_+ F, \tag{20.11}$$

$$\delta\psi_- = \sqrt{2}\bar{\epsilon}_- \sigma\phi. \tag{20.12}$$

Here $D_\pm = D_0 \pm D_1$, and D and F are the auxiliary fields in the vector and chiral multiplets (note that F is *not* a gauge field strength here, as the field strength's two-form indices have been indicated explicitly), while λ is the fermion in the gauge multiplet. Continuing to Euclidean signature by

Wick rotation ($D_+ \to 2D_{\bar{z}}$, $D_- \to -2D_z$, $F_{01} \to -iF_{12}$), the condition for the fermionic variations to be zero is

$$\sigma = 0, \tag{20.13}$$

$$D_{\bar{z}}\phi = 0, \tag{20.14}$$

$$F_{12} = e^2(|\phi|^2 - r_0). \tag{20.15}$$

The solutions to these equations are called "vortices." These are minimum-action solutions in the topological class given by

$$c_1 = k = \frac{1}{2\pi}\int F_{12} d^2x.$$

The trivial solution ($k = 0$) corresponds to $\phi =$ constant, and localization to that sector gives the classical superpotential $\widetilde{W}$ we have already computed. Consider the case $k = 1$. To understand what this solution means, note that the bosonic potential is

$$U(\phi) = \frac{e^2}{4\pi}(|\phi|^2 - r_0)^2.$$

There is a circle of minima of this potential in the ϕ plane, and to have a configuration with finite action, we must demand that ϕ goes to the minimum at space-time infinity. However, we see that there can be a non-trivial configuration in which the ϕ field winds around the circle of minima at infinity, i.e., ϕ defines a map

$$\phi : S^1_\infty \to S^1_{U=0}$$

with winding number 1. However, in this configuration the phase of ϕ varies at ∞, and so to ensure that the gauge covariant derivative $D_{\bar{z}}\phi = 0$ we need to turn on a gauge field with

$$k = \frac{1}{2\pi}\int F_{12} d^2x = \int_{S^1_\infty} A = 1.$$

This solution is called the vortex, and is a BPS configuration, which means that it is invariant under half of the supercharges. For the $k = 1$ vortex, ϕ has a single, simple zero, and the moduli space of this solution is complex one-dimensional, parametrized by the location of the zero of ϕ.

To do the analysis of the contribution of the vortices to the deformation of the superpotential, it is simplest to first study the case where the charge Q of the field Φ is $Q = 1$. In the $k = 1$ vortex background, there are two fermionic zero modes which induce, as can be checked by explicit calculation, a nonzero two-point function

$$\langle \chi_+ \overline{\chi}_- \rangle \sim e^{i\theta},$$

where the χ's are the fermionic partners of Y and in the original field variables are given by

$$\chi_+ = 2\overline{\psi}_+ \phi, \overline{\chi}_- = -2\phi^\dagger \psi_-.$$

On the other hand, a superpotential term $\int d^2\tilde{\theta} e^{-Y}$ would contain a term $-e^{-Y}\overline{\chi}_+\chi_-$ and this would give a contribution to the above two-point function of precisely this form. This is because the imaginary part of Y can be identified with the "theta angle" of the gauge theory, which is encoded in the $\int \Sigma Y$ term in the action that we discussed above. This proves that the term e^{-Y} is in fact generated.

Since the product of the two χ's carries axial R-charge -2, only vortex backgrounds with $k = 1$ can contribute to the two-point function. Hence higher k configurations, which might have generated terms like e^{-kY}, do not contribute to the superpotential. Note, however, that the one-vortex contribution encodes configurations with multiple $k = 1$ vortices to correlation functions, because the action is exponentiated in the path-integral. So we have shown that the vortices (which can also be called instantons of this two-dimensional theory) generate a term e^{-Y} in the superpotential $\widetilde{W}$. The (twisted) superpotential in the dual formulation is then

$$\widetilde{W} = \Sigma(Y - t) + e^{-Y}. \tag{20.16}$$

For a general charge Q, the answer turns out to be

$$\widetilde{W} = \Sigma(QY - t) + e^{-Y}.$$

One way to see this is to note that changing the quantum of charge is effectively the same as $V \to QV$ or $\Sigma \to Q\Sigma$ (up to a redefinition of t).[3]

There is a faster way to see why precisely such a term is generated, which also sheds light on its role and the meaning of T-duality in this context. As we have seen before in Ch. 18, superpotential terms encode information about BPS solitons (or kinks) that interpolate between vacua, labeled by the critical points of the superpotential. One can also reverse this: by examining what BPS solitons we have, we can try to reconstruct the superpotential. In the present case it is convenient to consider taking the limit $e \to 0$. Note that as far as variations of the superpotential this is irrelevant because it is a D-term variation. However, this has the effect of freezing out all fluctuations in the field Σ, because of the kinetic term

$$\frac{1}{e^2}\int d^4\theta \overline{\Sigma}\Sigma,$$

so we can set Σ to a constant. In the original formulation, this induces a central charge in the supersymmetry algebra proportional to $q\Sigma$, where q denotes the sector of the Hilbert space with charge q. In particular, if we consider a sector consisting of n excitations of the field Φ, this corresponds to $q = nQ$. Thus the Hilbert space decomposes into charge sectors in which the mass $M \geq n|Q\Sigma|$, which is saturated by the BPS solutions. Regardless of whether we actually do have a stable kink or not in any given sector, or how many of each we have, we should have infinitely many vacua for this theory labeled by an integer, and the central term in the supersymmetry algebra between any pair is given by the difference in the integer labels n times $Q\Sigma$, i.e.,

$$Z_n = 2\pi i n Q\Sigma$$

, How are these sectors appearing in the dual formulation in terms of the Y fields? If we only had $\widetilde{W} = \Sigma(QY - t)$, there would not be infinitely many sectors, as there would not be any critical points of $\widetilde{W}$ as a function of Y. With a superpotential $\widetilde{W} = \Sigma(QY - t) + e^{-Y}$, it is easy to see that we have infinitely many critical points given by

$$\partial_Y \widetilde{W} = 0 \to e^{-Y} = Q\Sigma.$$

If Y_0 is a solution of this equation, we will have infinitely many of them given by $Y_0 + 2\pi i n$. In a sector where Y changes from left to right by $2\pi i n$, we see that

$$\Delta\widetilde{W} = 2\pi i(nQ\Sigma)$$

which is exactly of the expected form. It takes a little bit more work to show that no other addition to the superpotential does the trick, so that this is the unique deformation. Note in particular that when Σ is frozen out, Φ acquires a mass $|\Sigma|$, which is precisely the mass of the kink solution in the dual formulation. In other words, the duality interchanges what is a fundamental field in the original description with what is a soliton in the

[3]In terms of vortices, "fractional" vortices contribute in this case. What this means is that the field configurations with fractional c_1 (multiples of $1/Q$) contribute to the superpotential.

dual description. Moreover, vortices (winding modes in Φ) correspond to momentum modes of Y and vice versa — this is a manifestation of the momentum-winding exchange that is characteristic of $R \to 1/R$ duality.

20.4. Step 2: The Mirror for Toric Varieties

In Step 2, we consider $U(1)^n$ gauge theory with n chiral multiplets Φ_i of charge Q_i (with respect to the ith $U(1)$). Of course this is just equivalent to n decoupled copies of the theory we studied in Step 1. Thus, after we dualize the fields Φ_i into fields Y_i we immediately see that we have an effective superpotential in terms of the Y_i given by

$$\widetilde{W} = \sum_{i=1}^{n}(Q_iY_i - t_i)\Sigma_i + \sum_{i=1}^{n} e^{-Y_i}. \tag{20.17}$$

Now freeze all except the diagonal $U(1)$ by sending the corresponding gauge coupling parameters e_i to zero, by D-term variations. Note that this does not affect what we are interested in, the F-terms. This reduces the problem to that of a single $U(1)$, and so we have

$$\widetilde{W} = (\sum_{i=1}^{n} Q_iY_i - t)\Sigma + \sum_{i=1}^{n} e^{-Y_i}, \tag{20.18}$$

after setting $\Sigma_i = \Sigma$ and $t = \sum_{i=1}^{n} t_i$. Note that Σ is still a dynamical superfield. Integrating out Σ, which as discussed above is the same as solving for $\partial_\Sigma \widetilde{W} = 0$, we get the superpotential

$$\widetilde{W} = \sum_{i=1}^{n} e^{-Y_i}$$

subject to the constraint

$$\sum_{i=1}^{n} Q_iY_i = t.$$

On the other hand, we know that the $U(1)$ theory with n chiral fields Φ_i with charges Q_i has a low-energy limit which is the NLSM with target space $W\mathbb{CP}(Q_1, \cdots, Q_n)$. So we have established that the mirror of weighted projective space is an LG theory with a specific superpotential. This proof extends to arbitrary toric varieties, and completes Step 2.

As an example, consider the case of $\mathbb{CP}^{n-1}$. The gauged linear sigma model description is in terms of a $U(1)$ gauge field coupled to n chiral fields of charge 1. The equivalent (dual) Landau–Ginzburg description is in terms of $X_i = e^{-Y_i}$, $i = 1, \ldots, n$, with superpotential $\widetilde{W} = X_1 + \cdots + X_n$ subject

to the constraint $\prod_i X_i = e^{-t}$. Equivalently (solving for X_n), we have $n-1$ fields X_i with

$$\widetilde{W}(X_i) = X_1 + \cdots + X_{n-1} + \frac{e^{-t}}{X_1 \cdots X_{n-1}}. \tag{20.19}$$

We should remember, though, that the fundamental variables (dictated by the measure of the path-integral) are the Y_i and not the X_i. Note that t started out life as the complexified Kähler parameter of $\mathbb{CP}^{n-1}$, but now appears in the Landau–Ginzburg description in a transcendental expression.

It is easy to see that there are n critical points of the superpotential, corresponding to the n massive vacua of the $\mathbb{CP}^{n-1}$ model. Indeed, they are obtained by

$$\partial_i \widetilde{W} = 0 \to X_i = \omega e^{-t/n},$$

where ω is an n-th root of unity.

Identifying $H = -\partial_t \widetilde{W}$ with the Kähler class of $\mathbb{CP}^{n-1}$, we recover the quantum cohomology ring of $\mathbb{CP}^{n-1}$ realized here as the twisted chiral ring of this Landau–Ginzburg theory:

$$H^n = e^{-t}.$$

Notice that as $t \to \infty$, the $\mathbb{CP}^{n-1}$ becomes flat and the quantum corrections coming from worldsheet wrappings (sigma model instantons) get suppressed, and we recover the classical cohomology ring relation, $H^n = 0$. However there is more to this story than the quantum cohomology ring, as we have seen. One can compute the soliton spectrum corresponding to kinks interpolating between different vacua) in the LG description.

EXERCISE 20.4.1. *Prove that between any pair of vacua with $\omega_i/\omega_j = e^{-2\pi i r/n}$ there are $n!/r!(n-r)!$ kinks.*

The n kinks between neighboring vacua are identified with the fundamental fields Φ_i in the original description and, more generally, the higher solitons correspond to antisymmetric products of the fields Φ_i. This picture was also anticipated, as we discussed before, from the large n analysis of the $\mathbb{CP}^{n-1}$ sigma model.

In the next step we will prove mirror symmetry for hypersurfaces in projective space.

20.5. Step 3: The Hypersurface Case

We saw in Step 1 that the non-linear sigma model on the weighted projective space $W\mathbb{CP}(Q_1,\cdots,Q_n)$ is equivalent to an LG theory with the superpotential

$$\widetilde{W}(Y_i)=\sum_{i=1}^{n}e^{-Y_i},$$

with the constraint $\sum_i Q_iY_i = t$. We can equivalently write $\widetilde{W}(X_i) = \sum_{i=1}^n X_i$ with a constraint $\prod_{i=1}^n X_i^{Q_i} = e^{-t}$ (this latter form is more familiar from the toric geometry viewpoint).

The machinery we have developed also allows us to consider non-compact toric varieties, as we have not assumed that all the charges Q_i are positive. For instance consider the gauged linear sigma model with $(n+2)$ chiral fields $(P,\Phi_1,\cdots,\Phi_{n+1})$ with charges $(-d,1,1,\cdots,1)$. This corresponds to the total space of $\mathcal{O}(-d)$ on $\mathbb{CP}^n$ (i.e., the low-energy limit of this gauged linear sigma model is the non-linear sigma model on this space). In the special case where $d=n+1$, we have a non-compact Calabi–Yau manifold.[4] Let us define $X_0=e^{-P}, X_i=e^{-Y_i}$ for $i=1$ to $(n+1)$. Then in the mirror LG description $\widetilde{W}=\sum_{i=0}^{n+1}X_i$ with the constraint $X_0^{-d}X_1\cdots X_{n+1}=e^{-t}$. We can use the constraint to eliminate X_0 by defining

$$\widetilde{X}_i = X_i^{1/d}$$

for $i=1$ to $(n+1)$ so that

$$X_0 = e^{t/d}\widetilde{X}_1\cdots\widetilde{X}_{n+1}$$

and the superpotential becomes

$$\widetilde{W}=\sum_{i=1}^{n+1}\widetilde{X}_i^d + e^{t/d}\widetilde{X}_1\cdots\widetilde{X}_{n+1}.$$

Note that the change of variables $\widetilde{X}_i = X_i^{1/d}$ means that we are redefining $\widetilde{Y}_i = Y_i/d$. Remembering that the Y_i were periodic variables ($Y_i\sim Y_i+2\pi i$), we see that we need to identify $\widetilde{X}_i\sim e^{2\pi i/d}\widetilde{X}_i$. This means that the $\widetilde{X}_i$ are only well defined up to multiplication by a dth root of unity, so our theory is really the "orbifold" of the LG theory with $\widetilde{W}=\sum_{i=1}^{n+1}\widetilde{X}_i^d+e^{t/d}\widetilde{X}_1\cdots\widetilde{X}_{n+1}$

[4] As we will discuss briefly later, the physical motivation for considering such non-compact Calabi–Yau's, i.e., the "local models," is that non-compact CY's capture some interesting aspects of the resulting field theory near a singularity of a CY in string theory.

by the group $(\mathbb{Z}_d)^n$. Note that it is not $(\mathbb{Z}_d)^{n+1}$ because the combination $e^{-t/d}X_0 = \widetilde{X}_1\cdots\widetilde{X}_{n+1}$ is well defined in the original theory, and so the orbifold group is the subgroup $(\mathbb{Z}_d)^n$ of $(\mathbb{Z}_d)^{n+1}$ corresponding to phase rotations preserving this monomial. The meaning of this orbifolding is the same as that of a discrete gauge symmetry: in the path-integral description, one sums over all flat $(\mathbb{Z}_d)^n$ bundles, and the fields (depending on how they transform according to the discrete group) are appropriate sections.

For $d=n+1$, we have an $\mathcal{O}(-n-1)$ bundle on $\mathbb{CP}^n$ and in this case the Landau–Ginzburg theory is the $(\mathbb{Z}_d)^n$ orbifold of the one with superpotential

$$\widetilde{W}=\widetilde{X}_1^{n+1}+\cdots+\widetilde{X}_{n+1}^{n+1}+e^{t/(n+1)}\widetilde{X}_1\cdots\widetilde{X}_{n+1} \tag{20.20}$$

In precisely this case, $\widetilde{W}$ is homogeneous and there is an extra $U(1)$ R-symmetry. This looks exactly like the Landau–Ginzburg description of the *mirror* to a compact Calabi–Yau hypersurface in $\mathbb{CP}^n$, with t now identified with the complex structure modulus of the mirror. But what we started out describing was the non-compact total space of $\mathcal{O}(-n-1)$ on $\mathbb{CP}^n$! What is the relation between these two?

In order to address this question, recall that in order to get a compact hypersurface of degree d in $\mathbb{CP}^n$ we need to consider a GLSM with a superpotential $W=PG_d(\Phi_i)$, where G_d is a homogeneous polynomial of degree d. This is a term that is integrated over chiral superspace. We can think of this additional term as a (c,c) perturbation to the gauged linear sigma model of the non-compact space. In the mirror description, i.e., in terms of the Y_i variables, this corresponds to a term $\int d^2\theta\sigma_P G(\sigma_{\Phi_i})$ where the fields σ_P, σ_{Φ_i} correspond to fields creating kinks, and will have a complicated description in terms of the fields Y_i. (Recall that the fundamental fields Φ_i in the original description corresponded to kinks in the dual description in terms of Y_i.) As a result of this perturbation in the (c,c) superspace terms, the theory will flow in the IR to the *compact theory*. In the UV, this theory has a non-compact $(n+1)$-dimensional target, and in the IR it has $n-1$ dimensions, so this is obviously a drastic change. However, quantities which are sensitive only to the twisted F-terms (such as the quantum cohomology) should not depend on this perturbation, as was discussed in detail in the context of decoupling theorems. This explains why we are getting the same form for the superpotential for both the compact hypersurface as well as the non-compact toric space. However, there must be some difference between them. It turns out that there is a subtle difference between the compact and the non-compact theory: the field space may have a different topology in the IR. As we will argue below, while in the non-compact case the good variables were the Y_i, the good variables in the compact case are the $X_i = e^{-Y_i}$.

We can probe the relation between the non-compact and compact theories in the following way. The ground states of the sigma model for $\mathcal{O}(-d)$ on $\mathbb{CP}^n$ are in one-to-one correspondence with the cohomology classes $1, k, \ldots, k^n$, with axial charge $-(n+1)/2, \cdots, (n+1)/2$. If we turn on the F-term perturbation and follow the RG flow to the IR, some of these ground states will flow to ground states of the compact theory, and some will be lifted. The ground states that correspond to normalizable forms in the non-compact theory are bound to survive in the compact theory, whereas those that correspond to non-normalizable forms might disappear from the spectrum (they can run infinitely far away in field space). The ground state with the lowest axial charge $-\frac{d}{2}$ (recall again $d = n+1$) corresponds to the cohomology class dual to the total space of the line bundle, and this is clearly a non-normalizable state (cannot be expressed as a form with compact support). The ground state corresponding to the fundamental class of $\mathbb{CP}^n$ (zero section) in the line bundle has axial charge $-\frac{d}{2}+1$ and survives in the compact theory as the unique ground state with that charge, which we shall label $|1\rangle_c$. In other words, the ground state of the compact theory corresponds to the form k, or the state $|k\rangle_{nc}$, in the non-compact theory.[5]

Now we can use the path-integral to identify the good variables in the compact case in terms of those of the non-compact theory. In the non-compact theory, the good variables (on the mirror side) are the $\widetilde{Y}_i = -\log(\widetilde{X}_i)$. In particular, as discussed in Ch. 32, the D-brane BPS masses correspond to periods, and are given by

$$\Pi^{\gamma}_{nc} = \langle\gamma|1\rangle_{nc} = \int_\gamma d\widetilde{Y}_1 \cdots d\widetilde{Y}_{n+1} e^{-\widetilde{W}} \tag{20.21}$$

$$= \int_\gamma \frac{d\widetilde{X}_1 \cdots d\widetilde{X}_{n+1}}{\widetilde{X}_1 \cdots \widetilde{X}_{n+1}} e^{-\widetilde{W}}. \tag{20.22}$$

[5] Strictly speaking, the notion of axial charge as a conserved quantity makes sense only for Calabi–Yau (or when the size of $\mathbb{CP}^n$ is large) but one still can define the state corresponding to $|k\rangle_{nc}$ by more abstract means.

Now the variation of the Kähler class in the non-compact theory is realized by taking the derivative of the superpotential with respect to t, and this gives

$$k = \partial_t \widetilde{W} = \frac{1}{d} e^{t/d} \prod_{i=1}^{n+1} \widetilde{X}_i.$$

Inserting this in the integral gives

$$\langle\gamma|k\rangle_{nc} = \langle\gamma|1\rangle_c = \frac{1}{d} e^{t/d} \int_\gamma d\widetilde{X}_1 \cdots d\widetilde{X}_{n+1} e^{-\widetilde{W}}. \tag{20.23}$$

In other words,

$$\frac{\partial \Pi_{nc}}{\partial t} = \Pi_c.$$

Note that the final integral of the period can be interpreted in the context of the compact theory as implying that the correct variables in the compact theory are the $\widetilde{X}_i$ variables. Thus the effect of the RG flow is that the good variables in the compact case are the $\widetilde{X}_i = e^{-\widetilde{Y}_i}$ instead of $\widetilde{Y}_i$.

To summarize, we have found that the mirror of a degree d hypersurface in $\mathbb{CP}^n$ is a Landau–Ginzburg theory with (twisted) superpotential $\widetilde{W} = X_1^d + \cdots + X_{n+1}^d + e^{t/d} X_1 \cdots X_{n+1}$ modded out by the discrete gauge group $(\mathbb{Z}_d)^n$, consisting of all $\mathbb{Z}_d$ phase rotations of X_i preserving $X_1 \cdots X_{n+1}$. When $d = n+1$, the Calabi–Yau case, $\widetilde{W}$ is homogeneous and the Landau–Ginzburg theory can be identified with a non-linear sigma model on a mirror Calabi–Yau. However, what we have found is more general: if $d \neq n+1$, the superpotential is not homogeneous, and there is no geometric interpretation of the mirror theory. In other words, we cannot insist on the mirror of a sigma model on a manifold being another ordinary sigma model on a mirror manifold, in the general case. As mentioned before, sometimes even in the Calabi–Yau case the mirror is not a sigma model on a manifold.

Note that in the present example $c_1 = n + 1 - d$ and as noted before the sigma model is asymptotically free if $c_1 > 0$, conformal if $c_1 = 0$ and infrared-free if $c_1 < 0$. If $c_1 > 0$, we see from the LG description that there are $n+1-d$ massive vacua and a massless vacuum (with some multiplicity) at the origin (in field space).

EXERCISE 20.5.1. *Demonstrate the above statement.*

If we flow this theory to the IR, the massive vacua move out to infinity and the IR limit is an orbifold CFT with $\widetilde{W} = X_1^d + \cdots + X_{n+1}^d$. In the case $c_1 < 0$ the physical theory is not believed to make sense as it is not asymptotically free. The theory is given in the IR by a superpotential $\widetilde{W} = e^{t/d} X_1 \cdots X_{n+1}$, which looks rather pathological from a physical point of view (it has a highly non-degenerate critical locus). However one can still talk about the corresponding topological theory obtained by twisting, and study $\mathcal{R} = \mathbb{C}[X_i]/(\partial_i W)$, etc.

The original derivation of mirror pairs was rather different from the above proof. Let us connect our result to the original Greene–Plesser construction of mirror pairs. We have seen that the mirror of the quintic hypersurface in $\mathbb{CP}^4$ is given by $\widetilde{W} = X_1^5 + \cdots + X_5^5 \bmod (\mathbb{Z}_5)^4$. On the other hand, the Calabi–Yau/Landau–Ginzburg correspondence discussed in the context of gauged linear sigma models associates to a sigma model on the quintic threefold, for some choice of moduli, the Landau–Ginzburg theory with the same Fermat-type superpotential, but modded out by $\mathbb{Z}_5$ (acting simultaneously on all fields). What is the relation between these two Landau–Ginzburg theories? The equivalence between these two Landau–Ginzburg theories was known at the level of CFT as follows: The CFT associated to a theory with superpotential $W = X^5$ is believed to be a known minimal conformal theory. Moreover, one can check that modding out by $\mathbb{Z}_5$ symmetry gives back an equivalent minimal model. Thus in our context modding out by $(\mathbb{Z}_5)^5$ acting on all fields gives an equivalent theory. Moreover a curious general symmetry of all conformal theories implies that if we consider an orbifold of a conformal theory $\mathcal{C}_1$ by an abelian group G, denoted by $\mathcal{C}_2 = \mathcal{C}_1/G$, then there is an orbifold of the new theory by the same group which gives back the original theory

$$\mathcal{C}_1 = \mathcal{C}_2/G.$$

EXERCISE 20.5.2. *Verify this statement for $G = \mathbb{Z}_n$ at the level of the partition function of the conformal theories on T^2. Hint: Recall the definition of orbifolds which implies that there are n^2 bundles to consider on T^2, yielding n^2 terms in the partition function.*

Define the action of G on the orbifold theory $\mathcal{C}_2$ in terms of the Hilbert space sector labels of the orbifold theory, which are organized by elements of $\mathbb{Z}_n$. Now let $\mathcal{C}$ denote the conformal theory associated to the Landau–Ginzburg theory with Fermat quintic superpotential. Then from what we have said, it follows that

$$\mathcal{C}/{\mathbb{Z}_5}^5 = \mathcal{C}$$

and applying the above inversion of the orbifold action to $\mathcal{C}_1 = \mathcal{C}/\mathbb{Z}_5^4$ and $G = \mathbb{Z}_5$, we deduce that

$$\mathcal{C}/\mathbb{Z}_5^4 = \mathcal{C}/\mathbb{Z}_5.$$

Part 4

Mirror Symmetry: Mathematics Proof

CHAPTER 21

Introduction and Overview

The aim of chapters 21 to 30 is to introduce Gromov–Witten theory to both mathematicians and physicists. We begin with an expository introduction to curves, and moduli spaces of curves and maps, along with their relevant properties. We then present Givental's approach to the Mirror conjecture for hypersurfaces. We conclude with a discussion of higher-genus phenomena related to the conjectures of Gopakumar and Vafa for Calabi–Yau threefolds.

21.1. Notation and Conventions

With the exception of Sec. 22.1, throughout this part we will deal only with complex geometry. Hence all dimensions will be complex dimensions unless otherwise noted. $\mathbb{CP}^m$ will usually be denoted $\mathbb{P}^m$.

If X is a non-singular complex manifold, then the line bundle $\wedge^{\dim X}(TX)^*$ is called the *canonical* line bundle $\mathcal{K}_X$, and its cohomology class $K_X = c_1(\mathcal{K}_X)$ is called the canonical divisor class.

21.1.1. Homology and Cohomology. Let X be a non-singular complex projective variety of (complex) dimension n. The singular homology $H_*(X)$ and cohomology $H^*(X)$ theories will always be taken with $\mathbb{Q}$-coefficients.

We will identify $H^d(X)$ with $H_{2n-d}(X)$ (via Poincaré duality). A closed subvariety V of X of pure (complex) codimension d determines classes in $H_{2n-2d}(X)$ and $H^{2d}(X)$ via duality. Both of these classes are denoted by $[V]$. If $c \in H^*(X)$ and $\beta \in H_k(X)$, we denote by $\int_\beta c$ the degree of the class of the zero-cycle obtained by evaluating c_k on β, where c_k is the component of c in $H^k(X)$. When V is a closed, pure-dimensional subvariety of X, we write $\int_V c$ instead of $\int_{[V]} c$. We use cup-product notation $\cup$ for the product in $H^*(X)$. If $H_2(X,\mathbb{Z}) = \mathbb{Z}\beta$ with a natural generator β, the elements of $H_2(X,\mathbb{Z})$ will be identified with the integers. For example, $H_2(\mathbb{P}^m,\mathbb{Z}) = \mathbb{Z}[L]$ for the line class $[L]$. The class $d[L]$ will often be denoted by d.

The cohomology of a sheaf $\mathcal{F}$ (such as the structure sheaf $\mathcal{O}_X$ of algebraic functions on X) is computed by Čech cohomology in the Zariski topology (see Sec. 2.3.1). For algebraic sheaves on projective varieties, the algebraic cohomology theory coincides with the analytic cohomology theory. Cohomology vanishes above the (complex) dimension of the variety:

$$H^k(X, \mathcal{F}) = 0, \quad k > \dim X. \tag{21.1}$$

21.1.2. Deligne–Mumford Stacks. It will be necessary to work with a generalization of varieties (and schemes) known as *Deligne–Mumford stacks*. Although Deligne–Mumford stacks are quite technical to define, it is nonetheless possible to informally work with them without delving into their foundations.

Deligne–Mumford stacks are an algebraic generalization of *orbifolds*, and can be roughly thought of as an "orbivariety" or "orbischeme". An orbifold is locally the quotient of a manifold by a finite group. Similarly, a Deligne–Mumford stack is "locally" the quotient of a scheme by a finite group. Hence (like an orbifold) each point of a Deligne–Mumford stack has an associated *finite* group, called the *isotropy group* of the point. All of the Deligne–Mumford stacks which appear have the form $X/\mathbb{G}$ where X is a quasi-projective scheme and $\mathbb{G}$ is a complex algebraic Lie group acting with finite stabilizer.

In the same way that a complex orbifold has an "underlying complex variety" (with finite quotient singularities corresponding to the points with non-trivial isotropy groups), Deligne–Mumford stacks have an underlying space, called the *coarse moduli space* of the stack. In all the cases we will deal with, the coarse moduli spaces will be varieties. (In general, they can be "algebraic spaces".) The reader may find it psychologically easier to deal with the coarse moduli space, as it is a simpler algebraic object, but it will be important to keep track of the isotropy groups as well.

Deligne–Mumford stacks differ from orbifolds in two ways. They may be singular (for example, $\overline{\mathcal{M}}_{1,0}(\mathbb{P}^1, d)$, Exercise 24.3.1; see the Caution below), and the isotropy group of a general point need not be trivial (for example, $\mathcal{M}_2$, see Sec. 23.2).

As with orbifolds, cohomology must be taken with $\mathbb{Q}$-coefficients, not $\mathbb{Z}$-coefficients.

Caution: Physicists sometimes use orbifold to mean the coarse moduli space of the mathematical sense of orbifold. Thus orbifolds in the sense of physicists may be singular, but orbifolds in the mathematical sense are smooth.

21.1.3. Tips for Physicists Reading the Mathematical Literature. Physicists should be warned of terminology they will find in the mathematical literature. In algebraic geometry, *complete* or *proper* usually means compact (but they have more general, algebraic definitions). *Non-singular* means smooth (and smooth has a more technical meaning in algebraic geometry); we will use this terminology here. The *Chow ring* is an algebraic version of cohomology that contains more refined information.

21.1.4. Summary of Notation. For the reader's convenience, the following list contains the important notation and the section where it is introduced.

Σ, $\tilde{\Sigma}$	Nodal curve and its normalization, 22.2
$\mathcal{M}_g$, $\overline{\mathcal{M}}_g$, $\mathcal{M}_{g,n}$, $\overline{\mathcal{M}}_{g,n}$	Moduli spaces of curves, 23.2, 23.3, 23.4, 23.4
$\overline{\mathcal{M}}_{X,\beta}$	Moduli space of stable maps, 24
$D(g_1, A\vert g_2, B)$, $D(A, \beta_1\vert B, \beta_2)$,	Boundary divisors, 23.4.1, 24.3
ev_i	Evaluation maps, 24.3
Def, Aut, Ob	Vector spaces of infinitesimal deformations, automorphisms, and obstructions, 24.4
$\mathrm{vdim} = \mathrm{vdim}\,\overline{\mathcal{M}}_{g,n}(X,\beta)$	Virtual or expected dimension, 24.4
N_d	Gromov–Witten invariants, 25.1.1, 29.1
$\mathbb{L}_i$, $\psi_i = c_1(\mathbb{L}_i)$	Tautological line bundles, 25.2
$\mathbb{E}$, $\lambda_k = c_k(\mathbb{E})$	Hodge bundle, 25.3
$[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}$	Virtual fundamental class, 26.1
$\langle \tau_{a_1}(\gamma_1)\cdots\tau_{a_n}(\gamma_n)\rangle^X_{g,\beta}$	Gromov–Witten and descendant invariants, 26.2
$QH^*(X)$, $QH^*_s(X)$	Big and small quantum cohomology ring, 26.5 and 26.5.1
C, C_{ijk}	Gromov–Witten or quantum potential, 26.5

$T_0 = 1, T_1, \ldots, T_p, \ldots, T_m$	Basis for $H^*(X)$, 26.5
$\mathbb{T}$	Torus $(\mathbb{C}^*)^{m+1}$, 27.1
$\alpha_0, \ldots, \alpha_m$	Basis for characters of $\mathbb{T}$, generators of $H^*((\mathbb{CP}^\infty)^{m+1})$, 27.1
p_i, ϕ_i and	Fixed point of $\mathbb{T}$-action on $\mathbb{P}^m$, corresponding equivariant cohomology class, 27.1
$\overline{\mathcal{M}}_\Gamma$, $\mathbb{A}_\Gamma$	Moduli space corresponding to $\mathbb{T}$-fixed locus on $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$, and associated automorphism group, 27.3
$\mu(v)$, $\mu(F)$	Fixed point associated to a vertex v or flag F of Γ, 27.3
$d(e)$	Degree of cover corresponding to edge e of Γ, 27.3
ω_F, ψ_F	Equivariant classes associated to flag F of Γ, 27.3
$\partial_0, \ldots, \partial_m$	Tangent fields on $V = H^*(X)$, 28
$\hbar$	Generator of equivariant cohomology of $\mathbb{C}^*$
$\nabla_\hbar$	Connection on TV, 28
Ψ	Matrix giving fundamental solution to the quantum D.E., 28
S, Z_i, z_i, Z_i^*, z_i^*, S^*, Φ	Correlators arising in the Mirror conjecture for hypersurfaces, 29.2, 29.3, 29.4.2, 30.2
$E_d = \pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)$, E_d'	Obstruction bundles, 29.2
G_d^{i*}, G_d^{i0}, G_d^{il}	Graphs corresponding to $\mathbb{T}$-fixed loci on $\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m, d)$, 29.4
L_d, L_d'	Auxiliary space in proof of Mirror conjecture, 30.2
$\mathcal{P}$	"Polynomial" class of correlators, 30.3
$T(t)$	Mirror transformation or change of variables, 30.4

CHAPTER 22

Complex Curves (Non-singular and Nodal)

In this chapter, we will survey facts about complex curves that will be useful in discussing stable maps. No proofs will be given for the various deep theorems mentioned.

22.1. From Topological Surfaces to Riemann Surfaces

Topological surfaces are differentiable manifolds (see Sec. 1.2) of real dimension 2 that are oriented, compact, and connected. Such surfaces are classified by their *genus* (the number of "holes", see Fig. 1). Up to diffeomorphism, there is one such surface Σ_g for every genus $g \geq 0$.

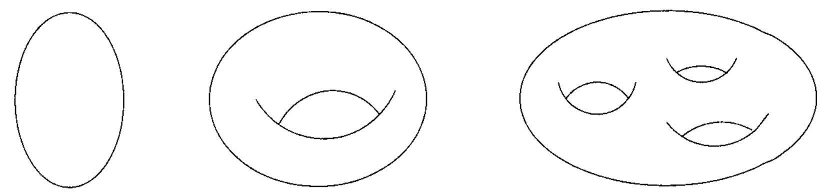


FIGURE 1. Topological surfaces of genus 0, 1, and 3

There are several natural additional structures one may place on a topological surface Σ_g.

(1) The first is a *Riemannian metric* (see Sec. 1.4.1), given by a positive definite symmetric two-tensor $g_{ij}dx^i \otimes dx^j$. This gives a notion of distance on the surface.

(2) A *conformal structure* is given by the data of a Riemannian metric up to multiplication by a (positive) function on Σ_g. A conformal structure determines angles between tangent vectors of Σ_g. Two Riemannian metrics giving the same conformal structure are said to be *conformally equivalent*.

(3) An *almost complex structure* is an automorphism of the tangent bundle $J : T\Sigma_g \to T\Sigma_g$ such that $J^2 = -1$. Every conformal structure

determines a canonical almost complex structure: J is defined by counterclockwise rotation by 90 degrees. An almost complex structure gives the fibers of $T\Sigma_g$ the structure of a one-dimensional complex vector space, where J is interpreted as multiplication by i.

EXERCISE 22.1.1. *Show that every almost complex structure on Σ_g is obtained from a canonically associated conformal structure.*

Thus there is a canonical correspondence between the set of conformal structures on Σ_g and the set of almost complex structures.

(4) An almost complex structure is *integrable* if there exist holomorphic charts for Σ_g. Every almost complex structure on a surface is integrable — this is a deep theorem in complex analysis. A surface Σ_g with a (given) complex structure (an integrable almost complex structure) is called a *non-singular complex curve* or *Riemann surface.*

(5) Finally, every Riemann surface is *algebraic* — it can be described as the vanishing set of polynomials in $\mathbb{CP}^m$ and may be studied using algebro-geometric tools (see Ch. 2). This result is also deep, and uses:

THEOREM 22.1.1 (Riemann–Roch). *Suppose V is a rank r vector bundle on a Riemann surface Σ_g. Then*

$$h^0(\Sigma_g, V) - h^1(\Sigma_g, V) = \deg c_1(V) + r(1-g).$$

Note that all higher cohomology groups of V vanish, by the dimensional vanishing Eq. (21.1).

This is a special case of the Grothendieck–Riemann–Roch formula, see Sec. 3.5.3.

In summary, the relationships for structures on topological surfaces are as follows.

Riemannian structure
$\Downarrow$
conformal structure
$\Updownarrow$
almost complex structure
$\Updownarrow$
complex structure
$\Updownarrow$
algebraic structure

These equivalences do not always hold for higher-dimensional spaces.

22.2. Nodal Curves

Singular objects play an essential role in algebraic geometry. The simplest singularity a complex curve can have is a *node*. A nodal point of a curve is a point that can be described analytically-locally by the equation $xy = 0$ in the complex plane $\mathbb{C}^2$. An example of a nodal curve is given in Fig. 2. (*Caution:* because of the difficulty of representing a node in a two-dimensional figure, it falsely appears that the branches of the node are tangent.)

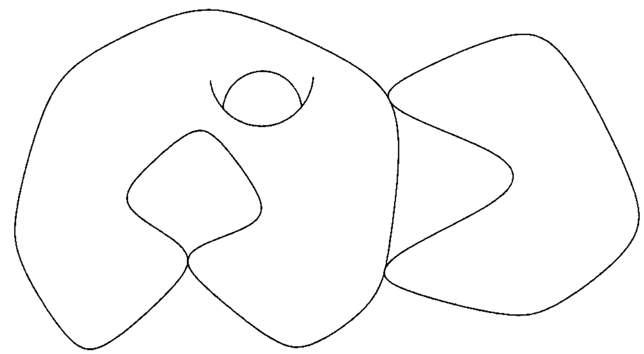


FIGURE 2. A nodal curve

DEFINITION 22.2.1. *If Σ is a nodal curve, define its* normalization $\tilde{\Sigma}$ *to be the Riemann surface obtained by "ungluing" its nodes. Let*

$$\nu : \tilde{\Sigma} \to \Sigma$$

denote the canonical normalization *map. The preimages in $\tilde{\Sigma}$ of the nodes of Σ are the* node-branches. *If $\tilde{\Sigma} = \cup\Sigma_i$ is the decomposition of $\tilde{\Sigma}$ into (connected) Riemann surfaces, $\nu(\Sigma_i)$ are the* irreducible components *of Σ.*

The normalization of the curve in Fig. 2 is given in Fig. 3. The node-branches are marked.

A "half-dimensional representation" of a nodal curve is given in Fig. 4. Each component is labelled with its genus.

Another convenient way of describing a nodal curve is by its *dual graph.* The vertices of the dual graph of Σ correspond to components of Σ (and are

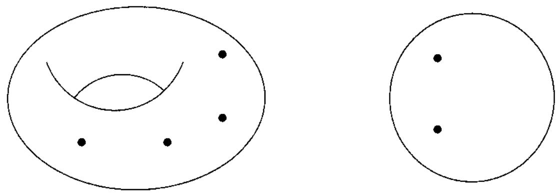


FIGURE 3. The normalization of the curve in Fig. 2, with node-branches marked

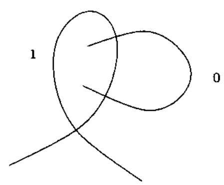


FIGURE 4. A "half-dimensional representation" of the nodal curve in Fig. 2

labelled with their genera), and the edges correspond to nodes. An example is given in Fig. 5.

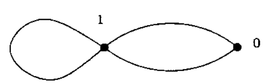


FIGURE 5. The dual graph of the curve in Fig. 2

The *arithmetic genus* of Σ, denoted by $p_a(\Sigma)$, is the genus of a "smoothing" of Σ (*not* the genus of the normalization). For example, the genus of the nodal curve in Fig. 2 is 3; its smoothing is shown in Fig. 6. Of course, the arithmetic genus of a non-singular curve coincides with the topological genus.

EXERCISE 22.2.1. *Suppose Σ is a curve with δ nodes such that $\tilde{\Sigma}$ has n components of genera $g_1, \ldots, g_n$. Show that $p_a(\Sigma) = \sum(g_i - 1) + \delta + 1$.*

This exercise also extends the definition of arithmetic genus to the case where Σ is not necessarily connected.

A more algebraic definition of arithmetic genus is

$$p_a(\Sigma) = 1 - \chi(\Sigma, \mathcal{O}_\Sigma) = 1 - h^0(\Sigma, \mathcal{O}_\Sigma) + h^1(\Sigma, \mathcal{O}_\Sigma),$$

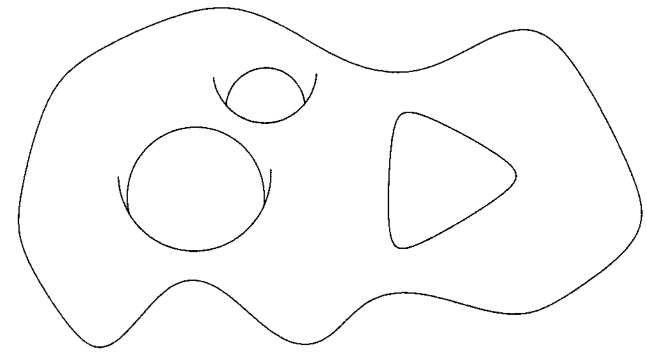


FIGURE 6. Smoothing of the nodal curve of Fig. 2

where $\mathcal{O}_\Sigma$ is the structure sheaf. Note that this definition also applies to disconnected curves.

Suppose the node-branches of the normalization $\nu : \tilde{\Sigma} \to \Sigma$ are $b_1, \ldots, b_{2\delta}$. Then the "normalization exact sequence"

$$0 \to \nu^* \mathcal{O}_\Sigma \to \mathcal{O}_{\tilde{\Sigma}} \to \oplus \mathcal{O}_{b_i} \to 0 \tag{22.1}$$

of sheaves on $\tilde{\Sigma}$ is a useful way to prove facts about nodal curves by reducing to the case of non-singular curves.

EXERCISE 22.2.2. *Using the related exact sequence*

$$0 \to \mathcal{O}_\Sigma \to \nu_* \mathcal{O}_{\tilde{\Sigma}} \to \oplus \mathcal{O}_{n_i} \to 0$$

of sheaves on Σ (where $n_1, \ldots, n_\delta$ are the nodes of Σ), redo Exercise 22.2.1 using the definition $p_a(\Sigma) = 1 - \chi(\Sigma, \mathcal{O}_\Sigma)$.

22.3. Differentials on Nodal Curves

A genus g non-singular curve has a g-dimensional vector space of (holomorphic) differentials. Define a differential on a nodal curve Σ to be a meromorphic differential w on each component satisfying:

(i) w is holomorphic away from the node-branches,
(ii) w has a pole of order at most 1 at each node-branch,
(iii) the residues of w at the two node-branches corresponding to a given node add to zero.

It is a fact that every connected genus g nodal curve has a g-dimensional vector space of differentials.

DEFINITION 22.3.1. *The* dualizing sheaf ω_Σ *of* Σ *is the sheaf of meromorphic differentials satisfying (i)–(iii).*

WARNING 22.3.1. Strictly speaking, sections of the dualizing sheaf should not be called differentials, as algebraic differentials have a different but related meaning.

CHAPTER 23

Moduli Spaces of Curves

23.1. Motivation: Projective Space as a Moduli Space

When studying objects of some sort (such as Riemann surfaces), it is helpful to construct a "moduli space" for such objects. For example, if one wants to study one-dimensional subspaces of a complex vector space V, one is naturally led to consider the projective space $\mathbb{P}V$ parametrizing such subspaces. Notice that $\mathbb{P}V$ is far more than the *set* of such subspaces. For example:

(1) It has the structure of a complex manifold, and even of an algebraic variety.
(2) It has natural cycles, homology, and cohomology classes, coming from the geometry of the subspaces. For example, for any subspace $W \subset V$, there is a cycle corresponding to one-dimensional subspaces contained in W. The homology class corresponding to W depends only on the dimension of W.
(3) There is a "universal family"

$$U \subset \mathbb{P}V \times V = \{([\ell], p) | p \in \ell\},$$

with a *structure morphism* $\pi : U \to \mathbb{P}V$ (that is the projection onto the first factor). The fiber of π above a point $[\ell]$ is the subspace in V corresponding to that point.

EXERCISE 23.1.1. *Verify that the universal family is a non-singular algebraic variety.*

(4) For every family F of one-dimensional subspaces parametrized by a variety B, there corresponds a map $a : B \to \mathbb{P}V$. Moreover, the family F can be recovered by pulling back the family U, that is, $F = a^*U$.

$\mathbb{P}V$ has another advantage: it is compact and non-singular — hence intersection theory is well defined. For example, if $\dim V = n + 1$ and H is the Poincaré dual of the locus of lines contained in a fixed codimension 1 subspace of V, then the fact $H^n = 1[pt]$ corresponds to the fact that there is one one-dimensional subspace contained in n generally chosen codimension 1 subspaces of V.

In the study of moduli of Riemann surfaces, or maps (as will later arise), our goal is threefold: (i) to construct a "reasonable" moduli space, (ii) to compactify it in a geometrically meaningful way, so that (iii) the resulting space is non-singular, or at least "not too singular".

23.2. The Moduli Space $\mathcal{M}_g$ of Non-singular Riemann Surfaces

Riemann surfaces of genus $g > 1$ have a well-behaved moduli space, denoted $\mathcal{M}_g$. As in the case of $\mathbb{P}V$, it is the set of non-singular genus g curves (up to isomorphism), endowed with additional geometric structure — in fact, $\mathcal{M}_g$ is a non-singular Deligne–Mumford stack.

In the case of $\mathcal{M}_g$, the isotropy group of the point $[\Sigma]$ (corresponding to the Riemann surface Σ) is the automorphism group of Σ. (Implicit here is the fact that the automorphism group of any Riemann surface Σ of genus $g > 1$ is finite.)

If $g > 2$, $\mathcal{M}_g$ is actually an orbifold. However, if $g = 2$, $\mathcal{M}_g$ isn't quite an orbifold, as every point has a non-trivial isotropy group — every genus 2 curve has a non-trivial automorphism. (Every genus 2 Riemann surface can be represented as a double cover of a line, and the involution of the double cover gives a non-trivial automorphism.)

EXERCISE 23.2.1. *Compute the dimension of* $\mathcal{M}_g$ *to be* $3g-3$ *as follows.*

(a) *Each genus g Riemann surface Σ has a g-dimensional family of line bundles of each degree d. For large d, the Riemann–Roch Theorem 22.1.1 tells us that any degree d line bundle has a $(d-g+1)$-dimensional vector space of sections, as h^1 vanishes. By choosing two general such sections s_0, s_1, we obtain a degree d genus g cover of $\mathbb{P}^1$. Compute the dimension of the space of such covers in terms of* $\dim \mathcal{M}_g$.

(b) *On the other hand, the Riemann–Hurwitz formula tells us that the general such cover has $2d + 2g - 2$ branch points, so we would expect the dimension of the space of such covers to be $2d + 2g - 2$ (corresponding to the independent motions of the branch point). Hence show that* $\dim \mathcal{M}_g = 3g - 3$.

REMARK 23.2.1. The above exercise can be extended to show that $\mathcal{M}_g$ is connected, by showing that one can connect two such covers by "moving the branch points".

$\mathcal{M}_g$ has a universal curve, which is best described in terms of "pointed curves", see Sec. 23.4.

WARNING 23.2.1. It is often said that genus 1 curves are parametrized by the j-line. More correctly, elliptic curves (which are genus 1 curves with the choice of a marked point as the identity in the group law) are (coarsely) parametrized by the j-line; the corresponding Deligne–Mumford stack is $\mathcal{M}_{1,1}$, described in Sec. 23.4.

23.3. The Deligne–Mumford Compactification $\overline{\mathcal{M}}_g$ of $\mathcal{M}_g$

In order to compactify $\mathcal{M}_g$ (for $g > 1$), we slightly extend the class of curves under consideration to include some nodal curves (see Sec. 22.2).

DEFINITION 23.3.1. A stable curve *is a connected nodal curve such that*

(i) *every irreducible component of geometric genus 0 has at least three node-branches,*

(ii) *every irreducible component of geometric genus 1 has at least one node-branch.*

Curves will be assumed to be connected unless otherwise noted.

Stability is equivalent to the condition that the automorphism group is finite; this is also true of all the stability conditions we'll see later. Note that stability can be quickly checked by looking at the dual graph of a curve.

EXERCISE 23.3.1. *Show that a stable curve must have genus at least* 2.

The moduli space of stable curves is a connected, irreducible, compact, non-singular Deligne–Mumford stack of dimension $3g - 3$.

23.3.1. Degenerations of Non-singular Curves. The space $\mathcal{M}_g$ of Riemann surfaces isn't compact, and there are good geometric reasons for suspecting this: we can visualize degenerations where the limit is nodal. For example, Fig. 1 shows a genus 3 curve that has degenerated into a nodal curve with a genus 2 component and a genus 1 component. Notice how the corresponding dual graph "degenerates", Fig. 2. Degenerations correspond to contracting loops on the Riemann surface. As another example, by contracting two loops in Fig. 6 of Ch. 22, we obtain the curve in Fig. 2 of that chapter.

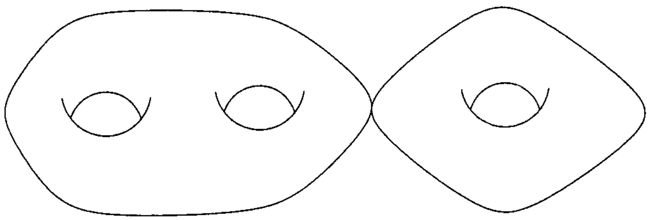


FIGURE 1. A degeneration of a genus 3 curve

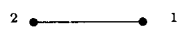


FIGURE 2. The dual graph degeneration corresponding to Fig. 1

WARNING 23.3.1. Fig. 3 shows a degeneration in which the limit curve is no longer stable (as it has, as a component, a sphere with only one node-branch). This may lead one to initially suspect that $\overline{\mathcal{M}}_g$, which is a moduli space of *stable* curves, is also not compact. However, it turns out that one can replace the limit of the degenerating family with a stable curve (in this case a non-singular genus 2 curve) — in fact, any such family has exactly one stable curve as a limit.

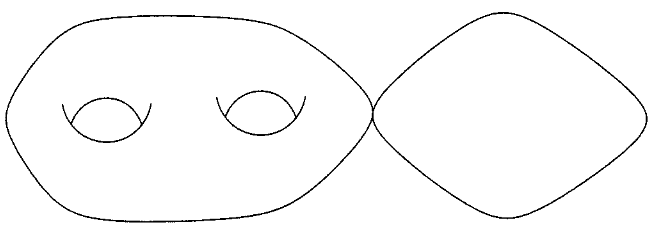


FIGURE 3. A degeneration with unstable limit

More precisely, suppose a family of stable curves over the punctured complex disc $|z| < 1$ is given. Then, perhaps after a base change $z \to z^r$, the family can be extended over 0 so that the curve mapping to 0 is also stable; such an extension is essentially unique.

23.4. The Moduli Spaces $\overline{\mathcal{M}}_{g,n}$ of Stable Pointed Curves

The constructions of Sec. 23.2 can be extended to give a compact, non-singular moduli space $\overline{\mathcal{M}}_{g,n}$ compactifying the moduli space $\mathcal{M}_{g,n}$ of n-pointed genus g Riemann surfaces.

DEFINITION 23.4.1. *An* n-pointed curve *is a nodal curve with* n *distinct labelled non-singular points. A* special point *of a component of a pointed curve is a point on the normalization of the component that is either a node-branch or (the pre-image of) a marked point (see Fig. 4 for an example). A pointed curve is* stable *if every genus 0 irreducible component has at least three special points, and every genus* 1 *irreducible component has at least one special point.*

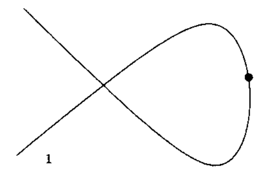


FIGURE 4. A marked curve of arithmetic genus 2 with three special points

The labels will often be taken to be $1, \ldots, n$, or $p_1, \ldots, p_n$. An n-pointed curve will be represented by a tuple, e.g. $(\Sigma, p_1, \ldots, p_n)$.

Once again, stability corresponds to having a finite automorphism group. In the *dual graph* of an n-pointed genus g curve, we use n labelled "tails" or half-edges to represent the marked points, as in Fig. 5.

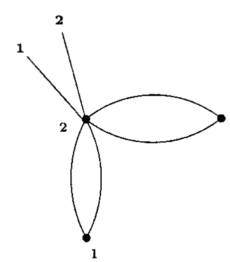


FIGURE 5. The dual graph Γ of a two-pointed genus 6 curve

EXERCISE 23.4.1. *Show that there are no stable n-pointed genus g curves if $2g-2+n \leq 0$. In addition to the cases $(g,n) = (0,0)$ and $(1,0)$ excluded in Exercise 23.3.1, this excludes the cases $(g,n) = (0,1)$ and $(0,2)$.*

The set of isomorphism classes of stable n-pointed genus g curves has a compact, non-singular moduli space, denoted $\overline{\mathcal{M}}_{g,n}$. It is irreducible, and has as an open set the moduli space $\mathcal{M}_{g,n}$ of n-pointed genus g Riemann surfaces. Also,

$$\dim \overline{\mathcal{M}}_{g,n} = \dim \overline{\mathcal{M}}_g + n = 3g - 3 + n$$

(each marked point gives one degree of freedom).

If $n_1 \geq n_2$, and $2g - 2 + n_2 > 0$, there is a forgetful morphism

$$\overline{\mathcal{M}}_{g,n_1} \to \overline{\mathcal{M}}_{g,n_2}.$$

Given a point $[(\Sigma, p_1, \dots, p_{n_1})] \in \overline{\mathcal{M}}_{g,n_1}$, the image point in $\overline{\mathcal{M}}_{g,n_2}$ is constructed by the following method. Consider first the pointed curve $(\Sigma, p_1, \dots, p_{n_2})$. If it isn't stable, then "contract" the "destabilizing" genus 0 components. There will never be a destabilizing genus 1 component — do you see why? Repeat this process until the curve is stable. An example of this stabilization process is shown in Fig. 6.

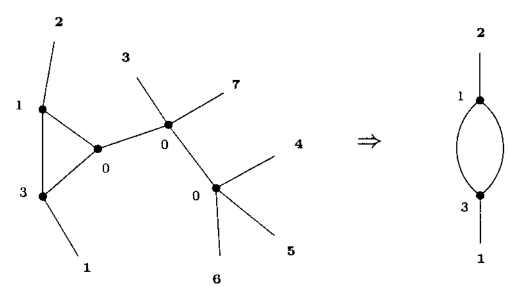


FIGURE 6. The stabilization process for the forgetful morphism $\overline{\mathcal{M}}_{4,7} \to \overline{\mathcal{M}}_{4,2}$

The morphism $\overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ can be identified with the universal curve over $\overline{\mathcal{M}}_{g,n}$. (In particular, $\overline{\mathcal{M}}_{g,1}$ is the universal curve over $\overline{\mathcal{M}}_g$.)

EXERCISE 23.4.2. *Suppose $(\Sigma, p_1, \dots, p_n)$ corresponds to a point of $\overline{\mathcal{M}}_{g,n}$.*

(a) *The point $p_1 \in \Sigma$ allegedly corresponds to a point of $\overline{\mathcal{M}}_{g,n+1}$, i.e., a stable $(n+1)$-pointed genus g curve. Which curve is it?*

(b) *Suppose Σ has a node N. Then $N \in \Sigma$ corresponds to a stable $(n+1)$-pointed genus g curve. Which curve is it?*

EXERCISE 23.4.3. *A new type of degeneration in $\overline{\mathcal{M}}_{g,n}$ comes up that did not arise in $\overline{\mathcal{M}}_g$, informally known as "bubbling". In both of the following cases, describe the "limit stable curve" in the family.*

(a) *Fix a stable pointed curve (Σ, p), and consider the family (Σ, p, q) where q is a point of Σ tending to p.*

(b) *Fix a stable nodal curve Σ, with node N, and consider the family (Σ, q) where q is a point of Σ tending to the node N.*

23.4.1. Boundary Strata. The "boundary" $\overline{\mathcal{M}}_{g,n} \setminus \mathcal{M}_{g,n}$ is as nice as one could hope: it consists of codimension 1 divisors intersecting transversely. The boundary is stratified by "dual graph type"; each stratum is non-singular. For example, the stratum S_Γ of $\overline{\mathcal{M}}_{6,2}$ corresponding to the dual graph Γ in Fig. 5 can be naturally identified with

$$(\mathcal{M}_{2,6} \times \mathcal{M}_{1,2} \times \mathcal{M}_{1,2}) / \operatorname{Sym}\Gamma,$$

where $\operatorname{Sym}\Gamma$ is the symmetry group of Γ ($\#\operatorname{Sym}\Gamma = 8$: the two genus 1 vertices can be switched, as can the two edges linking the genus 2 vertex with each genus 1 vertex).

EXERCISE 23.4.4. *Determine the action of Γ on $\mathcal{M}_{2,6} \times \mathcal{M}_{1,2} \times \mathcal{M}_{1,2}$ in the above example.*

EXERCISE 23.4.5. *Show that the codimension of a stratum S_β is the number of edges in the dual graph β, or equivalently the number of nodes of the general curve in S_β. For example, S_Γ has codimension 4 in $\overline{\mathcal{M}}_{6,2}$. (The normal bundle to a stratum at a point can be naturally identified; see Boundary Lemma 25.2.2.)*

The closure of each stratum is not in general non-singular. Here, by way of the explicit example of Γ, is a useful method of understanding the closure. There is a natural map

$$(\overline{\mathcal{M}}_{2,6} \times \overline{\mathcal{M}}_{1,2} \times \overline{\mathcal{M}}_{1,2}) / \operatorname{Sym}\Gamma \to \overline{\mathcal{M}}_{6,2} \tag{23.1}$$

which is a closed embedding on the open subset (where $\overline{\mathcal{M}}$ is replaced by $\mathcal{M}$ on the left side of the morphism). The left side has the advantage over $\overline{S}_\Gamma$ of being non-singular (as a Deligne–Mumford stack).

DEFINITION 23.4.2. *The boundary divisors will be named as follows.*

(a) *The divisor* Δ_0 *generically consists of irreducible curves with one node.*

(b) *If* $A_1 \coprod A_2 = \{1, \dots, n\}$ *and* $g_1 + g_2 = g$, *then the divisor* $D(g_1, A_1 | g_2, A_2)$ *consists generically of curves with two components, one of genus* g_1 *containing the marked points* A_1, *and the other of genus* g_2 *containing the marked points* A_2. *The genera may be omitted in* $D(0, A_1 | 0, A_2)$ *in genus 0.*

EXERCISE 23.4.6. *Write down the analogue of morphism (23.1) for all boundary divisors. For which boundary divisors is the general isotropy group non-trivial? Give an example of a codimension* 2 *stratum where* Δ_0 *has two branches.*

EXERCISE 23.4.7. *Show that* $\overline{\mathcal{M}}_{0,4}$ *is isomorphic to* $\mathbb{P}^1$ *(with the isomorphism given by the cross-ratio on* $\mathcal{M}_{0,4}$*). Describe the boundary divisors on* $\overline{\mathcal{M}}_{0,4}$. *As they are all points on* $\mathbb{P}^1$, *they are all homotopic, and even linearly equivalent. Hence prove that, on* $\overline{\mathcal{M}}_{0,n}$, *if* $\{i, j, k, l\} \subset \{1, \dots, n\}$,

$$\sum_{\substack{i,j \in A \\ k,l \in B}} D(A|B) \sim \sum_{\substack{i,k \in A \\ j,l \in B}} D(A|B) \sim \sum_{\substack{i,l \in A \\ j,k \in B}} D(A|B).$$

(In the sums, A *and* B *are varying, not* i, j, k, l.*)*

This equivalence on $\overline{\mathcal{M}}_{0,4}$ will be central to proving the WDVV equation, see Sec. 26.5.

CHAPTER 24

Moduli Spaces $\overline{\mathcal{M}}_{g,n}(X, \beta)$ of Stable Maps

The theory of moduli spaces of pointed curves predates Gromov–Witten theory, and indeed $\overline{\mathcal{M}}_{g,n}$ is one of the most studied objects in algebraic geometry. One of the key early developments in Gromov–Witten theory was Kontsevich's introduction of the *moduli space of stable maps*, a powerful generalization of $\overline{\mathcal{M}}_{g,n}$. In this section, we will define stable maps, describe the moduli space of such maps, and give some of its properties. A physical discussion appears in Sec. 16.4.

DEFINITION 24.0.3. *Let* X *be a non-singular projective variety. A morphism* f *from a pointed nodal curve to* X *is a* stable map *if every genus* 0 *contracted component of* Σ *(where* contracted *means mapping to a point) has at least three special points, and every genus* 1 *contracted component has at least one special point.*

As before, the genus 0 condition is the important one. Once again, stability corresponds to the condition that such a map has finite automorphism group. In order to make sense of "automorphism group", we need to define when two stable maps $(\Sigma, p_1, \dots, p_n, f)$ and $(\Sigma', p'_1, \dots, p'_n, f)$ are considered isomorphic. This is the case when there is an isomorphism

$$\tau : \Sigma \to \Sigma'$$

taking p_i to p'_i, with $f' \circ \tau = f$.

DEFINITION 24.0.4. *A stable map represents a homology class* $\beta \in H_2(X, \mathbb{Z})$ *if* $f_*[C] = \beta$.

DEFINITION 24.0.5. *The moduli space of stable maps from* n*-pointed genus* g *nodal curves to* X *representing the class* β *is denoted* $\overline{\mathcal{M}}_{g,n}(X, \beta)$. *The subscript* n *may be omitted if* $n = 0$.

The moduli space $\overline{\mathcal{M}}_{g,n}(X, \beta)$ is a Deligne–Mumford stack. However, we will see that $\overline{\mathcal{M}}_{g,n}(X, \beta)$ is not as well behaved as the moduli space of stable curves.

EXERCISE 24.0.8. *Show* $\overline{\mathcal{M}}_{g,n}(X,0) \cong \overline{\mathcal{M}}_{g,n} \times X$. *In particular, if* X *is a point, we recover the moduli space of stable curves.*

24.1. Example: The Grassmannian

A basic example is $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^m,1)$, which is the Grassmannian $\mathbb{G}(\mathbb{P}^1,\mathbb{P}^m)$ parametrizing lines in $\mathbb{P}^m$. If $n \geq 1$, $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m,1)$ is a locally trivial fibration over $\mathbb{G}(\mathbb{P}^1,\mathbb{P}^m)$ (with the "Fulton–MacPherson configuration space" $\mathbb{P}^1[n]$ as the fiber).

24.2. Example: The Complete (plane) Conics

As another example, we consider the space $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^2,2)$ in some detail. It is naturally stratified as follows. (i) An open set $\mathcal{M}_{0,0}(\mathbb{P}^2,2)$ is the space of non-singular conics, since for each such conic D there is an isomorphism $\mathbb{P}^1 \xrightarrow{\sim} D \subset \mathbb{P}^2$, unique up to equivalence. Hence $\mathcal{M}_{0,0}(\mathbb{P}^2,2)$ is naturally isomorphic to the projective space $\mathbb{P}^5$ of plane conics, minus the locus of singular conics. (ii) Singular conics D that are the unions of two (distinct) lines are similarly the isomorphic image $\Sigma \xrightarrow{\sim} D \subset \mathbb{P}^2$, where Σ is the union of two projective lines meeting transversally at a point. (iii) We also have maps from the same Σ to $\mathbb{P}^2$ sending each line in the domain onto the same line in $\mathbb{P}^2$. To determine this map up to isomorphism, however, the point that is the image of the intersection of the two lines must be specified, so the data for a point in this stratum is a line in $\mathbb{P}^2$ together with a point on it. (iv) Finally, there are maps from $\mathbb{P}^1$ that are branched coverings of degree 2 onto a line in the plane. These are determined by specifying the line together with the two distinct branch points.

EXERCISE 24.2.1. *Calculate the dimensions of the strata (i)–(iv) above, and determine which is in the closure of which.*

Thus we recover the classical space of *complete conics* — but in quite a different realization from the usual one. More precisely, the coarse moduli space (variety) of $\overline{\mathcal{M}}_{0,0}(\mathbb{P}^2,2)$ is the space of complete conics; the Deligne–Mumford stack is more refined, as strata (iii) and (iv) have isotropy groups $\mathbb{Z}/2$.

The same discussion is valid when $\mathbb{P}^2$ is replaced by $\mathbb{P}^m$, but this time the space is not the classical space of complete conics in $\mathbb{P}^m$. The classical space specifies a plane together with a complete conic contained in the plane; the space of stable maps "forgets" the data of this plane.

24.3. Seven Properties of $\overline{\mathcal{M}}_{g,n}(X,\beta)$

The moduli spaces $\overline{\mathcal{M}}_{g,n}(X,\beta)$ are in general quite ill behaved: possibly reducible, non-reduced, and of impure dimension.

EXERCISE 24.3.1 (Example of impure dimension). *Fix* $d > 1$ *and* $g > 0$*; we will consider* $\overline{\mathcal{M}}_{g,0}(\mathbb{P}^1,d)$. *One component consists generically of maps from non-singular curves; show that this component has dimension* $2d+2g-2$ *(see Exercise 23.2.1). Another component consists generically of maps from a nodal curve* $\Sigma_0 \cup \Sigma_1$*, where* Σ_0 *has genus 0 and maps with degree* d*, and* Σ_1 *has genus* g *and is collapsed to a point (and* Σ_0 *and* Σ_1 *meet at a point). Show that this component has dimension* $2d+3g-3$. *Why is the first component not in the closure of the second? Give an example of a map that lies in (the closure of) both components.*

However, these moduli spaces of stable maps do have some good geometric properties.

(1) There is an open subset (possibly empty) $\mathcal{M}_{g,n}(X,\beta)$ corresponding to maps from non-singular curves.

EXERCISE 24.3.2. *Give an example where this open subset is empty.*

(2) $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is compact.

EXERCISE 24.3.3. *Describe degenerations of stable maps in terms of "bubbling".*

(3) There are n "evaluation maps" $\mathrm{ev}_i : \overline{\mathcal{M}}_{g,n}(X,\beta)$ given by

$$\mathrm{ev}_i(\Sigma, p_1, \ldots, p_n, f) = f(p_i)$$

($1 \leq i \leq n$). (The ith evaluation of a map is the image of the ith point.)

(4) If $n_1 \geq n_2$, there is a "forgetful morphism"

$$\overline{\mathcal{M}}_{g,n_1}(X,\beta) \to \overline{\mathcal{M}}_{g,n_2}(X,\beta)$$

so long as the space on the right exists. (For example, there is no forgetful morphism $\overline{\mathcal{M}}_{0,7}(X,0) \to \overline{\mathcal{M}}_{0,0}(X,0)$, as the space on the right does not exist. However, there is a morphism to $\overline{\mathcal{M}}_{0,3}(X,0)$; this morphism is the projection $X \times \overline{\mathcal{M}}_{0,7} \to X$, see Exercise 24.0.8.)

(5) There is a "universal map" over the moduli space:

$$\begin{array}{ccc}(\tilde{\Sigma},\tilde{p}_1,\dots,\tilde{p}_n) & \xrightarrow{\tilde{f}} & X\\ \pi\downarrow & & \\ \overline{\mathcal{M}}_{g,n}(X,\beta). & & \end{array}$$

(Here $\tilde{p}_i$ are sections of the universal curve.) By a slight abuse of notation, the tildes are sometimes omitted.

EXERCISE 24.3.4. *Make explicit an identification of the universal curve with the moduli space $\overline{\mathcal{M}}_{g,n+1}(X,\beta)$ (see Exercise 23.4.2). In particular, identify the maps π and $\tilde{f}$.*

(6) Given a morphism $g : X \to Y$, there is an induced morphism

$$\overline{\mathcal{M}}_{g,n}(X,\beta) \to \overline{\mathcal{M}}_{g,n}(Y,g_*\beta)$$

so long as the space on the right exists. In case Y is a point, this gives the structure morphism to $\overline{\mathcal{M}}_{g,n}$, so long as $2g-2+n>0$.

(7) Under certain nice circumstances — if X is *convex*, to be defined in Definition 24.4.2 — $\overline{\mathcal{M}}_{0,n}(X,\beta)$ is non-singular of dimension

$$\int_\beta c_1(T_X) + \dim X + n - 3.$$

In the convex case, $\overline{\mathcal{M}}_{0,n}(X,\beta)$ has boundary divisors analogous to the divisors $D(0,A_1|0,A_2)$ on $\overline{\mathcal{M}}_{0,n}$ (see Sec. 23.4.1), similarly denoted

$$D(A_1,\beta_1|A_2,\beta_2),$$

where $A_1 \coprod A_2 = \{1,\dots,n\}$ and $\beta_1+\beta_2=\beta$. This divisor generically consists of maps where the source has two components, one with the points A_1, mapping to class β_1, and the other with the points A_2, mapping to class β_2. In positive genus and in the non-convex case, one can still make ("virtual") sense of this concept.

24.4. Automorphisms, Deformations, Obstructions

The deformation theory at a point x of a moduli space can often be interpreted in terms of cohomological data of the geometric object parametrized by x. In this section, we will put this intuition on firmer ground for the moduli space of maps. However, deformation theory is a rich and deep field, and we will barely scratch the surface.

As a motivating example, consider an immersion of a non-singular curve into a non-singular variety, $f:\Sigma\to X$. Consider the normal bundle exact sequence (of vector bundles on Σ):

$$0 \to T_\Sigma \to f^*T_X \to N_{\Sigma/X} \to 0.$$

Write the corresponding long exact sequence in cohomology as follows:

$$\begin{array}{ccccccc} & & 0 & \to & H^0(\Sigma,T_\Sigma) & \to\\ H^0(\Sigma,f^*T_X) & \to & H^0(\Sigma,N_{\Sigma/X}) & \to & H^1(\Sigma,T_\Sigma) & \to\\ H^1(\Sigma,f^*T_X) & \to & H^1(\Sigma,N_{\Sigma/X}) & \to & 0. & \end{array} \tag{24.1}$$

Recall that $H^2(\Sigma,T_\Sigma)=0$, by the dimensional vanishing Eq. (21.1).

We can reinterpret the terms of the long exact sequence as follows. The vector bundle T_Σ is related to the deformations of Σ itself: H^0 measures infinitesimal (first-order) automorphisms, H^1 measures infinitesimal deformations, and $H^2=0$ measures obstructions to deformations (all deformations are unobstructed in this case).

EXERCISE 24.4.1. *Using the Riemann–Roch Theorem 22.1.1, show that*

$$h^1(\Sigma,T_\Sigma)=3g-3$$

if $g>1$; this computation together with the vanishing of the obstruction space $H^2(\Sigma,T_\Sigma)$ implies $\mathcal{M}_g$ is non-singular of dimension $3g-3$.

The vector bundle $N_{\Sigma/X}$ is related to the deformations of the map f: $H^{-1}=0$ measures infinitesimal automorphisms, H^0 measures deformations, and H^1 measures obstructions.

The vector bundle f^*T_X is related to the deformations of the map f, where the structure of the source curve Σ is held fixed: $H^{-1}=0$ measures infinitesimal automorphisms, H^0 measures deformations, and H^1 measures obstructions.

Hence the exact sequence (24.1) can be rewritten as follows:

$$\begin{array}{ccccccc} & & 0 & \to & \mathrm{Aut}(\Sigma) & \to\\ \mathrm{Def}(f) & \to & \mathrm{Def}(\Sigma,f) & \to & \mathrm{Def}(\Sigma) & \to\\ \mathrm{Ob}(f) & \to & \mathrm{Ob}(\Sigma,f) & \to & 0. & \end{array}$$

(Again, $\mathrm{Aut}\,\Sigma$ refers to *infinitesimal* automorphisms of Σ. It is the Lie algebra of the automorphism group of Σ, although we are only interested in its vector space structure.) This exact sequence (suitably interpreted) is true for maps from pointed nodal curves in general:

The Deformation long exact sequence.

$$\begin{array}{lrcccc} & 0 & \to & \mathrm{Aut}(\Sigma,p_1,\ldots,p_n,f) & \to & \mathrm{Aut}(\Sigma,p_1,\ldots,p_n) \\ (24.2) & \mathrm{Def}(f) & \to & \mathrm{Def}(\Sigma,p_1,\ldots,p_n,f) & \to & \mathrm{Def}(\Sigma,p_1,\ldots,p_n) \\ & \mathrm{Ob}(f) & \to & \mathrm{Ob}(\Sigma,p_1,\ldots,p_n,f) & \to & 0. \end{array}$$

This is sometimes also called the tangent-obstruction exact sequence.

EXERCISE 24.4.2. *Interpret the exactness of the deformation long exact sequence at each step. For example, exactness at* $\mathrm{Def}(\Sigma,p_1,\ldots,p_n,f)$ *means that every deformation of a map induces a deformation of the pointed source curve, and those deformations that keep the pointed source curve fixed must come from a deformation only of* f.

As before, $\mathrm{Def}(f) = H^0(\Sigma, f^*T_X)$ and $\mathrm{Ob}(f) = H^1(\Sigma, f^*T_X)$; the 0 at the start of the sequence arises because $\mathrm{Aut}(f) = H^{-1}(\Sigma, f^*T_X) = 0$. If Σ is non-singular, then the role of T_Σ in the previous discussion is played by

$$T_\Sigma(-p_1 - \cdots - p_n),$$

the sheaf of holomorphic vector fields vanishing at the marked points. The 0 at the end of the sequence arises because deformations of nodal curves are unobstructed.

EXERCISE 24.4.3 (Stability and automorphisms).

(a) *Show that a marked nodal curve is stable if and only if it has no infinitesimal automorphisms.*

(b) *Show that a map from a marked nodal curve to X is stable if and only if it has no infinitesimal automorphisms. Hence for stable maps,*

$$\mathrm{Aut}(\Sigma,p_1,\ldots,p_n,f) = 0.$$

The deformation theory of stable maps is often obstructed — the moduli space $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is often singular. Nevertheless, we now compute the dimension of the moduli space in the unobstructed case, and find a criterion under which this assumption of unobstructedness holds.

As the alternating sum of dimensions in a long exact sequence is zero, the unobstructed deformation space of the map has dimension

$$h^0(\Sigma, f^*T_X) - h^1(\Sigma, f^*T_X) + \dim \mathrm{Def}(\Sigma,p_1,\ldots,p_n) - \dim \mathrm{Aut}(\Sigma,p_1,\ldots,p_n).$$

By the Riemann-Roch Theorem 22.1.1,

$$h^0(\Sigma, f^*T_X) - h^1(\Sigma, f^*T_X) = \int_\beta c_1(T_X) + (\dim X)(1-g).$$

It is also not hard to compute

$$\dim \mathrm{Def}(\Sigma,p_1,\ldots,p_n) - \dim \mathrm{Aut}(\Sigma,p_1,\ldots,p_n) = 3g-3+n.$$

Hence the expected dimension is

$$\int_\beta c_1(T_X) + (\dim X - 3)(1-g) + n. \tag{24.3}$$

As the above dimension is independent of the map f, it is an invariant of the moduli space.

DEFINITION 24.4.1. *Let* $\mathrm{vdim}\,\overline{\mathcal{M}}_{g,n}(X,\beta)$ *denote the expected (or virtual) dimension Eq. (24.3).*

From the deformation long exact sequence (24.2), a stable map $(\Sigma,p_1,\ldots,p_n,f)$ has obstruction space 0 if $h^1(\Sigma, f^*T_X) = 0$.

DEFINITION 24.4.2. X *is* convex *if* $h^1(\Sigma, f^*T_X) = 0$ *for every genus 0 stable map* $f : \Sigma \to X$.

Hence, in genus 0, the unobstructedness assumption holds for the class of convex varieties X. If X is convex, then $\overline{\mathcal{M}}_{0,n}(X,\beta)$ is a non-singular stack of dimension $\int_\beta c_1(T_X) + \dim X + n - 3$.

EXERCISE 24.4.4. *Suppose f is a genus 0 stable map $f : \Sigma \to X$ and f^*T_X is generated by global sections. Prove that $h^1(\Sigma, f^*T_X) = 0$.* (Hint: *Let V be the vector space of global sections, and $\mathcal{V}$ the trivial sheaf on Σ with fiber V. Consider the long exact sequence of*

$$0 \to \mathcal{U} \to \mathcal{V} \to f^*T_X \to 0,$$

where $\mathcal{U}$ is the kernel of the natural map $\mathcal{V} \to f^*T_X$.)

Therefore, if f^*T_X is generated by global sections, then X is convex. All algebraic homogeneous spaces are easily seen to be convex by the global generation of their tangent bundles — recall that a homogeneous space is a linear algebraic group modulo a parabolic subgroup. For example, projective spaces, Grassmannians, and flag varieties are all homogeneous spaces. We will repeatedly use the convexity of projective space:

LEMMA 24.4.3. *If $\Sigma \to \mathbb{P}^m$ is a map from a nodal genus 0 curve, then*

$$h^1(\Sigma, f^*T_{\mathbb{P}^m}) = 0.$$

EXERCISE 24.4.5 (Easy). *There are a few other cases where the moduli space is of the expected dimension.*

(a) *If $g = 0$ and $\beta = 0$, show that $h^1(\Sigma, f^*T_X) = 0$, and hence that $\overline{\mathcal{M}}_{0,n}(X,0)$ is unobstructed for arbitrary X.*
(b) *Show that $\overline{\mathcal{M}}_{g,n}(X,0)$ is unobstructed for all genera g only if X is a point.*

REMARK 24.4.1. The terms of the deformation long exact sequence can be defined cohomologically as follows:

$$\begin{aligned}
\mathrm{Aut}(\Sigma, p_1, \ldots, p_n) &= \mathrm{Hom}(\Omega_C(p_1 + \cdots + p_m), \mathcal{O}_C),\\
\mathrm{Def}(\Sigma, p_1, \ldots, p_n) &= \mathrm{Ext}^1(\Omega_C(p_1 + \cdots + p_m), \mathcal{O}_C),\\
\mathrm{Aut}(\Sigma, p_1, \ldots, p_n, f) &= \mathbb{H}\mathrm{om}(f^*\Omega_X \to \Omega_C(p_1 + \cdots + p_n), \mathcal{O}_C),\\
\mathrm{Def}(\Sigma, p_1, \ldots, p_n, f) &= \mathbb{E}\mathrm{xt}^1(f^*\Omega_X \to \Omega_C(p_1 + \cdots + p_n), \mathcal{O}_C),\\
\mathrm{Ob}(\Sigma, p_1, \ldots, p_n, f) &= \mathbb{E}\mathrm{xt}^2(f^*\Omega_X \to \Omega_C(p_1 + \cdots + p_n), \mathcal{O}_C),
\end{aligned}$$

where Ω is the sheaf of algebraic differentials, and $\mathbb{H}\mathrm{om}$ and $\mathbb{E}\mathrm{xt}$ are hypercohomology functors. We will avoid using these facts in our calculations.

CHAPTER 25

Cohomology Classes on $\overline{\mathcal{M}}_{g,n}$ and $\overline{\mathcal{M}}_{g,n}(X,\beta)$

Our next goal is to describe certain naturally defined cohomology classes on the moduli space of stable maps.

25.1. Classes Pulled Back from X

The easiest such classes are those pulled back from X: $\mathrm{ev}_i^*(\gamma)$, where $\gamma \in H^*(X)$. If γ is (Poincaré-dual to) a cycle Γ on X, then intuitively $\mathrm{ev}_i^*(\gamma)$ can be thought of as the locus of maps where the ith point maps to Γ. While the set $\mathrm{ev}_i^{-1}(\Gamma)$ may be of the wrong dimension, the class $\mathrm{ev}_i^*(\gamma)$ is always well-defined.

25.1.1. Recursions for Rational Plane Curves.

As an aside, we include here a recursion for rational plane curves due to Kontsevich and Ruan-Tian. This material is not a prerequisite to any later topic.

From Property **(7)** of Sec. 24.3, the space of degree d maps of rational curves to $\mathbb{P}^2$ is a compact orbifold of dimension $3d-1$. It can informally be thought of as parametrizing degree d rational plane curves. Hence there are a finite number of such curves through $3d-1$ general points in $\mathbb{P}^2$; let N_d be this number.

EXERCISE 25.1.1. *Interpret this number as*

$$N_d = \int_{\overline{\mathcal{M}}_{0,3d-1}(\mathbb{P}^2,d)} \mathrm{ev}_1^*(P) \cup \mathrm{ev}_2^*(P) \cup \cdots \cup \mathrm{ev}_{3d-1}^*(P) \tag{25.1}$$

where P is Poincaré dual to the point class. (Warning: *One still needs to check that the right hand side of the above equation counts maps correctly. This requires a transversality result, such as the Kleiman–Bertini theorem.*)

Clearly, $N_1 = 1$, the number of lines through two points. Starting from this trivial "base case", N_d is determined for $d \geq 2$ by a recursion formula:

THEOREM 25.1.1. *For $d > 1$,*

$$N_d = \sum_{\substack{d_1+d_2=d \\ d_1,d_2>0}} N_{d_1} N_{d_2} \left(d_1^2 d_2^2 \binom{3d-4}{3d_1-2} - d_1^3 d_2 \binom{3d-4}{3d_1-1} \right). \tag{25.2}$$

For example,

$$N_2 = 1,\ N_3 = 12,\ N_4 = 620,\ N_5 = 87304,\ N_6 = 26312976, \ldots.$$

Theorem 25.1.1 is a consequence of the WDVV equations in quantum cohomology, explained in Sec. 26.5. A proof by straightforward calculation is presented here. The strategy is to use fundamental linear relations among boundary components of $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^2, d)$, arising from the cross-ratio relations on $\overline{\mathcal{M}}_{0,4}$. We obtain Eq. (25.2) by restricting these linear equivalences to a curve Y in $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^2, d)$. This technique exactly produces the WDVV equations.

PROOF. *Step 1.* Consider the moduli space $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^2, d)$, where $n = 3d$ (*not* $3d-1$). Label the marked points by the set

$$\{1, 2, 3, \ldots, n-4, q, r, s, t\},$$

and consider the forgetful morphism

$$\pi : \overline{\mathcal{M}}_{0,n}(\mathbb{P}^2, d) \to \overline{\mathcal{M}}_{0,\{q,r,s,t\}}.$$

Recall the linear equivalence between the boundary divisors

$$D(\{q,r\}|\{s,t\}) \sim D(\{q,s\}|\{r,t\})$$

on $\overline{\mathcal{M}}_{0,\{q,r,s,t\}}$ (Exercise 23.4.7). The pull-back by π of this relation to $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^2, d)$ is

$$\pi^{-1}D(\{q,r\}|\{s,t\}) \sim \pi^{-1}D(\{q,s\}|\{r,t\}). \tag{25.3}$$

These pull-backs may be easily identified:

$$\pi^{-1}D(\{q,r\}|\{s,t\}) = \sum_{\substack{q,r\in A \\ s,t\in B \\ d_1+d_2=d}} D(A, d_1|B, d_2). \tag{25.4}$$

$$\pi^{-1}D(\{q,s\}|\{r,t\}) = \sum_{\substack{q,s\in A \\ r,t\in B \\ d_1+d_2=d}} D(A, d_1|B, d_2). \tag{25.5}$$

(This is analogous to part of Exercise 23.4.7.)

Let $z_1, \ldots, z_{n-4}, z_s, z_t$ be $n-2$ general points in $\mathbb{P}^2$ and let l_q, l_r be general lines. Define the curve $Y \subset \overline{\mathcal{M}}_{0,n}(\mathbb{P}^2, d)$ to be the intersection

$$Y = \mathrm{ev}_1^{-1}(z_1) \cap \cdots \cap \mathrm{ev}_{n-4}^{-1}(z_{n-4}) \cap \mathrm{ev}_q^{-1}(l_q) \cap \mathrm{ev}_r^{-1}(l_r) \cap \mathrm{ev}_s^{-1}(z_s) \cap \mathrm{ev}_t^{-1}(z_t).$$

Intuitively, Y parametrizes rational curves passing through the points $z_1, \ldots, z_{n-4}, z_s, z_t$, and with marked points on the lines l_q and l_r. (Do you see why it is of the "expected dimension" 1?)

Step 2. We compute the intersection of Y with a summand of Eq. (25.4). One can show that the points of this intersection correspond to maps from curves with (only) two components Σ_A and Σ_B.

The set $Y \cap D(A, 0|B, d)$ is non-empty only when $A = \{q, r\}$. In this case, the component Σ_A is required to map to the point $l_q \cap l_r$. If Σ_B is the other component (containing the marked points B), the restriction $f : \Sigma_B \to \mathbb{P}^2$ must map the $3d-2$ markings on Σ_B to the $3d-2$ given points, and in addition, f maps the point $\Sigma_A \cap \Sigma_B$ to $l_q \cap l_r$. Therefore,

$$\#Y \cap D(\{q,r\}, 0|\{1, \ldots, n-4, s, t\}, d) = N_d.$$

For $1 \le d_1 \le d-1$, $Y \cap D(A, d_1|B, d_2)$ is non-empty only when $|A| = 3d_1+1$. (Reason: a degree d_1 curve Σ_A can only be required to pass through at most $3d_1 - 1$ points, and a degree d_2 curve Σ_B can only be required to pass through at most $3d_2 - 1$ points. However, their union must pass through

$$(3d_1 - 1) + (3d_2 - 1) = n - 2$$

points $z_1, \ldots, z_{n-4}, z_s, z_t$. The two extra marked points in A are q and r, which are just required to map to lines in $\mathbb{P}^2$.)

There are $\binom{3d-4}{3d_1-1}$ partitions satisfying $q, r \in A$, $s, t \in B$, and $|A| = 3d_1 + 1$.

We now count the points of $Y \cap D(A, d_1|B, d_2)$. There are N_{d_1} choices for the image of Σ_A and N_{d_2} choices for the image of Σ_B. The points labelled q and r map to any of the d_1 intersection points of $f(\Sigma_A)$ with l_q and l_r respectively. Finally, there are $d_1 d_2$ choices for the image of the intersection point $\Sigma_A \cap \Sigma_B$ corresponding to the intersection points of $f(\Sigma_A) \cap f(\Sigma_B) \subset \mathbb{P}^2$. Thus

$$\#Y \cap D(A, d_1|B, d_2) = N_{d_1} N_{d_2} d_1^3 d_2. \tag{25.6}$$

The last case is simple: $Y \cap D(A, d|B, 0) = \emptyset$.

Therefore,

$$\#Y \cap \pi^{-1}D(\{q,r\}|\{s,t\}) = N_d + \sum_{\substack{d_1+d_2=d\\ d_1,d_2>0}} N_{d_1}N_{d_2}d_1^3d_2\binom{3d-4}{3d_1-1}.$$

Step 3. Now consider a summand of Eq. (25.5). $Y \cap D(A,d|B,0)$ and $Y \cap D(A,0|B,d)$ are both empty. For $1 \le d_1 \le d-1$, $Y \cap D(A,d_1|B,d_2)$ is non-empty only when $|A| = 3d_1$. There are $\binom{3d-4}{3d_1-2}$ such partitions and

$$\#Y \cap D(A,d_1|B,d_2) = N_{d_1}N_{d_2}d_1^2d_2^2$$

for each (by a similar calculation as for Eq. (25.6)). Therefore,

$$\#Y \cap \pi^{-1}D(\{q,s\}|\{r,t\}) = \sum_{\substack{d_1+d_2=d\\ d_1,d_2>0}} N_{d_1}N_{d_2}d_1^2d_2^2\binom{3d-4}{3d_1-2}.$$

Step 4. Finally, the linear equivalence Eq. (25.3) gives

$$\#Y \cap \pi^{-1}D(\{q,r\}|\{s,t\}) = \#Y \cap \pi^{-1}D(\{q,s\}|\{r,t\}),$$

and Kontsevich's recursion follows immediately. □

It is very enlightening to follow this argument through explicitly in a special case:

EXERCISE 25.1.2. *In the case $d=3$, describe the points of*

$$Y \cap \pi^{-1}D(\{q,r\}|\{s,t\}) \quad \textit{and} \quad Y \cap \pi^{-1}D(\{q,s\}|\{r,t\}).$$

25.2. The Tautological Line Bundles $\mathbb{L}_i$, and the Classes ψ_i

DEFINITION 25.2.1. *At each point $[\Sigma, p_1, \dots, p_n, f]$ of $\overline{\mathcal{M}}_{g,n}(X,\beta)$ (or each point $[\Sigma, p_1,\dots,p_n]$ of $\overline{\mathcal{M}}_{g,n}$), the cotangent line to Σ at point p_i is a one-dimensional vector space; these spaces "patch together" to give a line bundle $\mathbb{L}_i$, called the ith* tautological line bundle. *Define $\psi_i := c_1(\mathbb{L}_i)$.*

Thus ψ_i is a complex codimension 1 (real codimension 2) cohomology class.

EXERCISE 25.2.1. *What is ψ_1 on the one-dimensional space $\overline{\mathcal{M}}_{0,4} \cong \mathbb{P}^1$?*

The ψ-classes arise naturally in the geometry of the moduli of curves and maps. Consider the boundary divisor $D(g_1, A|g_2, B)$ on $\overline{\mathcal{M}}_{g=g_1+g_2,n}$,

where $A \coprod B = \{1,\dots,n\}$. It is best to describe the divisor in terms of the map

$$i : (\overline{\mathcal{M}}_{g_1,A\cup\{p\}} \times \overline{\mathcal{M}}_{g_2,B\cup\{q\}})/G \to \overline{\mathcal{M}}_{g,n}$$

(where G is the trivial group except in the case $g_1 = g_2$ and $n=0$, in which case $G = \mathbb{Z}/2$), as described in Sec. 23.4.1. Let $\overline{\mathcal{M}}$ be the source of the map i.

The "normal bundle" N_i to this immersion is defined to be the cokernel of the injection $T_{\overline{\mathcal{M}}} \to i^*T_{\overline{\mathcal{M}}_{g,n}}$. The normal bundle can be expressed in terms of the intrinsic geometry of (the curves parametrized by) $\overline{\mathcal{M}}$, through the following important lemma:

LEMMA 25.2.2 (Boundary Lemma). $N_i \cong (\mathbb{L}_p \otimes \mathbb{L}_q)^*$.

It is perhaps better to write the boundary lemma as $N_i \cong (p_1^*\mathbb{L}_p \otimes p_2^*\mathbb{L}_q)^*$, or $(\mathbb{L}_p \boxtimes \mathbb{L}_q)^*$.

EXERCISE 25.2.2. *The moduli space $\overline{\mathcal{M}}_{0,5}$ is a non-singular variety. (It has no non-trivial orbifold structure, as the automorphism group of every stable 5-pointed genus 0 curve is trivial.) Prove that each component of the boundary of $\overline{\mathcal{M}}_{0,5}$ is a (-1)-curve, i.e., is isomorphic to $\mathbb{P}^1$, with normal bundle $\mathcal{O}_{\mathbb{P}^1}(-1)$.* (Hint: *Use the previous exercise.*) *In fact, $\overline{\mathcal{M}}_{0,5}$ is isomorphic to $\mathbb{P}^2$ blown up at 4 points.*

EXERCISE 25.2.3. *Essentially the same story is true for the boundary divisor Δ_0 (although here there will always be a $\mathbb{Z}/2$-action), and also for any boundary stratum (although here the normal bundle is not necessarily rank 1). Figure out the details. (The notation Δ_0 was introduced in Definition 23.4.2.)*

The ψ-classes don't "commute" with forgetful morphisms. In other words, if $\pi : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ is the forgetful map, then $\psi_i \neq \pi^*\psi_i$ (where $1 \le i \le n$, and the ψ_i's are classes on different spaces). The difference is due to the "bubbling" phenomenon. Precisely:

LEMMA 25.2.3 (Comparison Lemma). *If $1 \le i \le n$,*

$$\psi_i - \pi^*\psi_i = [D(0,\{i,n+1\}|g,\{1,\dots,\hat{\imath},\dots,n\})].$$

EXERCISE 25.2.4. *Prove the Comparison lemma.*
Hint: *First show that $\pi^*\mathbb{L}_i$ can be identified with $\mathbb{L}_i$ on*

$$\overline{\mathcal{M}}_{g,n+1} \setminus [D(0,\{i,n+1\}|g,\{1,\dots,\hat{\imath},\dots,n\})].$$

EXERCISE 25.2.5. *Express ψ_1 explicitly as a sum of boundary divisors on $\overline{\mathcal{M}}_{0,n}$.*
Hint: *On $\overline{\mathcal{M}}_{0,3}$, ψ_1 is necessarily trivial. Pull back to $\overline{\mathcal{M}}_{0,n}$ using the Comparison Lemma $n-3$ times.*

EXERCISE 25.2.6. *Prove the "string equation for $\overline{\mathcal{M}}_{g,n}$":*

$$\int_{\overline{\mathcal{M}}_{g,n+1}} \psi_1^{\beta_1}\cdots\psi_n^{\beta_n} = \sum_{i=1}^{n}\int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{\beta_1}\cdots\psi_i^{\beta_i-1}\cdots\psi_n^{\beta_n}.$$

Thus intersections in ψ-classes, where one point doesn't "take part", can be reduced to intersections on the space of curves with one less point.

EXERCISE 25.2.7. *Prove the "dilaton equation for $\overline{\mathcal{M}}_{g,n}$":*

$$\int_{\overline{\mathcal{M}}_{g,n+1}} \psi_1^{\beta_1}\cdots\psi_n^{\beta_n}\psi_{n+1} = (2g-2+n)\int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{\beta_1}\cdots\psi_n^{\beta_n}$$

if $2g-2+n>0$.

The previous two exercises are special cases of the string and dilaton equation for maps, see Sec. 26.3.

EXERCISE 25.2.8. *Use the string equation to prove that if $\beta_1+\cdots+\beta_n = n-3$, then*

$$\int_{\overline{\mathcal{M}}_{0,n}} \psi_1^{\beta_1}\cdots\psi_n^{\beta_n} = \binom{n-3}{\beta_1,\ldots,\beta_n}.$$

Hence verify the dilaton equation in genus 0.

The genus 0 moduli spaces $\overline{\mathcal{M}}_{0,n}$ are essentially combinatorial objects; this is not true for higher genus.

EXERCISE 25.2.9. *Show that any integral*

$$\int_{\overline{\mathcal{M}}_{1,n}} \psi_1^{\beta_1}\cdots\psi_n^{\beta_n}$$

can be computed using the string and dilaton equations from the base case $\int_{\overline{\mathcal{M}}_{1,1}}\psi_1 = 1/24$.

DEFINITION 25.2.4. *For convenience, we will use Witten's notation*

$$\langle\tau_{\beta_1}\cdots\tau_{\beta_n}\rangle_g = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{\beta_1}\cdots\psi_n^{\beta_n}. \tag{25.7}$$

EXERCISE 25.2.10. *Show that for $g>0$, any integral $\langle\tau_{\beta_1}\cdots\tau_{\beta_n}\rangle_g$ can be computed using the string and dilaton equations knowing a finite number of base cases. Show that the number of base cases required is the number of partitions of $3g-3$.*

25.2.1. Aside: Witten's Conjecture (Kontsevich's Theorem) and Virasoro Constraints. All ψ-integrals can be efficiently computed using Witten's conjecture, proved by Kontsevich. The natural generating function for the genus g ψ-integrals described above is

$$F_g = \sum_{n\geq 0}\frac{1}{n!}\sum_{k_1,\ldots,k_n} t_{k_1}\cdots t_{k_n}\langle\tau_{k_1}\cdots\tau_{k_n}\rangle_g.$$

Then summing up over all genera $F(t,\lambda)=\sum F_g\lambda^{2g-2}$, we obtain Witten's *free energy of a point*.

The first system of differential equations conjectured by Witten are the KdV equations. Let $F(t)=F(t,\lambda=1)$. The KdV equations for $F(t)$ may be written in the following simple form. First, define the functions:

$$\langle\langle\tau_{k_1}\tau_{k_2}\cdots\tau_{k_n}\rangle\rangle = \frac{\partial}{\partial t_{k_1}}\frac{\partial}{\partial t_{k_2}}\cdots\frac{\partial}{\partial t_{k_n}}F.$$

Note that $\langle\langle\tau_{k_1}\tau_{k_2}\cdots\tau_{k_n}\rangle\rangle|_{t_i=0} = \langle\tau_{k_1}\tau_{k_2}\cdots\tau_{k_n}\rangle$. Then the KdV equations are equivalent to the set of equations for $n\geq 1$:

$$\begin{aligned}(2n+1)\langle\langle\tau_n\tau_0^2\rangle\rangle &= \\ \langle\langle\tau_{n-1}\tau_0\rangle\rangle\langle\langle\tau_0^3\rangle\rangle &+ 2\langle\langle\tau_{n-1}\tau_0^2\rangle\rangle\langle\langle\tau_0^2\rangle\rangle + \frac{1}{4}\langle\langle\tau_{n-1}\tau_0^4\rangle\rangle.\end{aligned} \tag{25.8}$$

As an example, consider (25.8) for $n=3$ evaluated at $t_i=0$. We obtain:

$$7\langle\tau_3\tau_0^2\rangle_1 = \langle\tau_2\tau_0\rangle_1\langle\tau_0^3\rangle_0 + \frac{1}{4}\langle\tau_2\tau_0^4\rangle_0.$$

Repeated use of the string equation yields:

$$7\langle\tau_1\rangle_1 = \langle\tau_1\rangle_1 + \frac{1}{4}\langle\tau_0^3\rangle_0.$$

Hence, we conclude $\langle\tau_1\rangle_1 = 1/24$. Equation (25.8) and the string equation together determine all the products (25.7) and thus uniquely determine $F(t)$.

The string and dilaton equations may be written as differential operators annihilating $e^{F(t,\lambda)}$ in the following way. Define

$$\begin{aligned} L_{-1} &= -\frac{\partial}{\partial t_0} + \frac{\lambda^{-2}}{2} t_0^2 + \sum_{i=0}^{\infty} t_{i+1} \frac{\partial}{\partial t_i}, \\ L_0 &= -\frac{3}{2}\frac{\partial}{\partial t_1} + \sum_{i=0}^{\infty} \frac{2i+1}{2} t_i \frac{\partial}{\partial t_i} + \frac{1}{16}. \end{aligned}$$

EXERCISE 25.2.11. *Prove that the string equation (Exercise 25.2.6) and* $\langle \tau_0^3 \rangle_0 = 1$ *imply the equation* $L_{-1} e^F = 0$.

EXERCISE 25.2.12. *Prove that the dilaton equation (Exercise 25.2.7) and* $\langle \tau_1 \rangle_1 = 1/24$ *imply the equation* $L_0 e^F = 0$.

Witten's conjecture and the string equation also formally imply that $e^{F(t,\lambda)}$ is annihilated by a sequence of differential operators corresponding to part of the Virasoro algebra, beginning with L_{-1} and L_0 described above (the "point Virasoro theorem"). One of the fundamental open questions in the field is the *Virasoro conjecture* (due to Eguchi, Hori, and Xiong, as well as S. Katz), which generalizes Virasoro constraints from $\overline{\mathcal{M}}_{g,n}$ to $\overline{\mathcal{M}}_{g,n}(X,\beta)$.

25.3. The Hodge Bundle $\mathbb{E}$, and the Classes λ_k

Recall from Sec. 22.3 that each nodal curve of arithmetic genus g has a canonical g-dimensional vector space of holomorphic differentials (or, more precisely when the curve is nodal, sections of the dualizing sheaf ω).

DEFINITION 25.3.1. *These rank g vector spaces "patch together" to give a rank g vector bundle $\mathbb{E}$, called the* Hodge bundle, *on* $\overline{\mathcal{M}}_{g,n}$ *and* $\overline{\mathcal{M}}_{g,n}(X,\beta)$. *Define the λ-classes by* $\lambda_k = c_k(\mathbb{E})$.

Thus λ_k is a complex codimension k (real codimension $2k$) cohomology class. Unlike the $\mathbb{L}_i$, $\mathbb{E}$ pulls back well under forgetful morphisms, including the moduli morphism $\overline{\mathcal{M}}_{g,n}(X,\beta) \to \overline{\mathcal{M}}_{g,n}$.

EXERCISE 25.3.1. *Prove this, by explicitly showing that "bubbling does not affect sections of the dualizing sheaf".*

25.4. Other Classes Pulled Back from $\overline{\mathcal{M}}_{g,n}$

If $2g-2+n>0$, then any class on $\overline{\mathcal{M}}_{g,n}$ can be pulled back to $\overline{\mathcal{M}}_{g,n}(X,\beta)$. For example, it is often helpful to pull back boundary strata. This trick was used in the proof of Theorem 25.1.1.

CHAPTER 26

The Virtual Fundamental Class, Gromov–Witten Invariants, and Descendant Invariants

In the previous chapter, several cohomology classes on the space of stable maps, $\overline{\mathcal{M}}_{g,n}(X,\beta)$, were constructed. In Sec. 25.1.1, the moduli space $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^2,d)$ was shown to be non-singular and equidimensional, and thus supported a fundamental class. Although $\overline{\mathcal{M}}_{g,n}(X,\beta)$ may be quite ill behaved in general, this moduli space always supports a canonical "virtual fundamental class" of the "expected" dimension. Gromov–Witten invariants are defined by capping the cohomology classes against the virtual fundamental class of the space of stable maps.

In this chapter, we will discuss the virtual fundamental class, and then define Gromov–Witten and descendant invariants.

26.1. The Virtual Fundamental Class

Recall the definition of the "expected" or "virtual" (complex) dimension of $\overline{\mathcal{M}}_{g,n}(X,\beta)$, from Eq. (24.3):

$$\operatorname{vdim}\overline{\mathcal{M}}_{g,n}(X,\beta)=\int_\beta c_1(T_X)+(\dim X-3)(1-g)+n.$$

In case the target is a Calabi–Yau threefold (where $\mathcal{K}_X=\mathcal{O}_X$) and no marked points are taken ($n=0$), the virtual dimension vdim equals 0 for any genus g. This is one indication of the special role that Calabi–Yau threefolds play in Gromov–Witten theory.

It is a fundamental and highly non-trivial fact that the space of maps carries a *virtual fundamental class*, denoted $[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}$, which lies in the expected dimension $H_{2\,\mathrm{vdim}}(X,\mathbb{Q})$. The full construction of the virtual class will not be given here. In three special cases, however, the virtual class has a simple interpretation. These cases will be discussed below.

26.1.1. Special Case: The Moduli Space is Unobstructed. If the moduli space is unobstructed (that is, $\mathrm{Ob}(\Sigma,p_1,\dots,p_n,X)=0$ for all stable maps parametrized by the moduli space), then

$$[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\text{vir}} = [\overline{\mathcal{M}}_{g,n}(X,\beta)].$$

The virtual fundamental class is the ordinary fundamental class in this case. Examples include: (i) $g=0$, X convex, (ii) $g=0$, $\beta=0$, (iii) X is a point (see Sec. 24.4).

26.1.2. Special case: The Moduli Space is Non-singular. If the moduli space is non-singular, but not of the expected dimension, then the virtual fundamental class is the Euler class of a canonical obstruction bundle Ob. The fiber of Ob at the moduli point $(\Sigma, p_1, \ldots, p_n, f)$ is the obstruction space $\mathrm{Ob}(\Sigma, p_1, \ldots, p_n, f)$. Since the moduli space is non-singular, these obstruction spaces are of constant dimension and form a vector bundle. The virtual fundamental class is the Euler class:

$$[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\text{vir}} = \mathbf{e}(\mathrm{Ob}) \cap [\overline{\mathcal{M}}_{g,n}(X,\beta)].$$

An example of this special case can be found by studying maps to a $\mathbb{P}^1$ in a Calabi–Yau threefold (see for example Sec. 27.5). Another example is $\overline{\mathcal{M}}_{g,n}(X,0)$.

26.1.3. Special Case: $g=0$ and X is a Hypersurface. Suppose $g=0$, and X is a degree l hypersurface in $\mathbb{P}^m$. This case is crucial to the proof of the Mirror conjecture for hypersurfaces. The statements here can be generalized somewhat, to complete intersections in projective space, and to a certain extent to complete intersections in toric varieties.

Every stable map to X is naturally a stable map to $\mathbb{P}^m$, so there is an inclusion

$$i : \overline{\mathcal{M}}_{0,n}(X,d) \hookrightarrow \overline{\mathcal{M}}_{0,n}(\mathbb{P}^m,d). \tag{26.1}$$

WARNING 26.1.1. The "d" in $\overline{\mathcal{M}}_{0,n}(X,d)$ does not necessarily refer to a well-defined homology class on X. When $m>3$, by the Lefschetz hyperplane theorem, there is a unique homology class of X that pushes forward to the class d (times the class of a line) in $\mathbb{P}^m$. However, when $m=3$, this is not necessarily true. (Consider the case $l=3$, $d=1$, for example: there are 27 lines on the cubic surface, and no two are homologous.) In this case, $\overline{\mathcal{M}}_{0,n}(X,d)$ should be interpreted as the union of $\overline{\mathcal{M}}_{0,n}(X,\beta)$ over all classes β that push forward to the class d (times the class of a line) in $\mathbb{P}^m$.

In the next set of exercises, we work out the (virtual) dimensions of both sides of Eq. (26.1):

EXERCISE 26.1.1. *Prove that the dimension of the right side is* $d(m+1)+(m-3)+n$. Hint: *First use the Euler sequence (cf. Eq. (2.1))*

$$0 \to \mathcal{O}_{\mathbb{P}^m} \to \mathcal{O}_{\mathbb{P}^m}(1)^{m+1} \otimes V \to T\mathbb{P}^m \to 0 \tag{26.2}$$

to show that $\mathcal{K}_{\mathbb{P}^m} = \mathcal{O}_{\mathbb{P}^m}(-m-1)$, *and then use Eq. (24.3) and Sec. 26.1.1.*

EXERCISE 26.1.2 (Adjunction formula). *Suppose D is a non-singular divisor (a complex codimension 1 submanifold) of the complex manifold M. Use the exact sequence*

$$0 \to T_D \to T_M \to N_{D/M} \to 0$$

and the fact that $N_{D/M} \cong \mathcal{O}_M(D)|_D$ *to prove that* $\mathcal{K}_D = (\mathcal{K}_M(D))|_D$.

EXERCISE 26.1.3. *Use Eq. (24.3) and the adjunction formula to prove that*

$$\begin{aligned} \operatorname{vdim} \overline{\mathcal{M}}_{0,n}(X,d) &= d(m+1-l)+(m-1)-3+n \\ &= \dim \overline{\mathcal{M}}_{0,n}(\mathbb{P}^m,d) - (dl+1). \end{aligned}$$

EXERCISE 26.1.4 (The quintic threefold is Calabi–Yau). *Use the adjunction formula to show that a non-singular quintic threefold X in $\mathbb{P}^4$ is Calabi–Yau ($\mathcal{K}_X = \mathcal{O}_X$), and that a non-singular complete intersection of two cubics in $\mathbb{P}^5$ is also.*

Although the class $[\overline{\mathcal{M}}_{0,n}(X,d)]^{\text{vir}}$ is difficult to get hold of, we can identify $i_*[\overline{\mathcal{M}}_{0,n}(X,d)]^{\text{vir}}$ as a cycle class on $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m,d)$. This identification will suffice to compute various invariants of X. The virtual class (pushed forward) is the Euler class of a canonical rank $dl+1$ vector bundle on $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m,d)$.

If $f:\Sigma \to \mathbb{P}^m$ is a degree d map from an irreducible genus 0 curve (a Riemann sphere), then $f^*\mathcal{O}_{\mathbb{P}^m}(l) = \mathcal{O}_\Sigma(dl)$, so

$$H^0(\Sigma, f^*\mathcal{O}_{\mathbb{P}^m}(l)) = dl+1. \tag{26.3}$$

This dimension calculation remains valid if Σ is a nodal genus 0 curve. The rank $dl+1$ vector spaces (26.3) can be "patched together" to give a vector bundle on $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d)$, which we will denote $\pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)$. To be more precise, $\pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)$ is defined through the diagram

$$\begin{array}{ccc} \Sigma & \xrightarrow{f} & \mathbb{P}^m \\ \pi \downarrow & & \\ \overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d) & & \end{array}$$

of the "universal map" (property **(5)** of Sec. 24.3).

Suppose the hypersurface X is defined by the degree l equation $s = 0$. The form s can be interpreted as a section of $\mathcal{O}_{\mathbb{P}^m}(l)$. This section can be pulled back to a section $\pi_* f^* s$ of $\pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)$.

EXERCISE 26.1.5. *Show that the section $\pi_* f^* s$ vanishes precisely on the set of stable maps to X.*

The Euler class of $\pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)$ gives a codimension $dl+1$ homology class. If a section s of the bundle vanishes on a locus of the expected codimension $dl+1$, then its fundamental class (with appropriate multiplicities) can be identified with the Euler class. If the section vanishes on a locus not of the expected dimension, then one can still associate a "virtual class" of the expected dimension, supported on the zero-locus of s. (This is the idea of a localized Chern class. See Sec. 3.5 for an introduction to Chern classes.) Hence there is a canonical homology class

$$\mathrm{e}(\pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)) \cap [\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d)]$$

naturally associated to $\overline{\mathcal{M}}_{0,n}(X, d)$, of the "expected" dimension. This motivates the following:

THEOREM 26.1.1.

$$i_* \left[\overline{\mathcal{M}}_{0,n}(X,d)\right]^{\mathrm{vir}} = \mathrm{e}(\pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)) \cap \overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d).$$

This fundamental formula will allow us to reduce computations on the moduli space of stable maps to a hypersurface to computations on the moduli space of stable maps to $\mathbb{P}^m$ (which has the advantage of being a non-singular space with a torus action). Theorem 26.1.1 can be proven from the constructions of the virtual class, but we will take Theorem 26.1.1 as the definition of $i_* \left[\overline{\mathcal{M}}_{0,n}(X,d)\right]^{\mathrm{vir}}$.

26.1.4. Relations among the Special Cases, and Witten's Formula. In this section, we will work through some of the connections between the previous special cases, as well as a useful formula predicted by Witten. This will also lend insight into the virtual fundamental class, and will give some of the ideas which led to its original construction. To begin with, note that the second special case (Sec. 26.1.2) generalizes the first (Sec. 26.1.1).

Suppose now that X is a quintic hypersurface in $\mathbb{P}^4$. Let i be the inclusion

$$i : \overline{\mathcal{M}}_0(X, d) \hookrightarrow \overline{\mathcal{M}}_0(\mathbb{P}^4, d),$$

as in Sec. 26.1.3. As X is Calabi–Yau, the virtual dimension of $\overline{\mathcal{M}}_0(X,d)$ is 0 — so the degree of the virtual class is a number. Suppose, however, that $\overline{\mathcal{M}}_0(X,d)$ is a non-singular r-dimensional family of maps. Suppose further that they are all maps from non-singular spheres; denote this assumption by (†). Assumption (†) is essentially never satisfied, but it can be removed in good situations.

The terms of the deformation long exact sequence (24.2),

$$\begin{array}{ccccccc} & & 0 & \to & \mathrm{Aut}(\Sigma) & \to \\ H^0(\Sigma, f^*T_X) & \to & T\overline{\mathcal{M}}_0(X,d) & \to & \mathrm{Def}(\Sigma) & \to \\ H^1(\Sigma, f^*T_X) & \to & \mathrm{Ob}(\overline{\mathcal{M}}_0(X,d)) & \to & 0, & \end{array} \tag{26.4}$$

can be "patched together" to form an exact sequence of vector bundles on $\overline{\mathcal{M}}_0(X,d)$. As usual, we denote the vector bundles by their fibers. As $\dim(\mathrm{Aut}\,\Sigma) = 3$ and $\dim(\mathrm{Def}\,\Sigma) = 0$ for a Riemann sphere Σ, the ranks of the vector bundles in the sequence are 3, $r+3$, r, 0, r, and r respectively.

The two previous special cases determine the virtual fundamental class in this setting:

(i) from Sec. 26.1.2, the virtual class is the (codimension r) Euler class on $\overline{\mathcal{M}}_0(X,d)$ of the obstruction bundle $\mathrm{Ob}(\overline{\mathcal{M}}_0(X,d))$,

(ii) from Sec. 26.1.3, the push-forward of the virtual class is the (codimension $5d+1$) Euler class of the vector bundle $\pi_* f^* \mathcal{O}_{\mathbb{P}^4}(5)$ on $\overline{\mathcal{M}}_0(\mathbb{P}^4, d)$.

Also, by Witten's study of this geometry,

(iii) the virtual class is the (codimension r) Euler class of the vector bundle $H^1(\Sigma, f^*T_X)$ on $\overline{\mathcal{M}}_0(X,d)$.

We will now see why these three definitions agree.

First, as the rank r vector bundles $H^1(\Sigma, f^*T_X)$ and $\mathrm{Ob}(\overline{\mathcal{M}}_0(X,d))$ on $\overline{\mathcal{M}}_0(X,d)$ are isomorphic by Eq. (26.4), Witten's definition (iii) agrees with the definition (i) of Sec. 26.1.2.

We also see from Eq. (26.4) how to relax assumption (†): if

$$T\overline{\mathcal{M}}_0(X,d) \to \mathrm{Def}(\Sigma)$$

is always surjective, then Witten's formula still holds. As we are in genus 0, $\mathrm{Def}(\Sigma)$ parametrizes smoothings of the nodes of Σ (that is, it has dimension equal to the number of nodes of Σ). If for every map $\Sigma \to X$ from a nodal curve in the component of $\overline{\mathcal{M}}_0(X,d)$ in question, there are deformations smoothing each of the nodes separately, then $T\overline{\mathcal{M}}_0(X,d)$ surjects onto $\mathrm{Def}(\Sigma)$, and Witten's formula applies. This is the case, for example, in the proof of the Aspinwall–Morrison formula (Proposition 27.5.1), where the component of $\overline{\mathcal{M}}_0(X,d)$ can be identified with $\overline{\mathcal{M}}_0(\mathbb{P}^1,d)$; the surjectivity of $T\overline{\mathcal{M}}_0(X,d)$ onto $\mathrm{Def}(\Sigma)$ corresponds to the fact that the boundary divisors of $\overline{\mathcal{M}}_0(\mathbb{P}^1,d)$ meet transversely.

The rest of this section will be occupied with an explanation of why (i) and (ii) are the same.

Recall from Sec. 26.1.3 that $\overline{\mathcal{M}}_0(X,d)$ is the zero-locus of a section s of the rank $5d+1$ vector bundle $\pi_* f^*\mathcal{O}_{\mathbb{P}^4}(5)$ on $\overline{\mathcal{M}}_0(\mathbb{P}^4,d)$ induced by the equation defining X. For convenience, we will denote the non-singular space $\overline{\mathcal{M}}_0(\mathbb{P}^4,d)$ by M, the bundle $\pi_* f^*\mathcal{O}_{\mathbb{P}^4}(5)$ by E, and the zero-locus $\overline{\mathcal{M}}_0(X,d)$ of the section s of E by Z.

From Eq. (26.4) and the ensuing discussion, we have the short exact sequence of vector bundles of the top row of Eq. (26.5) below. The same analysis applies with X replaced by $\mathbb{P}^4$, giving the bottom row of Eq. (26.5) below, and there is a natural map from the top to the bottom.

$$\begin{array}{ccccccc} 0 \to & \mathrm{Aut}\,\Sigma & \to & H^0(\Sigma, f^*T_X) & \to & TZ & \to 0 \\ & \sim\downarrow & & \downarrow & & \downarrow & \\ 0 \to & \mathrm{Aut}\,\Sigma & \to & H^0(\Sigma, f^*T_{\mathbb{P}^4}) & \to & TM|_Z & \to 0. \end{array} \tag{26.5}$$

By restricting the exact sequence

$$0 \to T_X \to T_{\mathbb{P}^4}|_X \to \mathcal{O}_{\mathbb{P}^4}(5) \to 0$$

(see Exercise 26.1.2) to Σ and taking the long exact sequence in cohomology, we find

$$0 \to H^0(\Sigma, f^*T_X) \to H^0(\Sigma, f^*T_{\mathbb{P}^4}) \to E|_Z \to H^1(\Sigma, f^*T_X) \to 0$$

is an exact sequence of vector bundles (of rank $r+3$, $5d+4$, $5d+1$, and r respectively). Here we use the vanishing $h^1(\Sigma, f^*T_{\mathbb{P}^4}) = 0$ of Lemma 24.4.3.

By applying the Snake Lemma to Eq. (26.5) and the vector bundle isomorphism $H^1(\Sigma, f^*T_X) \cong \mathrm{Ob}\,Z$ described above, the sequence of vector bundles

$$0 \to TZ \to TM|_Z \xrightarrow{\psi_s} E|_Z \to \mathrm{Ob}\,Z \to 0 \tag{26.6}$$

is exact, where ψ_s is the central morphism.

We now focus our attention on

$$\psi_s : TM|_Z \to E|_Z$$

which is essential to understanding (and properly defining) the virtual fundamental class. Note that:

- The kernel of ψ_s can be identified with TZ.
- The cokernel of ψ_s can be identified with $\mathrm{Ob}\,Z$.
- ψ_s can be interpreted as differentiation. Note that there is no map $TM \to E$ in general; it is essential that Z is the zero-locus of a section s of E. (Can you make sense of ψ_s if Z is singular?)
- ψ_s is a morphism of vector bundles on $\overline{\mathcal{M}}_0(X,d)$ even if assumption (†) is removed.

To complete the identification of cases (i) and (ii), we use (a special case of) the excess intersection formula.

THEOREM 26.1.2 (Excess intersection formula). *Suppose M is a non-singular space, E is a vector bundle on M, Z is the non-singular vanishing locus of a global section s of E. There is a natural "differentiation" morphism*

$$\psi_s : TM|_Z \to E|_Z$$

with kernel TZ, and cokernel denoted $\mathrm{Ob}\,Z$ (the "excess bundle"). Then

$$\mathbf{e}(E) = i_*\mathbf{e}(\mathrm{Ob}\,Z)$$

where $i : Z \hookrightarrow M$ is the inclusion and $\mathrm{Ob}\,Z = \mathrm{Coker}\,\psi_s$.

The following exercise motivates the excess intersection formula.

EXERCISE 26.1.6. *Suppose X is a compact manifold, with two submanifolds Y_1 and Y_2 of complementary codimension (so they "should" intersect in a finite number of points). Suppose Z is a connected component of $Y_1 \cap Y_2$*

that is a manifold. Define the "excess" bundle $\mathcal{E}$ on Z (of rank $\dim Z$*) so that the following sequence is exact:*

$$0 \to TZ \to TY_1|_Z \oplus TY_2|_Z \to TX|_Z \to \mathcal{E} \to 0.$$

Explain why $\mathrm{e}(\mathcal{E})$ *is the number of points Z should "count for" in the intersection $Y_1 \cap Y_2$. (It may help to assume that Y_1 and Y_2 can be deformed so as to intersect transversely.)*

To prove the Excess intersection formula, generalize the exercise so that $\mathcal{E}$ is of arbitrary rank, and apply it to the case where X is the total space of the bundle E over M, Y_1 is the zero section of E in X, and Y_2 is the section s of E.

26.2. Gromov–Witten Invariants and Descendant Invariants

The virtual fundamental class of the moduli space of maps may be paired against the cohomology classes defined earlier to obtain invariants of X as follows.

Given classes $\gamma_1, \ldots, \gamma_n$ in $H^*(X)$, the corresponding *Gromov–Witten invariant* is defined by:

$$\langle \gamma_1 \cdots \gamma_n \rangle^X_{g,\beta} := \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*(\gamma_1) \cup \cdots \cup \mathrm{ev}_n^*(\gamma_n) \tag{26.7}$$

The genus subscript will often be suppressed if $g = 0$. Intersections over $\overline{\mathcal{M}}_{g,n}(X,\beta)$ are called *genus g, n-point invariants.* The entries in $\langle \cdot \rangle$ are often written in product notation. Thus, for example, the numbers N_d computed in Sec. 25.1.1 are Gromov–Witten invariants, as Eq. (25.1) can be rewritten as

$$N_d = \langle [pt]^{3d-1} \rangle^{\mathbb{P}^2}_d. \tag{26.8}$$

EXERCISE 26.2.1. *Prove that $\langle \gamma_1\gamma_2\gamma_3 \rangle^X_0 = \int_X \gamma_1 \cup \gamma_2 \cup \gamma_3$. Thus three-point invariants include all triple intersections in $H^*(X)$.*

EXERCISE 26.2.2. *Compute all the Gromov–Witten invariants of a point. (Almost all vanish for trivial dimensional reasons.)*

A generalization of the Gromov–Witten invariants, the *gravitational descendant invariants* or *descendant invariants*, are defined by:

(26.9)
$$\langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_n}(\gamma_n) \rangle^X_{g,\beta} := \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^*(\gamma_1) \cup \psi_1^{a_1} \cup \cdots \cup \mathrm{ev}_n^*(\gamma_n) \cup \psi_n^{a_n},$$

where $\gamma_i \in H^*(X)$ and the a_i are non-negative integers. (This combines earlier notation, reflecting the fact that descendant invariants couple Gromov–Witten invariants (26.7) with topological gravity (25.7).)

As usual, the invariants are defined to vanish unless the dimension of the integrand is correct. For simplicity, $\tau_0(\gamma)$ will often be denoted by γ in Eq. (26.9). Again, the genus subscript will often be suppressed if $g = 0$.

26.3. String, Dilaton, and Divisor Equations for $\overline{\mathcal{M}}_{g,n}(X,\beta)$

Let the map

$$\nu : \overline{\mathcal{M}}_{g,n+1}(X,\beta) \to \overline{\mathcal{M}}_{g,n}(X,\beta)$$

be the forgetful morphism, forgetting the last point. (Recall that this morphism exists when the space on the right exists.) Three basic equations hold for descendant invariants: the string, dilaton, and divisor equations. They apply when ν exists, and the class assigned to the last marking is of total codimension 0 or 1. In these formulas, any term with a negative exponent on a cotangent line class is defined to be 0.

I. *The string equation.* Let $T_0 \in H^*(X)$ be the unit:

$$\langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_n}(\gamma_n) T_0 \rangle_{g,\beta} = \sum_{i=1}^n \langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_{i-1}}(\gamma_{i-1}) \tau_{a_i - 1}(\gamma_i) \tau_{a_{i+1}}(\gamma_{i+1}) \cdots \tau_{a_n}(\gamma_n) \rangle_{g,\beta}.$$

II. *The dilaton equation.*

$$\langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_n}(\gamma_n) \tau_1(T_0) \rangle_{g,\beta} = (2g - 2 + n) \langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_n}(\gamma_n) \rangle_{g,\beta}.$$

The string and dilaton equations specialize to the versions described in Sec. 25.2 by taking X to be a point.

III. *The divisor equation.* Let $\gamma \in H^2(X)$. Then

$$\begin{aligned} \langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_n}(\gamma_n) \gamma \rangle_{g,\beta} &= \left(\int_\beta \gamma \right) \langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_n}(\gamma_n) \rangle_{g,\beta} \\ &+ \sum_{i=1}^n \langle \tau_{a_1}(\gamma_1) \cdots \tau_{a_{i-1}}(\gamma_{i-1}) \tau_{a_i - 1}(\gamma_i \cup \gamma) \tau_{a_{i+1}}(\gamma_{i+1}) \cdots \tau_{a_n}(\gamma_n) \rangle_{g,\beta}. \end{aligned} \tag{26.10}$$

In the case of Gromov–Witten invariants (where no ψ-classes occur, or equivalently where only τ_0's appear), notice that the divisor equation has a simple intuitive interpretation. The second summand doesn't occur, and the equation informally says that the number of maps of a certain sort in class β, with an additional marked point p required to map to a divisor γ, is precisely the number of maps of that sort, times the number of choices $\gamma \cap \beta$ of where the point p could map.

EXERCISE 26.3.1. *Compute all the Gromov–Witten invariants of* $\mathbb{P}^1$. *Note (as in Exercise 26.2.2) that most invariants vanish trivially for dimensional reasons.*

EXERCISE 26.3.2. *Assuming Eq. (26.8), compute all the* genus 0 *Gromov–Witten invariants of* $\mathbb{P}^2$, *in terms of* N_d.

The proofs of equations I–III rely upon the analogue of the comparison lemma 25.2.3. In the convex case, we have:

LEMMA 26.3.1 (Comparison lemma for genus 0 stable maps to convex X). *If* X *is convex,* $\pi : \overline{\mathcal{M}}_{0,n+1}(X,\beta) \to \overline{\mathcal{M}}_{0,n}(X,\beta)$ *is the forgetful map, and* $1 \leq i \leq n$, *then*

$$\psi_i = \pi^*\psi_i + D(\{i, n+1\}, 0|\{1, \dots, \hat{i}, \dots, n\}, \beta).$$

EXERCISE 26.3.3. *Derive Equations I–III in the special case* $g = 0$, X *convex. What properties must the virtual class have for your argument to hold?*

EXERCISE 26.3.4. *How should this generalize to the higher genus or non-convex case?*

26.4. Descendant Invariants from Gromov–Witten Invariants in Genus 0

In genus 0, the descendant integrals actually carry no more information than the Gromov–Witten invariants:

PROPOSITION 26.4.1 (Genus 0 descendant reconstruction, Dubrovin). *The genus 0 descendants of* X *can be uniquely reconstructed from the genus 0 Gromov–Witten invariants.*

We prove this in the convex case; the general case is essentially identical , given a good answer to Exercise 26.3.4.

The key idea is:

LEMMA 26.4.2 ("ψ_1 is boundary in genus 0"). *If* $n \geq 3$, *then* $\psi_1 = \sum_\Gamma D_\Gamma$, *where the sum is over all boundary divisors with point splitting separating* 1 *from* $\{2, 3\}$.

PROOF. Consider the forgetful morphism

$$\nu : \overline{\mathcal{M}}_{0,n}(X,\beta) \to \overline{\mathcal{M}}_{0,3},$$

forgetting all data except the first three markings. By a comparison result analogous to Lemmas 25.2.3 and 26.3.1, $\psi_1 - \nu^*(\psi_1)$ is equivalent to a linear combination of boundary divisors of $\overline{\mathcal{M}}_{0,n}(X,\beta)$. Since ψ_1 is 0 in $H^2(\overline{\mathcal{M}}_{0,3})$, ψ_1 is a boundary class on $\overline{\mathcal{M}}_{0,n}(X,\beta)$. As in Exercise 25.2.5, the divisors that occur in $\psi_1 - \nu^*(\psi_1)$ are those with point splitting $A \cup B$ where $1 \in A$ and $\{2,3\} \subset B$. □

Using the above lemma together with the recursive structure of the boundary, we obtain a topological recursion relation among genus 0 descendant integrals. First, let $T_0, \dots, T_m$ be a basis of $H^*(X)$ (we assume here the cohomology is all even-dimensional to avoid signs). Let $g_{ij} = \int_X T_i \cup T_j$ be the intersection pairing, and let (g^{ij}) be the inverse matrix. The recursion relation is:

$$\begin{aligned}(26.11)\quad &\langle \tau_{a_1}(\gamma_1)\tau_{a_2}(\gamma_2)\tau_{a_3}(\gamma_3)\prod_{i\in S}\tau_{d_i}(\delta_i)\rangle_\beta = \\ &\sum \langle \tau_{a_1-1}(\gamma_1)\prod_{i\in S_1}\tau_{d_i}(\delta_i)T_e\rangle_{\beta_1} g^{ef}\langle T_f\tau_{a_2}(\gamma_2)\tau_{a_3}(\gamma_3)\prod_{i\in S_2}\tau_{d_i}(\delta_i)\rangle_{\beta_2}.\end{aligned}$$

The sum is over all stable splittings $\beta_1 + \beta_2 = \beta$, $S_1 \cup S_2 = S$, and over the diagonal splitting indices e, f; note that the class $\sum g^{ef} T_e \otimes T_f$ is the Poincaré dual of the diagonal $\Delta \subset X \times X$.

The proof of Proposition 26.4.1 follows easily from Eq. (26.11), by induction on the number of cotangent line classes. A descendant with no cotangent line classes is a Gromov–Witten invariant by definition. All $\beta = 0$ invariants are determined by the classical cohomology of X together with the formula for cotangent line class integrals on $\overline{\mathcal{M}}_{0,n}$ of Exercise 25.2.8. The topological recursion relations reduce descendants with at least three markings to integrals with fewer cotangent line classes. Let $\langle I \rangle_{\beta\neq 0}$ be a descendant integral with only two markings. Let H be an ample divisor on X. Add an extra marking subject to the divisor H condition: $\langle I \cdot H \rangle_\beta$. The divisor equation, Eq. (26.10), then relates $\langle I \rangle_\beta$ and $\langle I \cdot H \rangle_\beta$ modulo descendants with fewer cotangent lines. Since $\langle I \cdot H \rangle_\beta$ has three markings, Eq. (26.11) equates $\langle I \cdot H \rangle_\beta$ with an expression involving descendants with fewer cotangent lines. Similarly, if $\langle I\rangle_{\beta\neq 0}$ is an integral with only one marking, then consider $\langle I\cdot H\cdot H\rangle_\beta$. This completes the proof of Proposition 26.4.1. □

26.5. The Quantum Cohomology Ring

The *quantum cohomology ring*, a deformation of the usual cohomology ring, can be defined using Gromov–Witten invariants. This ring has played an important role in the history of the subject, but it will not be used later in these notes, so the reader may wish to skip this section on first reading. Some notation introduced here will be needed later, however.

For the rest of this section, the genus g will be assumed to be 0. Let $T_0 = 1 \in H^0(X,\mathbb{Z})$, and let $T_1,\dots,T_m$ be a homogeneous basis for the other cohomology groups, where $T_1,\dots,T_p$ are a homogeneous basis for the Kähler part $H^{1,1}(X,\mathbb{Z})$ of $H^2(X,\mathbb{Z})$, and T_m is Poincaré dual to T_0. (The Kähler classes will be used in the definition of the small quantum cohomology ring in Sec. 26.5.1.)

The $(\sum n_i)$-point Gromov–Witten invariant

$$\langle {T_1}^{n_1}\cdots {T_m}^{n_m}\rangle_\beta$$

is nonzero only when

$$\sum n_i\,(\operatorname{codim}(T_i)-1) = \dim X + \int_\beta c_1(T_X) - 3.$$

In this case, it is the (possibly virtual) number of pointed genus 0 maps meeting n_i general representatives of (the Poincaré dual of) T_i for each i.

As in the proof of Proposition 26.4.1, define the numbers g_{ij}, $0\le i$, $j\le m$, by the equations

$$g_{ij} = \int_X T_i\cup T_j$$

((g_{ij}) is the intersection matrix of $H^*(X)$), and (g^{ij}) as the inverse matrix to (g_{ij}). Note that

$$T_i\cup T_j = \sum_{e,f}\left(\int_X T_i\cup T_j\cup T_e\right) g^{ef}T_f,$$

and recall that $\int_X T_i\cup T_j\cup T_e$ is the Gromov–Witten invariant $\langle T_i\cdot T_j\cdot T_e\rangle_0^X$ (see Exercise 26.2.1).

We will define a "quantum deformation" of the usual cup multiplication

$$T_i\cup T_j = \sum_{e,f}\langle T_i\cdot T_j\cdot T_e\rangle_{\beta=0}^X g^{ef}T_f$$

by allowing nonzero classes β.

Let $\gamma = \sum t_i T_i$, where the t_i are supercommuting variables: if t_j and t_k correspond to odd cohomology classes, then $t_jt_k = -t_kt_j$.

Definition 26.5.1. *The* Gromov–Witten potential *or* quantum potential $C(\gamma) = C(t_0,\dots,t_m)$ *is a formal power series in* $\mathbb{Q}[[t]] = \mathbb{Q}[[t_0,\dots,t_m]]$ *given by*

$$C(t_0,\dots,t_m) = \sum_{n,\beta}\frac{\langle\gamma^n\rangle_\beta}{n!} = \sum_{n_0+\cdots+n_m\ge 3}\sum_\beta\langle {T_0}^{n_0}\cdots {T_m}^{n_m}\rangle_\beta\frac{{t_0}^{n_0}}{n_0!}\cdots\frac{{t_m}^{n_m}}{n_m!}.$$

The first sum is over (n,β) where $\langle\gamma^n\rangle_\beta$ is defined, i.e., $(n,\beta)\neq(0,0)$, $(1,0)$, $(2,0)$. (The Gromov–Witten potential is sometimes denoted Φ in the literature.)

Strictly speaking, a free variable should be included indexing the curve class β, for example q^β. For simplicity of notation, we will omit it. In any case, such a term will appear naturally later (see Eq. (26.13)).

Define C_{ijk} to be the partial derivative

$$C_{ijk} = \frac{\partial^3 C}{\partial t_i\,\partial t_j\partial t_k}.$$

Exercise 26.5.1. *Prove that*

$$C_{ijk} = \sum_{n\ge 0}\sum_\beta\frac{1}{n!}\langle\gamma^n\cdot T_i\cdot T_j\cdot T_k\rangle_\beta.$$

(This is just a formal manipulation.) In general, derivatives of the Gromov–Witten potential correspond to adding terms to the bracket.

We define a new "quantum" product $*$ by the rule:

$$T_i * T_j = \sum_{e,f} C_{ije}\,g^{ef}\,T_f. \qquad (26.12)$$

Extend the product in Eq. (26.12) $\mathbb{Q}[[t]]$-linearly to the $\mathbb{Q}[[t]]$-module $H^*(X)\otimes_{\mathbb{Q}}\mathbb{Q}[[t]]$, thus making it a $\mathbb{Q}[[t]]$-algebra. The product is commutative if X has no odd cohomology classes, or if we restrict to the even part of $H^*(X)$.

It is not difficult to see that $T_0 = 1$ is a unit for the $*$-product:

EXERCISE 26.5.2. *Show that*

$$C_{0jk} = \langle T_0 \cdot T_j \cdot T_k \rangle_0 = \int_X T_j \cup T_k = g_{jk},$$

and hence that $T_0 * T_j = \sum g_{je} g^{ef} T_f = T_j$.

The essential point, however, is the associativity of $*$:

THEOREM 26.5.2. *This definition makes* $H^*(X) \otimes \mathbb{Q}[[t]]$ *into an associative* $\mathbb{Q}[[t]]$*-algebra, with unit* T_0.

DEFINITION 26.5.3. *This ring* $H^*(X) \otimes \mathbb{Q}[[t]]$, *with the unusual quantum product structure, is called the* quantum cohomology ring, *or the* big quantum cohomology ring, *denoted* $QH^*(X)$.

Associativity is a formal consequence of divisor relations on $\overline{\mathcal{M}}_{0,4}$ (see Exer. 23.4.7). These associativity relations are called the *Witten–Dijkgraaf–Verlinde–Verlinde* (WDVV) equations.

Associativity is clearly equivalent to the statement that the coefficient of T_l in $((T_i * T_j) * T_k)$ is the same as the coefficient of T_l in $(T_i * (T_j * T_k))$, or equivalently

$$C_{ije} g^{ef} C_{fkl} = C_{ile} g^{ef} C_{fjk}.$$

This in turn is true if the coefficients of

$$\frac{t_0^{n_0}}{n_0!} \frac{t_1^{n_1}}{n_1!} \cdots \frac{t_m^{n_m}}{n_m!}$$

on both sides are equal. Let $n = \sum_{i=0}^m n_i$.

EXERCISE 26.5.3. *Show this, using the same strategy as the proof of Theorem 25.1.1. If you wish, assume that X is convex. Feel free to make assumptions about the virtual fundamental class, but make explicit what those assumptions are.*

Hint: *Consider* $\overline{\mathcal{M}}_{0,n+4}(X,\beta)$, *where the points will be called* p_{ab} ($0 \le a \le m$, $1 \le b \le n_a$), i, j, k, l. *Define, in analogy with the proof of Theorem 25.1.1, a one-dimensional homology class Y by intersecting pull-backs of evaluation maps* $\mathrm{ev}_{p_{ab}}^* T_a$ *with the virtual fundamental class of* $\overline{\mathcal{M}}_{0,n+4}(X,\beta)$. *Let ϕ be the forgetful map* $\overline{\mathcal{M}}_{0,n+4}(X,\beta) \to \overline{\mathcal{M}}_{0,\{i,j,k,l\}}$. *Restrict the equivalence* $\phi^*(D(ij|kl)) \sim \phi^*(D(jk|il))$ *to Y.*

EXERCISE 26.5.4. *Reinterpret the recursion (Theorem 25.1.1) as the WDVV equations for* $\mathbb{P}^2$.

26.5.1. The Small Quantum Cohomology Ring. There is also a "small" quantum cohomology ring, $QH_s^*(X)$, that incorporates only the three-point Gromov–Witten invariants in its product. $QH_s^*(X)$ is obtained by restricting the $*$-product to the parameters $T_1, \dots, T_p$ corresponding to the Kähler classes. Most computations of quantum cohomology rings in the literature have been of this small ring, which is often easier to describe.

It is simplest to define $QH_s^*(X)$ in the basis $T_0, \dots, T_m$. Define

$$\overline{C}_{ijk} := C_{ijk}(t_0, t_1, \dots, t_p, 0, \dots, 0) = \int_X T_i \cup T_j \cup T_k + \overline{\Gamma}_{ijk}.$$

To avoid convergence issues, for simplicity assume these cohomology classes are Poincaré dual to a basis of effective curve classes in X. The modified quantum potential $\overline{\Gamma}_{ijk}$ is given by

$$\overline{\Gamma}_{ijk} = \sum_{n \ge 0} \frac{1}{n!} \sum_{\beta \ne 0} \langle \gamma^n \cdot T_i \cdot T_j \cdot T_k \rangle_\beta$$

where $\gamma = t_1 T_1 + \cdots + t_p T_p$.

EXERCISE 26.5.5. *Use the divisor equation, Eq. (26.10), to show that*

$$\overline{\Gamma}_{ijk} = \sum_{\beta \ne 0} \langle T_i \cdot T_j \cdot T_k \rangle_\beta \, q_1^{\int_\beta T_1} \cdots q_p^{\int_\beta T_p}, \tag{26.13}$$

where $q_i = e^{t_i}$.

Note that only three-point invariants occur in Eq. (26.13).

The product

$$T_i * T_j = \sum_{e,f} \overline{C}_{ije} g^{ef} T_f = T_i \cup T_j + \sum_{e,f} \overline{\Gamma}_{ije} g^{ef} T_f \tag{26.14}$$

then makes the $\mathbb{Q}[q_1, \dots, q_p]$-module $H^*(X) \otimes_{\mathbb{Q}} \mathbb{Q}[q_1, \dots, q_p]$ into a commutative, associative $\mathbb{Q}[q_1, \dots, q_p]$-algebra with unit T_0. By Eq. (26.13), the small quantum cohomology is a deformation of $H^*(X)$.

26.5.2. Example: The Small Quantum Cohomology of $X = \mathbb{P}^m$. If T_i is the class of a linear subspace of codimension i, and β is d times the class of a line, then the number $\langle T_i \cdot T_j \cdot T_k \rangle_\beta$ can be nonzero only if $i + j + k = m + (m+1)d$ by an easy dimension count; this can happen only for $d = 0$ or $d = 1$, and in each case the number is 1. For simplicity, let $q = q_1$. It follows that

(i) if $i + j \le m$, then $T_i * T_j = T_{i+j}$,
(ii) if $m + 1 \le i + j \le 2m$, then $T_i * T_j = q T_{i+j-m-1}$.

Therefore the small quantum cohomology ring is:

$$QH_s^*(\mathbb{P}^m) = \mathbb{Q}[T, q]/(T^{m+1} - q), \tag{26.15}$$

where $T = T_1$ is the class of a hyperplane.

CHAPTER 27

Localization on the Moduli Space of Maps

We now introduce the techniques of torus localization on the moduli spaces of stable maps to $\mathbb{P}^m$. The torus action on $\mathbb{P}^m$ naturally lifts to an action on the space of maps, and integrals over the moduli space can be reduced (via the Localization formula described in Ch. 4) to integrals over the space of maps fixed by the torus. These integrals are much easier, and they can be combinatorially manipulated in genus 0.

We assume the reader is comfortable with the contents of Ch. 4. In order to apply these methods to $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d)$, we identify the fixed loci with decorated graphs, and compute their (equivariant) normal bundles. As an example, localization is used to prove the Aspinwall–Morrison formula for the contributions of genus 0 multiple covers of $\mathbb{P}^1 \subset X$, where X is Calabi–Yau. Next, we discuss the localization techniques in higher genus in the context of the virtual class ("virtual localization"). A sketch of the proof of the full multiple cover formula for $\mathbb{P}^1 \subset X$ is then given. We will later see a connection between the full multiple cover formula and the conjectural Gopakumar–Vafa invariants (Ch. 34).

27.1. The Equivariant Cohomology of Projective Space

In this section, we will establish facts about the equivariant cohomology of $\mathbb{P}^m$ that will later prove essential.

Let $\mathbb{T}$ be the complex torus $\mathbb{C}^* \times \cdots \times \mathbb{C}^*$ (where there are $m+1$ factors, indexed 0 through m). Suppose $\mathbb{T}$ acts on $V = \mathbb{C} \oplus \cdots \oplus \mathbb{C}$ ($m+1$ times) diagonally:

$$(t_0, \ldots, t_m) : (x_0, \ldots, x_m) \mapsto (t_0 x_0, \ldots, t_m x_m). \tag{27.1}$$

On $B\mathbb{T} = (\mathbb{CP}^\infty)^{m+1}$, let $\mathcal{L}_i$ be the pull-back of $\mathcal{O}_{\mathbb{P}^\infty}(1)$ from the ith factor (for $0 \le i \le m$). By definition α_i is $c_1(\mathcal{L}_i)$ in the standard presentation

$$H_{\mathbb{T}}^* = H^*((\mathbb{CP}^\infty)^{m+1}) = \mathbb{Q}[\alpha_0, \ldots, \alpha_m].$$

Note that α_i corresponds to the dual of the standard representation obtained by projecting $\mathbb{T}$ to the i^{th} factor. The vector bundle on $B\mathbb{T}$ corresponding to V is precisely $\mathcal{L}_0^* \oplus \cdots \oplus \mathcal{L}_m^*$.

The $\mathbb{T}$-action on V induces an action on $\mathbb{P}^m$. This corresponds to an isomorphism of $\mathbb{P}^m_{\mathbb{T}}$ with the projective bundle $\mathbb{P} := \mathbb{P}(\mathcal{L}_0^* \oplus \cdots \oplus \mathcal{L}_m^*)$ on $(\mathbb{CP}^\infty)^{m+1}$. From the splitting formula,

$$H^*_{\mathbb{T}}(\mathbb{P}^m) = H^*(\mathbb{P}) = \mathbb{Q}[H, \alpha_0, \ldots, \alpha_m] / \left(\prod_{i=0}^{m} (H - \alpha_i) \right)$$

where H is the hyperplane class (complex codimension 1) of the projective bundle $\mathbb{P}$.

Exercise 27.1.1. *Interpret the action (27.1) of $\mathbb{T}$ on V as a $\mathbb{T}$-action on the total space of the "tautological bundle" $\mathcal{O}_{\mathbb{P}^m}(-1)$. This induces a natural $\mathbb{T}$-action on $\mathcal{O}_{\mathbb{P}^m}(l)$ for all l, corresponding to the line bundle $\mathcal{O}_{\mathbb{P}}(lH)$ on the projective bundle $\mathbb{P}$ over $B\mathbb{T}$.*

Exercise 27.1.2. *Show that the weights of the $\mathbb{T}$-action (given in the previous exercise) on $H^0(\mathbb{P}^m, \mathcal{O}_{\mathbb{P}^m}(1))$ are $\alpha_0, \ldots, \alpha_m$. In other words, the vector space $H^0(\mathbb{P}^m, \mathcal{O}_{\mathbb{P}^m}(1))$ splits (as a $\mathbb{T}$-representation) into the one-dimensional $\mathbb{T}$-representations $\alpha_0, \ldots, \alpha_m$.*

Choosing a different action on the total space of $\mathcal{O}_{\mathbb{P}^m}(-1)$ would give a different $\mathbb{T}$-action on $\mathcal{O}_{\mathbb{P}^m}(l)$. For example, the action

$$(t_0, \ldots, t_m) : (x_0, \ldots, x_m) \mapsto (x_0, t_0^{-1} t_1 x_1, \ldots, t_0^{-1} t_m x_m),$$

corresponding to the line bundle $\mathcal{O}_{\mathbb{P}}(\alpha_0 - H)$ on $\mathbb{P}$, is a different linearization.

Let $p_0, \ldots, p_m$ be the fixed points of the $\mathbb{T}$-action on $\mathbb{P}^m$. Define $\phi_i \in H^{2m}_{\mathbb{T}}(\mathbb{P}^m)$ as the equivariant class of p_i $(0 \leq i \leq m)$.

Exercise 27.1.3. *Show that the weights of the (canonical) $\mathbb{T}$-action on $T_{\mathbb{P}^m}|_{p_i}$ are $\langle \alpha_i - \alpha_j \rangle$ $(j \neq i)$.*

Exercise 27.1.4. *Show that the weights of the $\mathbb{T}$-action on $H^0(\mathbb{P}^1, T\mathbb{P}^1)$ are $\alpha_0 - \alpha_1$, 0, and $\alpha_1 - \alpha_0$.*

For $a, b \in H^*_{\mathbb{T}}(\mathbb{P}^m)$, define the pairing

$$(a, b) := \int_{\mathbb{P}^m} (a \cup b) \in H^*_{\mathbb{T}} \in \mathbb{Q}[\alpha_0, \ldots, \alpha_m].$$

The following exercise contains important tools for working with the equivariant cohomology of projective space.

Exercise 27.1.5.

(a) *Show that $(\phi_i, f(H, \alpha)) = f(\alpha_i, \alpha)$.*

(b) *Show that $\phi_j = \prod_{i \neq j} (H - \alpha_i)$. Hence $(\phi_j, \phi_j) = \prod_{i \neq j} (\alpha_j - \alpha_i)$.*

(c) *Suppose $a, b \in H^*_{\mathbb{T}}(\mathbb{P}^m)$. Show that $a = b$ if and only if for all i, $(\phi_i, a) = (\phi_i, b)$.*

(d) The Bott residue formula. *Interpret the localization formula (4.4) in this case as*

$$\int_{\mathbb{P}^m} f(H, \alpha) = \sum_{\operatorname{Res} H = \alpha_0, \ldots, \alpha_m} \frac{f(H, \alpha)}{\prod (H - \alpha_i)} = \frac{1}{2\pi i} \oint dH \frac{f(H, \alpha)}{\prod (H - \alpha_i)}.$$

As an example, we compute the intersection $h \cdot h = 1$ on $\mathbb{P}^2$ using localization, where h is the class of a line in $\mathbb{P}^2$, so $h = c_1(\mathcal{O}_{\mathbb{P}^2}(1))$. We take the "natural linearization" H on $\mathcal{O}_{\mathbb{P}^2}(1)$ (see Exercises 27.1.1 and 27.1.2). The Bott residue formula gives

$$\begin{aligned} h \cdot h &= \sum_{\operatorname{Res} H = \alpha_0, \alpha_1, \alpha_2} \frac{H^2}{(H - \alpha_0)(H - \alpha_1)(H - \alpha_2)} \\ &= \frac{\alpha_0^2}{(\alpha_0 - \alpha_1)(\alpha_0 - \alpha_2)} + \frac{\alpha_1^2}{(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_0)} + \frac{\alpha_2^2}{(\alpha_2 - \alpha_0)(\alpha_2 - \alpha_1)} \\ &= 1. \end{aligned}$$

This computation required some almost magical cancellation. However, by choosing linearizations more carefully, we can make the computation much simpler. On the "first" copy of h (in $h \cdot h$), choose the linearization $H - \alpha_0$, and on the "second" copy, choose the linearization $H - \alpha_1$. Then the Bott residue formula gives

$$\begin{aligned} h \cdot h &= \sum_{\operatorname{Res} H = \alpha_0, \alpha_1, \alpha_2} \frac{(H - \alpha_0)(H - \alpha_1)}{(H - \alpha_0)(H - \alpha_1)(H - \alpha_2)} \\ &= 0 + 0 + 1. \end{aligned}$$

The moral of this story is that careful choice of linearization can drastically simplify localization calculations. Another example with this flavor is the proof of the Aspinwall–Morrison formula, Proposition 27.5.1.

Exercise 27.1.6. *Use localization to prove that $h^m = 1$ in $\mathbb{P}^m$, where h is the hyperplane class.*

27.2. Example: Branched Covers of $\mathbb{P}^1$

For future reference, we now compute the equivariant Chern classes (or weights) of various vector bundles. In other words, we decompose the vector bundles into $\mathbb{T}$-representations.

Consider the usual action of $\mathbb{T} = (\mathbb{C}^*)^2$ on $\mathbb{P}^1$, induced by

$$(t_0, t_1) : (x_0, x_1) \mapsto (t_0 x_0, t_1 x_1). \tag{27.2}$$

Suppose $f : \Sigma \to \mathbb{P}^1$ is a d-to-1 cover totally branched over 0 and ∞ (i.e., $(0, 1)$ and $(1, 0)$), and branched nowhere else.

EXERCISE 27.2.1. *Using the Riemann-Hurwitz formula, show that $\Sigma \cong \mathbb{P}^1$, and that the map is given by $(z_0, z_1) \mapsto (z_0^d, z_1^d) = (x_0, x_1)$.*

The weights of $\mathbb{T}$ on x_0 and x_1 (considered as elements of $H^0(\mathbb{P}^1, \mathcal{O}_{\mathbb{P}^1}(1))$) are α_0 and α_1 respectively. Hence the weights of $\mathbb{T}$ on z_0 and z_1 are α_0/d and α_1/d respectively. (To think of the fractional weight, it may help to assume that α_0 and α_1 are divisible by d.)

EXERCISE 27.2.2. *Show that the torus acts on $T_\Sigma f^{-1}(p_0)$ with weight $(\alpha_0 - \alpha_1)/d$, and on $T_\Sigma f^{-1}(p_1)$ with weight $(\alpha_1 - \alpha_0)/d$. The special case $d = 1$ was part of Exercise 27.1.3.*

If $\mathcal{O}_\Sigma(l)$ is a line bundle on Σ with some $\mathbb{T}$-action ($l \geq 0$), then the $\mathbb{T}$-action on the vector space $H^0(\Sigma, \mathcal{O}_\Sigma(l))$ can be easily found by the following useful method.

EXERCISE 27.2.3. *Let W_i be the weight of the $\mathbb{T}$-action on the one-dimensional vector space $\mathcal{O}_\Sigma(l)|_{p_i}$ (i = 0, 1). Show that the weights of the $\mathbb{T}$-action on $H^0(\Sigma, \mathcal{O}_\Sigma(l))$ are*

$$\frac{iW_0 + (l - i)W_1}{l}, \quad (0 \leq i \leq l).$$

As an example, the $\mathbb{T}$-weights on $H^0(\Sigma, T_\Sigma)$ are

$$(\alpha_0 - \alpha_1)/d,\ 0,\ (\alpha_1 - \alpha_0)/d,$$

which may be verified directly as well — the $d = 1$ case was Exercise 27.1.4.

EXERCISE 27.2.4. *Show that the weights of the $\mathbb{T}$-action on $H^0(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(1))$, with $\mathbb{T}$-action induced by the action (27.2) on $\mathcal{O}_{\mathbb{P}^1}(1)$ (corresponding to the equivariant class H, see Exercise 27.1.1), are*

$$\alpha_0 + \frac{i(\alpha_1 - \alpha_0)}{d}, \quad 0 \leq i \leq d.$$

We could, however, choose a different $\mathbb{T}$-action on $H^0(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(1))$ by choosing a different action on $\mathcal{O}_{\mathbb{P}^1}(1)$. For example, consider the action

$$(t_0, t_1) : (x_0, x_1) \mapsto (t_1^{-1} t_0 x_0, x_1) \tag{27.3}$$

on the total space of the tautological bundle $\mathcal{O}_{\mathbb{P}^1}(-1)$, corresponding to $\mathcal{O}_{\mathbb{P}}(-H + \alpha_1)$, inducing a new linearization on the dual bundle $\mathcal{O}_{\mathbb{P}^1}(1)$, corresponding to $\mathcal{O}_{\mathbb{P}}(H - \alpha_1)$.

EXERCISE 27.2.5. *Show that the weights of the $\mathbb{T}$-action on $H^0(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(1))$, with $\mathbb{T}$-action induced by the action Eq. (27.3) on $\mathcal{O}_{\mathbb{P}^1}(1)$ are*

$$\frac{i(\alpha_0 - \alpha_1)}{d}, \quad 0 \leq i \leq d.$$

We can compute the weights on other cohomology groups as well. For example, note that the vector space $H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1))$ has rank $d-1$ (by the Riemann–Roch theorem 22.1.1 or Serre duality, see Sec. 3.5.4). Consider the $\mathbb{T}$-action induced on it by (27.3). To compute the characters of this action, note that by Serre duality,

$$H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1)) \cong H^0(\Sigma, \omega_\Sigma \otimes f^*\mathcal{O}_{\mathbb{P}^1}(1))$$

canonically (where $\omega_\Sigma \cong \mathcal{O}_\Sigma(-2)$ is the sheaf of differentials on Σ, see Sec. 22.3, with $\mathbb{T}$-action induced by Eq. (27.2)).

EXERCISE 27.2.6. *Show that the weights of $H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1))$ are*

$$\frac{i}{d}(\alpha_0 - \alpha_1), \quad 1 \leq i \leq d - 1.$$

Hence the Euler class of the corresponding bundle on $(\mathbb{CP}^\infty)^m$ is

$$(d-1)!(\alpha_0 - \alpha_1)^{d-1}/d^{d-1}.$$

(Hint: *The weight at p_0 coming from ω_Σ is $(\alpha_1 - \alpha_0)/d$, and the weight at p_0 coming from $\mathcal{O}_{\mathbb{P}^1}(1)$ is $\alpha_0 - \alpha_1$.)*

The above exercise will be useful in proving the Aspinwall–Morrison formula (Proposition 27.5.1, see Exercise 27.5.1).

27.3. Determination of Fixed Loci

We now apply localization to the moduli space of stable maps $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d)$. Many of our preliminary comments will apply for arbitrary genus g (though the moduli space is singular), in order to set up our discussion of the virtual localization formula in Sec. 27.6.

Throughout, the $\mathbb{T}$-action on $\mathbb{P}^m$ will be taken to be the "standard" one described in Sec. 27.1. This induces a $\mathbb{T}$-action on $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$.

In order to apply localization, we first need to identify the fixed loci. This is not difficult, and the reader may prefer to do this on his or her own before reading on. Following Kontsevich, we will identify the components of the fixed point locus of the $\mathbb{T}$-action on $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$ with decorated graphs. See Fig. 1 for an example.

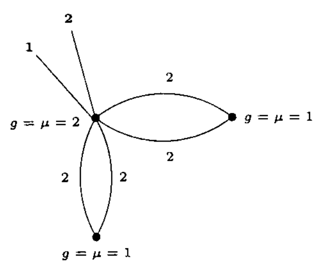


FIGURE 1. An example of a marked graph Γ corresponding to a component of $\overline{\mathcal{M}}_{6,2}(\mathbb{P}^2, 8)$ ($\#\operatorname{Aut}\Gamma = 8$, $\#\mathbb{A}_\Gamma = 2^4 \cdot 8$)

Let $f : (\Sigma, p_1, \ldots, p_n) \to \mathbb{P}^m$ be a $\mathbb{T}$-fixed stable map. The image of Σ is a $\mathbb{T}$-invariant curve in $\mathbb{P}V$, and the images of all marked points, nodes, contracted components, and ramification points are $\mathbb{T}$-fixed points. The $\mathbb{T}$-fixed points on $\mathbb{P}^m$ are $p_0, \ldots, p_m$, and the only invariant curves are the lines joining the points p_i. It follows that each non-contracted component of Σ must map onto one of these lines, and be ramified only over the two fixed points. By Exercise 27.2.1, such a component must be rational, and the map restricted to this component is completely determined by its degree.

We identify the components of the fixed locus with decorated graphs. The construction will be reminiscent of the dual graphs of Sec. 22.2. To an invariant stable map f, we associate a decorated graph Γ as follows. Γ has one edge for each non-contracted component. The edge e is marked with the degree $d(e)$ of the map from that component to its image line. Γ has one vertex for each connected component of $f^{-1}(\{p_0, \ldots, p_m\})$. Define the labelling map

$$\mu : \text{Vertices} \to \{0, \ldots, m\} \tag{27.4}$$

by $f(v) = p_{\mu(v)}$. The vertices have an additional labelling $g(v)$ by the arithmetic genus of the associated component. (If the component is a single point, we take the genus to be 0.) If $g = 0$, then each $g(v)$ is 0, so this labelling may be omitted. An edge is incident to a vertex if the two associated subsets of Σ are incident. Finally, Γ has n numbered "tails" or half-edges coming from the n marked points. These tails are attached to the appropriate vertex. (In figures, we will show the tail labels in bold face.)

The set of all invariant stable maps whose associated graph is Γ is naturally identified with a finite quotient of a product of moduli spaces of pointed curves. Define

$$\overline{\mathcal{M}}_\Gamma = \prod_{\text{vertices}} \overline{\mathcal{M}}_{g(v),\text{val}(v)}.$$

The valence of a vertex includes the markings/tails, as well as the incident edges. $\overline{\mathcal{M}}_{0,1}$ and $\overline{\mathcal{M}}_{0,2}$ are interpreted as points in this product. Let $\pi : \mathcal{C} \to \overline{\mathcal{M}}_\Gamma$ be the universal family of $\mathbb{T}$-fixed stable maps to $\mathbb{P}^m$ yielding a morphism

$$\gamma : \overline{\mathcal{M}}_\Gamma \to \overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d).$$

There is a natural automorphism group $\mathbb{A}_\Gamma$ acting on $\mathcal{C}$ and $\overline{\mathcal{M}}_\Gamma$. $\mathbb{A}_\Gamma$ is filtered by an exact sequence of groups

$$1 \to \prod_{\text{edges}} \mathbb{Z}/d(e) \to \mathbb{A}_\Gamma \to \operatorname{Aut}(\Gamma) \to 1$$

where $\operatorname{Aut}(\Gamma)$ is the automorphism group of Γ (as a decorated graph). $\operatorname{Aut}(\Gamma)$ naturally acts on $\prod_{\text{edges}} \mathbb{Z}/d(e)$ and $\mathbb{A}_\Gamma$ is the semidirect product. The induced map

$$\gamma/\mathbb{A}_\Gamma : \overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma \to \overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$$

is a closed immersion of Deligne–Mumford stacks. It should be noted that the subgroup $\prod_{\text{edges}} \mathbb{Z}/d(e)$ acts trivially on $\overline{\mathcal{M}}_\Gamma$ and that $\overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma$ is nonsingular. A component of the $\mathbb{T}$-fixed stack of $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$ is supported on $\overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma$.

For future use, we make the following definitions.

DEFINITION 27.3.1. *Define a flag F of the graph Γ to be an incident edge-vertex pair (e,v). Define a labelling map for flags $\mu(F) := \mu(v)$ (where μ is the labelling map on vertices, see Eq. (27.4)). Define*

$$\omega_F = \frac{\alpha_{\mu(F)} - \alpha_{\mu(v')}}{d(e)} \tag{27.5}$$

where v' is the other vertex of edge e. Let ψ_F denote the line bundle on $\overline{\mathcal{M}}_\Gamma$ whose fiber over a point is the cotangent space to the component associated to the flag F at the corresponding node (a ψ-class on the corresponding moduli space of pointed curves $\overline{\mathcal{M}}_{g(v),\mathrm{val}(v)}$ corresponding to the vertex v).

By Exercise 27.2.2, ω_F is the weight of the induced action of $\mathbb{T}$ on the tangent space to the rational component Σ_e of Σ corresponding to F at its preimage over p_v.

27.4. The Normal Bundle to a Fixed Locus

In order to apply the localization formula (4.4) in genus 0 (where the moduli space is a non-singular Deligne–Mumford stack), we must calculate the (equivariant) Euler class of the normal bundle N_Γ to each fixed locus. As usual, we do this by comparing the tangent bundle of each fixed locus $\overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma$ with the restriction of the tangent bundle of the entire moduli space $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m, d)$. It will be easier to perform these calculations on $\overline{\mathcal{M}}_\Gamma$ rather than $\overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma$; we must then remember to divide by $|\mathbb{A}_\Gamma|$ when applying localization.

Apply the deformation long exact sequence (24.2) for genus 0 stable maps to $\mathbb{P}^m$:

$$\begin{array}{ccccccc} 0 & \to & \mathrm{Aut}(\Sigma, p_1, \dots, p_n) & \to & \mathrm{Def}(f) & \to & \mathrm{Def}(\Sigma, p_1, \dots, p_n, f) \\ & \to & \mathrm{Def}(\Sigma, p_1, \dots, p_n) & \to & 0. & & \end{array} \tag{27.6}$$

Here we use the vanishing $h^1(\Sigma, f^*T\mathbb{P}^m) = 0$ of Lemma 24.4.3.

We consider the restriction of this exact sequence to the fixed locus $\overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma$ (or more correctly, the pull-back to $\overline{\mathcal{M}}_\Gamma$). If V is a vector bundle, denote its fixed weight-zero part by V^{fix}, and its moving (weight nonzero) part by V^{mov}. Note that $\mathrm{Def}(\Sigma, p_1, \dots, p_n, f)$ is the pull-back of the tangent bundle of $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^r, d)$, $\mathrm{Def}(\Sigma, p_1, \dots, p_n, f)^{\mathrm{fix}}$ is the tangent bundle to $\overline{\mathcal{M}}_\Gamma$, and $\mathrm{Def}(\Sigma, p_1, \dots, p_n, f)^{\mathrm{mov}}$ is the normal bundle N_Γ. In order to compute

$$\mathbf{e}(N_\Gamma) = \mathrm{Def}(\Sigma, p_1, \dots, p_n, f)^{\mathrm{mov}},$$

we compute the moving parts of the other bundles in the exact sequence (27.6), and combine them via

$$\mathbf{e}(N_\Gamma) = \frac{\mathbf{e}((\mathrm{Def}\, f)^{\mathrm{mov}})\mathbf{e}((\mathrm{Def}(\Sigma, p_1, \dots, p_n))^{\mathrm{mov}})}{\mathbf{e}((\mathrm{Aut}(\Sigma, p_1, \dots, p_n))^{\mathrm{mov}})} \tag{27.7}$$

The bundle $\mathrm{Aut}(\Sigma, p_1, \dots, p_n)$. We first determine the fixed automorphisms. For each non-contracted component of Σ, there is a torus-fixed piece coming from the infinitesimal automorphism of that component fixing the two special points. These automorphisms correspond to the Lie algebra of the $\mathbb{C}^*$'s that are the automorphism groups of the non-contracted $\mathbb{P}^1$'s, keeping the two preimages of 0 and ∞ fixed. (This fixed part will cancel with a similar weight-zero term in $\mathrm{Def}\, f$.)

There are, however, more automorphisms. They correspond to vertices v of valence 1, as in this case the point mapping to $\mu(v)$ (call it q_v) is not a special point (see Definition 23.4.1). Hence we have an additional automorphism corresponding to moving this point; this one-dimensional space is isomorphic (as a $\mathbb{T}$-representation) to the tangent space to Σ at q_v, i.e., ω_F (see Eq. (27.5)), where F is the (unique) flag incident to v. Thus

$$\mathbf{e}((\mathrm{Aut}(\Sigma, p_1, \dots, p_n))^{\mathrm{mov}}) = \prod_{\mathrm{val}(v)=1} \omega_F.$$

The bundle $\mathrm{Def}(\Sigma, p_1, \dots, p_n)$. A deformation of the contracted components (as marked curves) is a weight-zero deformation of the map that yields the tangent space $\mathrm{Def}(\Sigma, p_1, \dots, p_n, f)^{\mathrm{fix}}$ of $\overline{\mathcal{M}}_\Gamma/\mathbb{A}_\Gamma$ as a summand of $\mathrm{Def}(\Sigma)^{\mathrm{fix}}$. The other deformations of Σ come from smoothing nodes of Σ. This vector space splits into a product of spaces corresponding to deformations that smooth each node individually. By the boundary lemma 25.2.2, the one-dimensional space associated to each node is identified (as a line bundle) with the tensor product of the tangent spaces of the two components at the node.

(i) We deal first with the nodes that join contracted components to non-contracted components; these correspond to flags F where $\mathrm{val}(\mu(F)) \geq 3$. The tangent space to the non-contracted curve forms a trivial bundle with weight ω_F. The tangent space to the contracted curve has class $-\psi_F$ (see Definition 27.3.1). Thus the contribution from these nodes is

$$\prod_F (\omega_F - \psi_F)$$

where the product only contains factors where ψ_F is defined, that is, where $\mathrm{val}(\mu(F)) \geq 3$.

(ii) We next deal with nodes that join two non-contracted components; these correspond to vertices with valence 2, and no marked tails. By the same argument, these contribute

$$\prod_{\mathrm{val}(v)=2} (\omega_{F_{v,1}} + \omega_{F_{v,2}})$$

where $F_{v,1}$ and $F_{v,2}$ are the two flags incident to v. (The product is taken to only include factors that "make sense", that is, when v is incident to two edges, and hence has no marked tails.) Thus

$$\mathbf{e}(\mathrm{Def}(\Sigma, p_1, \dots, p_n)^{\mathrm{mov}}) = \prod_{\mathrm{flags}} (\omega_F - \psi_F) \prod_{\mathrm{val}(v)=2} (\omega_{F_{v,1}} + \omega_{F_{v,2}}).$$

The bundle $\mathrm{Def}\, f = H^0(\Sigma, f^*T\mathbb{P}^m)$. We consider the normalization exact sequence (22.1) resolving all of the nodes of Σ which are forced by the graph Γ:

$$0 \to \mathcal{O}_\Sigma \to \bigoplus_{\mathrm{vertices}} \mathcal{O}_{\Sigma_v} \oplus \bigoplus_{\mathrm{edges}} \mathcal{O}_{\Sigma_e} \to \bigoplus_{\mathrm{flags}} \mathcal{O}_{x_F} \to 0.$$

Tensoring with $f^*(T\mathbb{P}^m)$ and taking cohomology yields:

$$0 \to H^0(\Sigma, f^*T\mathbb{P}^m) \to \bigoplus_{\mathrm{vertices}} H^0(\Sigma_v, f^*T\mathbb{P}^m) \oplus \bigoplus_{\mathrm{edges}} H^0(\Sigma_e, f^*T\mathbb{P}^m) \to$$

$$\to \bigoplus_{\mathrm{flags}} T_{p_{\mu(F)}}\mathbb{P}^m \to 0$$

(here we use the vanishing $h^1(\Sigma, f^*T\mathbb{P}^m) = 0$ of Lemma 24.4.3). Also note that $H^0(\Sigma_v, f^*T\mathbb{P}^m) = T_{p_{\mu(v)}}\mathbb{P}^m$ since Σ_v is connected and f is constant on it. Hence, as representations (or in K-theory),

$$H^0(\Sigma, f^*T\mathbb{P}^m) = \begin{cases} + \bigoplus_{\mathrm{vertices}} T_{p_{\mu(v)}}\mathbb{P}^m + \bigoplus_{\mathrm{edges}} H^0(\Sigma_e, f^*T\mathbb{P}^m) \\ - \bigoplus_{\mathrm{flags}} T_{p_{\mu(F)}}\mathbb{P}^m \end{cases}$$

As non-contracted components are rigid, we see that $H^0(\Sigma_e, f^*T\mathbb{P}^m)$ is trivial as a bundle. We determine its weights using the Euler sequence on $\mathbb{P}^m$, Eq. (26.2):

$$0 \to \mathcal{O}_{\mathbb{P}^m} \to \mathcal{O}_{\mathbb{P}^m}(1) \otimes V \to T\mathbb{P}^m \to 0$$

where V is the vector space with $\mathbb{T}$-action such that $\mathbb{P}^m = \mathbb{P}V$. Pulling back to Σ_e and taking cohomology gives us

$$0 \to \mathbb{C} \to H^0(\Sigma_e, \mathcal{O}_{\Sigma_e}(d(e))) \otimes V \to H^0(\Sigma_e, f^*T\mathbb{P}^m) \to 0.$$

Here the weight on $\mathbb{C}$ is trivial. The weights on $H^0(\Sigma_e, \mathcal{O}_{\Sigma_e}(d(e)))$ are given by $\frac{a}{d(e)}\alpha_i + \frac{b}{d(e)}\alpha_j$ for $a + b = d(e)$ (by Exercise 27.2.4), and the weights on V are $-\alpha_0, \dots, -\alpha_m$, so the weights of the middle term are just the pairwise sums of these, $\frac{a}{d(e)}\alpha_i + \frac{b}{d(e)}\alpha_j - \alpha_k$. There are exactly two weight-zero terms here coming from $a = 0, k = j$ and $b = 0, k = i$; they cancel the weight-zero term from the $\mathbb{C}$ on the left, and the weight-zero term occurring in $\mathrm{Aut}(\Sigma)$. Breaking up the remaining terms into two groups corresponding to $k = i, j$ and $k \neq i, j$, we obtain the contribution of $H^0(\Sigma_e, f^*T\mathbb{P}^m)$ to the Euler class $\mathbf{e}(\mathrm{Def}(f)^{\mathrm{mov}})$:

$$(-1)^{d(e)} \frac{d(e)!^2}{d(e)^{2d(e)}} (\alpha_i - \alpha_j)^{2d(e)} \prod_{\substack{a+b=d(e)\\ k \neq i,j}} \left(\frac{a}{d(e)}\alpha_i + \frac{b}{d(e)}\alpha_j - \alpha_k \right).$$

Hence $1/\mathbf{e}(N_\Gamma)$ is given by:

$$\prod_{\mathrm{flags}} \frac{1}{\omega_F - \psi_F} \prod_{\nu \neq \mu(F)} (\alpha_{\mu(F)} - \alpha_\nu) \tag{27.8}$$

$$\prod_{\mathrm{vertices}} \prod_{\nu \neq \mu(v)} \frac{1}{\alpha_{\mu(v)} - \alpha_\nu} \prod_{\mathrm{val}(v)=2} \frac{1}{\omega_{F_{v,1}} + \omega_{F_{v,1}}} \prod_{\mathrm{val}(v)=1} \omega_F$$

$$\prod_{\mathrm{edges}} \frac{(-1)^{d(e)} d(e)^{2d(e)}}{(d(e)!)^2 (\alpha_i - \alpha_j)^{2d(e)}} \prod_{\substack{a+b=d(e)\\ k \neq i,j}} \frac{1}{\frac{a}{d(e)}\alpha_i + \frac{b}{d(e)}\alpha_j - \alpha_k}.$$

The following result will be used in the proof of the Aspinwall–Morrison formula in the next section.

EXERCISE 27.4.1. *Verify that for the one-edge graph given in Fig. 2,*

$$\frac{1}{\mathbf{e}(N^{\mathrm{vir}})} = \frac{(-1)^{d-1} d^{2d-2}}{(d!)^2 (\alpha_0 - \alpha_1)^{2d-2}}.$$

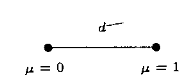


FIGURE 2. The one-edge fixed locus

27.5. The Aspinwall–Morrison Formula

A general quintic (Calabi–Yau) threefold X contains a finite number of lines, and all such lines have normal bundle $\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$. The normal bundle of any embedding $\mathbb{P}^1 \subset X$ is a rank 2 bundle of degree -2, as X has trivial canonical bundle and $\int_{\mathbb{P}^1} c_1(T_{\mathbb{P}^1}) = 2$. Since vector bundles on $\mathbb{P}^1$ split completely, and the "balanced" splitting is the most general, one may hope that a suitably general rational curve in X has balanced normal bundle. Embedded $\mathbb{P}^1$'s in a Calabi–Yau threefold (not necessarily lines) with normal bundle $\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$ are called *rigid*. The degree 1 (genus 0) Gromov–Witten invariant of X is the number of rigid lines. This number was classically known to be 2875.

The degree 2 Gromov–Witten invariant of a generic quintic has two contributions, one from the rigid conics, and one from double covers of the lines. Thus we see that even in the generic case, the moduli space $\overline{\mathcal{M}}_0(X, 2)$ is not equidimensional: it has 2875 components that are isomorphic to $\overline{\mathcal{M}}_0(\mathbb{P}^1, 2)$, of dimension 2. It is clearly important to calculate the contribution of these double covers. This question (and its generalization to higher degree) is answered by:

PROPOSITION 27.5.1 (Aspinwall–Morrison formula). *Let $\mathbb{P}^1 \subset X$ be a rigid curve of class β. The genus 0, degree d covers of $\mathbb{P}^1$ contribute $1/d^3$ to the genus 0, class $d\beta$ Gromov–Witten invariant of X.*

As described above, one naively expects all rational curves in X to be rigid, although this is not quite true (see Sec. 29.1.2).

In this section, we will use localization to prove the Aspinwall–Morrison formula. The first localization proof was given by Manin. The proof here uses an optimal localization choice leading to much simpler combinatorics (and, later, to computations of higher-genus contributions).

We first note that the contribution to the degree $d\beta$ Gromov–Witten invariant is naturally an integral over the non-singular space $\overline{\mathcal{M}}_0(\mathbb{P}^1, d)$, so we are in the situation of the special case of Sec. 26.1.2. Thus the desired contribution is the Euler class of the obstruction bundle. Hence to prove the Aspinwall–Morrison formula, we (i) identify the obstruction bundle, and then (ii) use localization to compute its Euler class.

We now identify the obstruction bundle. By Witten's formula of Sec. 26.1.4, the obstruction bundle is canonically $H^1(\Sigma, f^*T_X)$. (Recall that we are identifying bundles on $\overline{\mathcal{M}}_0(\mathbb{P}^1, d)$ by describing their fibers.) From the long exact sequence of

$$0 \to f^*T_{\mathbb{P}^1} \to f^*T_X \to f^*N_{\mathbb{P}^1/X} \to 0$$

and the vanishing $h^1(\Sigma, f^*T_{\mathbb{P}^1}) = 0$ of Lemma 24.4.3, we have:

LEMMA 27.5.2. *The fiber of the obstruction bundle on $\overline{\mathcal{M}}_0(\mathbb{P}^1, d)$ is canonically $H^1(\Sigma, f^*N_{\mathbb{P}^1/X}) = H^1(\Sigma, f^*(\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)))$.*

Note that (by the Riemann-Roch Theorem 22.1.1 or Serre duality) this vector bundle is rank $2d - 2 = \dim \overline{\mathcal{M}}_0(\mathbb{P}^1, d)$ as expected.

The Aspinwall–Morrison formula now follows from:

LEMMA 27.5.3.

$$\int_{\overline{\mathcal{M}}_0(\mathbb{P}^1,d)} \mathbf{e}(H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1)) \oplus H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1))) = 1/d^3.$$

PROOF. The idea is to choose different linearizations on the two copies of $\mathcal{O}_{\mathbb{P}^1}(-1)$ (and hence on the two copies of $f^*\mathcal{O}_{\mathbb{P}^1}(-1)$). On the first copy, choose $-H + \alpha_0$, and on the second choose $-H + \alpha_1$ (where as usual H is the hyperplane class on $\mathbb{P}^1$). By the localization formula (4.4), the integral in Lemma 27.5.3 has contributions from various decorated graphs Γ. We will see that only one graph contributes.

Suppose Γ is not the graph of Fig. 2. Then Γ must have more than one edge. We will see that the contribution of Γ is 0. This will follow by showing that $H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1)) \oplus H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1))$ has a (nonzero) fixed part (and hence trivial equivariant Euler class).

Let $f : \Sigma \to \mathbb{P}^1$ be any stable map in the fixed locus corresponding to Γ. Let $\tilde{\Sigma}$ be the normalization of Σ, and let the node-branches be b_i. Tensor the normalization exact sequence (22.1) by $f^*\mathcal{O}_{\mathbb{P}^1}(-1)$:

$$0 \to f^*\mathcal{O}_{\mathbb{P}^1}(-1) \to \bigoplus (f|_{\Sigma_i})^* \mathcal{O}_{\mathbb{P}^1}(-1) \to \bigoplus (f|_{b_i})^* \mathcal{O}_{\mathbb{P}^1}(-1) \to 0.$$

The long exact sequence in cohomology begins

$$0 \to \bigoplus H^0(b_i, (f|_{b_i})^* \mathcal{O}_{\mathbb{P}^1}(-1)) \to H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1)).$$

Suppose $f(b_1) = p_0$. If $\mathcal{O}_{\mathbb{P}^1}(-1)$ is linearized to be $\mathcal{O}(-H + \alpha_0)$ on $\mathbb{P}^1_{\mathbb{T}}$, then $\mathbb{T}$ acts trivially on $H^0(b_1, (f|_{b_1})^* \mathcal{O}_{\mathbb{P}^1}(-1))$, which is a sub-bundle of $H^1(\Sigma, f^*\mathcal{O}_{\mathbb{P}^1}(-1))$. The same is true if $f(b_1) = p_1$ and $\mathcal{O}_{\mathbb{P}^1}(-1)$ is linearized to be $\mathcal{O}(-H + \alpha_1)$.

Thus, as we have chosen the linearization $H - \alpha_0$ for one of the factors and $H - \alpha_1$ for the other, if Σ has any nodes at all, $H^1(\Sigma, f^*(\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)))$ has a non-trivial fixed sub-bundle. The only fixed locus with no node corresponds to the one-edge graph of Fig. 2. □

EXERCISE 27.5.1. *Using Exercises 27.2.6 and 27.4.1, finish the proof of Lemma 27.5.3 and hence the Aspinwall-Morrison formula.*

EXERCISE 27.5.2. *Suppose a Calabi–Yau threefold X contains a surface S isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$ or $\mathbb{P}^2$ (with inclusion $i : S \to X$). Given a homology class β on S, explain how to calculate the contribution of S to the Gromov–Witten invariant $\overline{\mathcal{M}}_0(X, i_*\beta)$ of $\overline{\mathcal{M}}_0(S,\beta)$. (*Hint: *Calculate $N_{S/X}$, then calculate the obstruction bundle on $\overline{\mathcal{M}}_0(S,\beta)$, and finally use localization.)*

27.6. Virtual Localization

Although the moduli space $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$ is not non-singular or even equidimensional, the technique of localization can be extended to deal with intersections against the virtual fundamental class. What follows is only a sketch.

The fixed loci of $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$ again correspond to graphs Γ, as described in Sec. 27.3. Note that, unlike the genus 0 case, the vertices of Γ have an additional labelling (of the arithmetic genus of the corresponding component). Each fixed locus has a "virtual normal bundle", which is found by a careful analysis of the deformation long exact sequence (24.2).

THEOREM 27.6.1 (Virtual localization formula for $\mathbb{P}^m$). *If ϕ is any equivariant cohomology class on $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)$,*

$$\phi \cap [\overline{\mathcal{M}}_{g,n}(\mathbb{P}^m, d)]^{\mathrm{vir}} = \sum_{\Gamma} \frac{\phi \cap [\overline{\mathcal{M}}_\Gamma]/\mathbb{A}_\Gamma}{\mathrm{e}(N_\Gamma^{\mathrm{vir}})}. \tag{27.9}$$

Virtual localization also applies to more general targets, in which case the fixed loci may also have virtual fundamental classes.

27.6.1. The Virtual Normal Bundle. In order to apply the virtual localization formula, we need to compute $\mathrm{e}(N_\Gamma^{\mathrm{vir}})$. The formula is given in Eq. (27.11); we sketch the derivation here.

As in the genus 0 case, we start with the deformation long exact sequence (which is two terms longer than in Eq. (27.6), as we no longer have the genus

0 vanishing of lemma 24.4.3):

(27.10)

$$\begin{array}{ccccccccc} 0 & \to & \mathrm{Aut}(\Sigma, p_1, \ldots, p_n) & \to & \mathrm{Def}(f) & \to & \mathrm{Def}(\Sigma, p_1, \ldots, p_n, f) & \to \\ & \to & \mathrm{Def}(\Sigma, p_1, \ldots, p_n) & \to & \mathrm{Ob}(f) & \to & \mathrm{Ob}(\Sigma, p_1, \ldots, p_n, f) & \to & 0. \end{array}$$

It turns out that the Euler class of the virtual normal bundle is given by

$$\mathrm{e}(N_\Gamma^{\mathrm{vir}}) = \frac{\mathrm{e}((\mathrm{Def}\, f)^{\mathrm{mov}})\mathrm{e}((\mathrm{Def}(\Sigma, p_1, \ldots, p_n))^{\mathrm{mov}})}{\mathrm{e}((\mathrm{Aut}\, f)^{\mathrm{mov}})\mathrm{e}((\mathrm{Aut}(\Sigma, p_1, \ldots, p_n))^{\mathrm{mov}})}.$$

This is a natural generalization of Eq. (27.7).

As before, one computes the fixed and moving parts of the four bundles above. The genus 0 arguments for $\mathrm{Aut}(\Sigma, p_1, \ldots, p_n)$ and $\mathrm{Def}(\Sigma, p_1, \ldots, p_n)$ carry through without change.

The analysis of $\mathrm{Def}(f)$ and $\mathrm{Ob}(f)$ (whose fibers are canonically $H^0(\Sigma, f^*T\mathbb{P}^m)$ and $H^1(\Sigma, f^*T\mathbb{P}^m)$ respectively) is slightly more subtle. As before, consider the normalization exact sequence resolving all of the nodes of Σ which are forced by the graph Γ:

$$0 \to \mathcal{O}_\Sigma \to \bigoplus_{\text{vertices}} \mathcal{O}_{\Sigma_v} \oplus \bigoplus_{\text{edges}} \mathcal{O}_{\Sigma_e} \to \bigoplus_{\text{flags}} \mathcal{O}_{x_F} \to 0.$$

Tensoring with $f^*(T\mathbb{P}^m)$ and taking cohomology yields:

$$0 \to H^0(\Sigma, f^*T\mathbb{P}^m) \to \bigoplus_{\text{vertices}} H^0(\Sigma_v, f^*T\mathbb{P}^m) \oplus \bigoplus_{\text{edges}} H^0(\Sigma_e, f^*T\mathbb{P}^m) \to$$

$$\to \bigoplus_{\text{flags}} T_{p_{\mu(F)}}\mathbb{P}^m \to H^1(\Sigma, f^*T\mathbb{P}^m) \to \bigoplus_{\text{vertices}} H^1(\Sigma_v, f^*T\mathbb{P}^m) \to 0.$$

Notice that our sequence is two terms longer than in the genus 0 case. Here we have used the fact that there will be no higher cohomology on the non-contracted components, as they are rational. As before, the following is true as representations, or in K-theory:

$$\begin{aligned} H^0(\Sigma, f^*T\mathbb{P}^m) - H^1(\Sigma, f^*T\mathbb{P}^m) = \\ \bigoplus_{\text{vertices}} T_{p_{\mu(v)}}\mathbb{P}^m + \bigoplus_{\text{edges}} H^0(\Sigma_e, f^*T\mathbb{P}^m) \\ - \bigoplus_{\text{flags}} T_{p_{\mu(F)}}\mathbb{P}^m - \bigoplus_{\text{vertices}} H^1(\Sigma_v, f^*T\mathbb{P}^m). \end{aligned}$$

As non-contracted components are rigid, we see that $H^0(\Sigma_e, f^*T\mathbb{P}^m)$ is trivial as a bundle. We determine its weights (exactly as in the genus 0 case) to obtain the contribution of $H^0(\Sigma_e, f^*T\mathbb{P}^m)$ to the Euler class ratio $\mathrm{e}(\mathrm{Def}(f)^{\mathrm{mov}})/\mathrm{e}(\mathrm{Ob}(f)^{\mathrm{mov}})$:

$$(-1)^{d(e)}\frac{d(e)!^2}{d(e)^{2d(e)}}(\alpha_i-\alpha_j)^{2d(e)}\prod_{\substack{a+b=d(e)\\ k\neq i,j}}\left(\frac{a}{d(e)}\alpha_i+\frac{b}{d(e)}\alpha_j-\alpha_k\right).$$

Finally, we evaluate the contribution of

$$H^1(\Sigma_v, f^*T\mathbb{P}^m) = H^1(\Sigma_v, \mathcal{O}_{\Sigma_v})\otimes T_{p_{\mu(v)}}\mathbb{P}^m.$$

As a bundle, $H^1(\Sigma_v, \mathcal{O}_{\Sigma_v})$ is the dual of the Hodge bundle $\mathbb{E} = H^0(\Sigma_v, \omega_{\Sigma_v})$ on the moduli space $\overline{\mathcal{M}}_{g(v),\mathrm{val}(v)}$ (Sec. 25.3). The bundle $H^1(\Sigma_v, \mathcal{O}_{\Sigma_v})\otimes T_{p_{\mu(v)}}\mathbb{P}^m$ splits into m copies of $\mathbb{E}^\vee$, twisted respectively by the r weights $\alpha_{\mu(v)}-\alpha_\nu$ for $\nu\neq\mu(v)$. Taking the equivariant Euler class of this bundle yields

$$\prod_{\nu\neq\mu(v)}\left(c(\mathbb{E}^\vee)\left(\frac{1}{\alpha_{\mu(v)}-\alpha_\nu}\right)\right)\cdot(\alpha_{\mu(v)}-\alpha_\nu)^{g(v)}$$

where for a bundle Q of rank q:

$$c(Q)(t) = 1+tc_1(Q)+\cdots+t^q c_q(Q).$$

Hence the inverse Euler class of the virtual normal bundle to the fixed point locus corresponding to the graph Γ is given by:

$$\begin{aligned}1/\mathrm{e}(N^{\mathrm{vir}}) = &\prod_{\mathrm{flags}}\frac{1}{\omega_F-\psi_F}\prod_{\nu\neq\mu(F)}(\alpha_{\mu(F)}-\alpha_\nu)\\ &\prod_{\mathrm{vertices}}\prod_{\nu\neq\mu(v)}\left(c(\mathbb{E}^\vee)\left(\frac{1}{\alpha_{\mu(v)}-\alpha_\nu}\right)\right)\cdot(\alpha_{\mu(v)}-\alpha_\nu)^{g(v)-1}\\ &\prod_{\mathrm{val}(v)=2}\frac{1}{\omega_{F_{v,1}}+\omega_{F_{v,1}}}\prod_{\mathrm{val}(v)=1}\omega_F\\ &\prod_{\mathrm{edges}}\frac{(-1)^{d(e)}d(e)^{2d(e)}}{(d(e)!)^2(\alpha_i-\alpha_j)^{2d(e)}}\prod_{\substack{a+b=d(e)\\ k\neq i,j}}\frac{1}{\frac{a}{d(e)}\alpha_i+\frac{b}{d(e)}\alpha_j-\alpha_k}.\end{aligned}$$

27.6.2. Application: Genus g Gromov–Witten Invariants of Projective Space. Virtual localization is a powerful tool. For example, it immediately gives a method for computing all Gromov–Witten invariants of projective space as follows.

EXERCISE 27.6.1. *Use virtual localization to show that*

$$\langle H^{l_1},\cdots,H^{l_n}\rangle_{g,d}^{\mathbb{P}^m} = \sum_{\Gamma}\frac{1}{|\mathbb{A}_\Gamma|}\int_{\overline{\mathcal{M}}_\Gamma}\frac{\prod_{j=1}^n\alpha_{\mu(j)}^{l_j}}{\mathrm{e}(N_\Gamma^{\mathrm{vir}})} \tag{27.11}$$

where H is the hyperplane class in $\mathbb{P}^m$.

To evaluate the integral, one expands the terms of the form $\frac{1}{\omega_F-\psi_F}$ as formal power series, and then integrates all terms of the appropriate degree. Each integral that is encountered splits as a product of integrals over the different moduli spaces of pointed curves $\overline{\mathcal{M}}_{g(v),\mathrm{val}(v)}$.

There are natural methods of reducing the integrals of Hodge classes arising in the localization formula Eq. (27.11) to the pure ψ-integrals determined by Witten's conjectures. At a computational level, Faber has implemented an algorithm that evaluates Hodge integrals using Mumford's Grothendieck–Riemann–Roch formulas and Witten's conjectures. Faber and Pandharipande have found a natural sequence of differential operators annihilating generating series of Hodge and ψ-classes, and have developed theoretical techniques for manipulating Hodge integrals. There is rich interaction between Gromov–Witten theory, localization, and Hodge integrals.

27.7. The Full Multiple Cover Formula for $\mathbb{P}^1$

Let $\mathbb{P}^1\subset X$ be a rigid embedding in a Calabi–Yau threefold. The multiple cover contributions of $\mathbb{P}^1$ may be expressed as integrals over the moduli spaces of stable maps. More precisely, the genus g, degree d contribution $C(g,d)$ is

$$C(g,d) = \int_{[\overline{\mathcal{M}}_g(\mathbb{P}^1,d)]^{\mathrm{vir}}} c_{\mathrm{top}}(R^1\pi_*\mu^*N), \tag{27.12}$$

where $N = \mathcal{O}_{\mathbb{P}^1}(-1)\oplus\mathcal{O}_{\mathbb{P}^1}(-1)$ denotes the normal bundle. (Here $R^1\pi_*$ is the first "higher direct image sheaf" which generalizes the vector bundle of Lemma 27.5.2.) The Aspinwall–Morrison formula computes the genus 0 case of Eq. (27.12).

THEOREM 27.7.1 (Faber–Pandharipande). *The multiple cover contributions are determined in higher genus by*

$$C(1,d) = \frac{1}{12d}$$

for $g = 1$, *and*

$$C(g,d) = \frac{|B_{2g}| \cdot d^{2g-3}}{2g \cdot (2g-2)!} = |\chi(\mathcal{M}_g)| \frac{d^{2g-3}}{(2g-3)!}, \tag{27.13}$$

for $g \geq 2$. B_{2g} *is the* $2g^{th}$ *Bernoulli number and* $\chi(\mathcal{M}_g) = B_{2g}/2g(2g-2)$ *is the orbifold Euler characteristic of* $\mathcal{M}_g$.

PROOF(SKETCH). We will evaluate (27.12) by the virtual localization formula. It is crucial to use the linearization choices introduced in Sec. 27.5.1. A study of the fixed component contributions yields a general genus vanishing principle: if a graph Γ contains *any* vertex of valence greater than 1, then the contribution of Γ to Eq. (27.12) vanishes.

Hence, the contributing graphs have exactly 1 edge. The graph sum reduces simply to a sum over partitions $g_1 + g_2 = g$ of the genus. The localization formula then yields the result for $g \geq 0$:

$$C(g,d) = d^{2g-3} \sum_{\substack{g_1+g_2=g \\ g_1, g_2 \geq 0}} b_{g_1} b_{g_2} \tag{27.14}$$

where b_g is the Hodge integral

$$b_g = \begin{cases} 1, & g = 0, \\ \int_{\overline{\mathcal{M}}_{g,1}} \psi_1^{2g-2} \lambda_g, & g > 0. \end{cases}$$

As expected, Hodge integrals play a central role in the virtual localization analysis.

The Aspinwall–Morrison formula is recovered from the genus 0 restriction of Eq. (27.14). The formula for $C(1,d)$ is immediately deduced from

$$b_1 = \int_{\overline{\mathcal{M}}_{1,1}} \lambda_1 = \frac{1}{24}.$$

The Hodge integrals b_g have been computed in generating series form:

$$\sum_{g \geq 0} b_g t^{2g} = \left(\frac{t/2}{\sin(t/2)} \right).$$

Eq. (27.14) implies

$$C(g,d) = d^{2g-3} C(g,1),$$

$$\sum_{g \geq 0} C(g,1) t^{2g} = \left(\frac{t/2}{\sin(t/2)} \right)^2.$$

Formula (27.13) is then obtained from the definition of the Bernoulli numbers. □

CHAPTER 28

The Fundamental Solution of the Quantum Differential Equation

We study here a solution of a differential equation that arises naturally in mirror symmetry and quantum cohomology. One is forced here to consider descendant invariants even if one is primarily interested in Gromov–Witten invariants (as in the statement of the Mirror conjecture). The correlator S_b defined via the fundamental solution will play an essential role in the proof of the Mirror conjecture.

As in Sec. 26.5, let $T_0, \ldots, T_m$ be a homogeneous basis of $V = H^*(X)$ such that T_0 is the ring identity and T_m is its Poincaré dual. The tangent space of V at every point is canonically identified with V. Let $\partial_0, \ldots, \partial_m$ be the corresponding tangent fields. Let $\gamma = \sum t_i T_i$ be coordinates on V defined by the basis. Let $F = \sum f^i \partial_i$ be a vector field. Let C be the Gromov–Witten or quantum potential (Definition 26.5.1):

$$C(t_0, \ldots, t_m) = \sum_{n \geq 3} \sum_{\beta} \frac{1}{n!} \langle \gamma^n \rangle_\beta.$$

The quantum product is defined by

$$\partial_i * \partial_j = \sum_{r,s} C_{ijr} g^{rs} \partial_s.$$

This is the same quantum product defined earlier, in slightly different notation. Define a (formal) connection $\nabla_\hbar$ on the tangent bundle of V by

$$\nabla_{\hbar,i}(F) = \hbar \frac{\partial F}{\partial t_i} - \partial_i * F = \sum_s \left(\hbar \frac{\partial f^s}{\partial t_i} - \sum_r C_{ijr} g^{rs} f^j \right) \partial_s.$$

WARNING 28.0.1. In Givental's notation $\hbar$ is the generator of the equivariant cohomology of $\mathbb{C}^*$ — it does not correspond to a physical constant.

EXERCISE 28.0.1. *Use the WDVV equations (Sec. 26.5) to show that* $\nabla_\hbar$ *is flat. Therefore, flat vector fields exist (at least formally).*

Rewrite the equations for flat solutions F as

$$\hbar\frac{\partial F}{\partial t_i} = \partial_i * F \quad \text{for } 1 \le i \le m. \tag{28.1}$$

This is called the *quantum differential equation.*

Define a matrix of formal functions in $\mathbb{Q}[[\hbar^{-1}, t_i]]$:

$$\Psi_{ab} = g_{ab} + \sum_{\substack{n\ge 0,\ \beta\\ (n,\beta)\neq(0,0)}} \frac{1}{n!}\langle T_a \cdot \frac{T_b}{\hbar - \psi}\cdot \gamma^n\rangle_\beta$$

where $0 \le a, b \le m$. (Intuitively, g_{ab} is the $(n,\beta) = (0,0)$ term.) The matrix may be written more explicitly as

$$\Psi_{ab} = g_{ab} + \sum_{\substack{n\ge 0,\ \beta\\ (n,\beta)\neq(0,0)}} \sum_{k\ge 0} \frac{\hbar^{-k-1}}{n!}\langle T_a \cdot \tau_k(T_b)\cdot \gamma^n\rangle_\beta. \tag{28.2}$$

PROPOSITION 28.0.2. *Ψ yields a fundamental solution of the quantum differential equation*

$$\nabla_{\hbar,i}\sum_{a,s}\Psi_{ab}g^{as}\partial_s = 0$$

for all b.

Note that the constant term of the solution $\sum_a \Psi_{ab}g^{as}$ is the identity matrix.

PROOF. A direct calculation (see Exercise 26.5.1) gives the left side of Eq. (28.1):

$$\sum_{a,s}\hbar\frac{\partial\Psi_{ab}}{\partial t_i}g^{as}\partial_s = \sum_{a,s}\sum_{\substack{n\ge 0\\ \beta}}\sum_{k\ge 0}\frac{\hbar^{-k}}{n!}\langle T_a\cdot\tau_k(T_b)\cdot T_i\cdot\gamma^n\rangle_\beta g^{as}\partial_s. \tag{28.3}$$

The coefficient of ∂_s of the right side of Eq. (28.1), $\sum_{a,j}\partial_i * \Psi_{ab}g^{aj}\partial_j$, is

$$\sum_{a,j,r} C_{ijr}g^{rs}\Psi_{ab}g^{aj} = \sum_r C_{ibr}g^{rs} +$$

$$\sum\sum_{a,j,r}\sum_{k\ge 0}\frac{\hbar^{-k-1}}{n_1!n_2!}\langle T_a\cdot\tau_k(T_b)\cdot\gamma^{n_1}\rangle_{\beta_1}g^{aj}\langle T_i\cdot T_j\cdot T_r\cdot\gamma^{n_2}\rangle_{\beta_2}g^{rs} \tag{28.4}$$

where the first sum in the last term is over stable splittings $n_1 + n_2 = n$, $\beta_1 + \beta_2 = \beta$.

The $k = 0$ terms of Eq. (28.3) sum to the first term $\sum C_{ibr}g^{rs}$ of Eq. (28.4). The $k \ge 1$ terms of Eq. (28.3) may be replaced via the topological recursion relation Eq. (26.11) relative to the first 3 markings to obtain the second term in Eq. (28.4). □

EXERCISE 28.0.2. *Verify Eq. (28.4), using Eq. (28.2) and the explicit formula for C_{ijr} of Exercise 26.5.1.*

28.1. The "Small" Quantum Differential Equation

The fundamental solution takes a simpler form after passing to small quantum cohomology (see Sec. 26.5.1). Let T denote the vector of cohomology classes $(T_1,\dots,T_p)$, and let t denote the vector of variables $(t_1,\dots,t_p)$. For $\beta\in H_2(X)$, let v_β denote the vector of constants $(\int_\beta T_1,\dots,\int_\beta T_p)$. As in Sec. 26.5.1, for simplicity we assume that T_1, …, T_p are dual to a basis of effective curve classes.

By Eq. (26.13) and Eq. (26.14), the small product may be written as

$$\partial_i * \partial_j = \sum_{\beta,r,s} e^{v_\beta\cdot t}\langle T_i\cdot T_j\cdot T_r\rangle_\beta g^{rs}\partial_s.$$

The matrix Ψ can be written after restriction to the Kähler classes as

$$\Psi_{ab} = \sum_\beta e^{v_\beta\cdot t}\langle T_a\cdot\frac{e^{T\cdot t/\hbar}T_b}{\hbar-\psi}\rangle_\beta. \tag{28.5}$$

The divisor equation Eq. (26.10) is again used, as in Exercise 26.5.5. In the case $\beta = 0$, we follow the convention

$$\langle T_a\cdot\frac{e^{T\cdot t/\hbar}T_b}{\hbar-\psi}\rangle_0 := \langle T_a\cdot e^{T\cdot t/\hbar}T_b\cdot T_0\rangle_0 \tag{28.6}$$

(as 2-point degree 0 invariants are not defined).

Note that the series Eq. (28.5) is a sum of two-point invariants (with the exception of the convention just described), and is an element of $\mathbb{Q}[\hbar^{-1}, t][[e^t]]$.

EXERCISE 28.1.1. *Show that, modulo the variables t_i and e^{t_i}, $\sum_a\Psi_{ab}g^{as}$ is the identity matrix.*

The *small quantum differential equation* is

$$1\le i\le p,\quad \hbar\frac{\partial F}{\partial t_i} = \partial_i * F$$

where F is a vector field function of only t_1, …, t_p, and the product is the small quantum product (cf. Eq. (28.1)). For $1\le i\le p$, the small analogue of Proposition 28.0.2 holds for (28.5):

$$\sum_a \hbar\frac{\partial\Psi_{ab}}{\partial t_i}g^{as}\partial_s = \sum_a \partial_i * \Psi_{ab}g^{as}\partial_s,$$

and the restricted matrix $\sum_a \Psi_{ab} g^{as}$ is a fundamental solution to this small quantum differential equation.

28.2. Example: Projective Space Revisited

We now revisit the results of Sec. 26.5.2. Let $X = \mathbb{P}^m$. Let H denote the hyperplane class in $H^2(\mathbb{P}^m)$, and let $T_i = H^i$ be the natural basis for cohomology. Let $t = t_1$. From Eq. (26.15), the small quantum ring structure is

$$QH_s^*(\mathbb{P}^m) = \mathbb{Q}[\partial_1, e^t]/(\partial_1^{m+1} - e^t).$$

Let $\sum_{i=0}^m f^i \partial_i$ be a vector field, where $f^i = f^i(t)$. (The superscripts do *not* denote powers!) The small quantum differential equation is then the system

$$i > 0, \quad \hbar \frac{\partial f^i}{\partial t} = f^{i-1},$$

$$\hbar \frac{\partial f^0}{\partial t} = e^t f^m.$$

The function f^m determines a vector solution if and only if it is annihilated by the operator

$$\mathcal{D} = \left(\hbar \frac{d}{dt}\right)^{m+1} - e^t. \tag{28.7}$$

A (formal) fundamental solution to the equation $\mathcal{D}f = 0$ is given by the expression

$$S \equiv \sum_{d \geq 0} \frac{e^{(H/\hbar + d)t}}{\prod_{j=1}^d (H + j\hbar)^{m+1}} \quad (\text{mod } H^{m+1}). \tag{28.8}$$

Expand S in powers of H as

$$S = \sum_{b=0}^m S_b H^{m-b}$$

where S_b is a formal series in $\mathbb{Q}[\hbar^{-1}, t][[e^t]]$.

EXERCISE 28.2.1. *Use Eq. (28.8) to check that $\mathcal{D}$ annihilates S_b.*

Define the matrix M of functions by

$$M_b^s = \left(\hbar \frac{d}{dt}\right)^{m-s} S_b.$$

Modulo t and e^t, the only contribution to M_b^s occurs in the $d = 0$ summand in Eq. (28.8); it is the identity matrix. $M_b^s \partial_s$ defines a fundamental solution

to the small quantum differential equation. By uniqueness of fundamental solutions,

$$\Psi_{ab} g^{as} = M_b^s. \tag{28.9}$$

(This uniqueness statement depends on the equality modulo t, e^t and the fact that the solutions lie in $\mathbb{Q}[\hbar^{-1}, t][[e^t]]$.)

Eqs. (28.8) and (28.9) together with the solution Eq. (28.5) compute all two-point invariants of $\mathbb{P}^m$ with a cotangent line class on one point. For example, tracing through the equations yields

$$\langle \tau_{dm+d-2}(T_m) \rangle_d = \int_{\overline{\mathcal{M}}_{0,1}(\mathbb{P}^m, d)} \mathrm{ev}_1^*(H^m) \cup \psi^{dm+d-2} = \frac{1}{(d!)^{m+1}}.$$

The solution to the small quantum differential equation provides an elegant organization of these two-point descendant invariants.

CHAPTER 29

The Mirror Conjecture for Hypersurfaces I: The Fano Case

In this chapter, we will describe the relationship between hypergeometric series and the quantum cohomology of hypersurfaces in projective space. While the most general context for such relationships has not yet been understood, tremendous progress has recently been made by numerous researchers. We will follow the approach of Givental.

29.1. Overview of the Conjecture

In the early 1990s, Candelas, de la Ossa, Green, and Parkes made a startling prediction of the number of rational curves on a quintic threefold $X \subset \mathbb{P}^4$. We describe here an equivalent form of their prediction.

The expected dimension $\operatorname{vdim} \overline{\mathcal{M}}_0(X,d)$ of the moduli space of genus 0 maps to X is 0. Let N_d be the genus 0 degree d Gromov–Witten invariant of X, given by the degree of the virtual fundamental class of the space of maps. The Aspinwall–Morrison formula (Proposition 27.5.1) already shows that N_d is not enumerative. One might naively hope that all rational curves in X are immersed with "generic" normal bundle $\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$. This is not the case (see Sec. 29.1.2), but even if it *were*, each degree d such curve would contribute $1/k^3$ to the degree dk invariants ($k \geq 1$). We correct for this by defining new numbers n_d (called *instanton numbers*) via

$$\sum_{d=1}^{\infty} N_d e^{dt} = \sum_{d=1}^{\infty} \sum_{k=1}^{\infty} n_d k^{-3} e^{kdt}.$$

We describe the instanton numbers n_d as *virtually enumerative* (see Sec. 29.1.2 for further discussion).

Define $I_i(t)$ by

$$\sum_{i=0}^{3} I_i H^i \equiv \sum_{d=0}^{\infty} e^{(H+d)t} \frac{\prod_{r=1}^{5d}(5H+r)}{\prod_{r=1}^{d}(H+r)^5} \pmod{H^4}. \tag{29.1}$$

H will later be interpreted as the hyperplane class in $\mathbb{P}^4$. These functions are a basis of the Picard–Fuchs differential equation (see e.g. Sec. 6.5.3)

$$\left(\frac{d}{dt}\right)^4 I = 5e^t\left(5\frac{d}{dt}+1\right)\left(5\frac{d}{dt}+2\right)\left(5\frac{d}{dt}+3\right)\left(5\frac{d}{dt}+4\right)I.$$

This equation arises in the B-model from the variation of Hodge structures of a specific family of Calabi–Yau threefolds.

Define a new coordinate T by $T(t) = I_1(t)/I_0(t)$. The functions $J_i = I_i(T)/I_0(T)$ in the new variable were predicted to equal an A-model series:

$$\sum_{i=0}^{3} J_i H^i \equiv e^{HT} + \frac{H^2}{5}\sum_{d=1}^{\infty} n_d d^3 \sum_{k=1}^{\infty} \frac{e^{(H+kd)T}}{(H+kd)^2} \pmod{H^4} \tag{29.2}$$

and satisfy the differential equation

$$\frac{d^2}{dT^2}\left(\frac{1}{5+\sum_{d=1}^{\infty} n_d d^3 \frac{e^{dT}}{1-e^{dT}}}\right)\frac{d^2}{dT^2}J_i = 0 \tag{29.3}$$

These formulas were completely unexpected, as the enumerative geometry of quintic threefolds was not known to have any structure at all.

Givental's approach is as follows. Let X be a hypersurface in $\mathbb{P}^m$ of degree $l \le m+1$. Motivated by the quantum product and related quantum differential equation (see Sec. 28), we define the "correlator" $S_X(t)$ encapsulating Gromov–Witten and descendant invariants (see Sec. 29.2). It is closely related to the hypergeometric series

$$S_X^*(t) \equiv \sum_{d=0}^{\infty} e^{(H+d)t}\frac{\prod_{r=1}^{ld}(lH+r)}{\prod_{r=1}^{d}(H+r)^{m+1}} \pmod{H^m}$$

which arises as a solution to the Picard–Fuchs differential equation

$$\left(\frac{d}{dt}\right)^m I \equiv le^t\left(l\frac{d}{dt}+1\right)\cdots\left(l\frac{d}{dt}+(l-1)\right)I \pmod{H^m}.$$

The precise relationship is divided into three cases.

(i) *Fano index >1 case.* If $l<m$, then $S_X(t) = S_X^*(t)$.
(ii) *Fano index 1 case.* If $l=m$, then $S_X(t) = e^{-m!e^t}S_X^*(t)$.
(iii) *Calabi–Yau case.* If $l=m+1$, then S_X and S_X^* are related by an explicit Mirror transformation (see Ch. 30).

Recall that X is a Fano variety if the anti-canonical bundle $\mathcal{K}_X^*$ is ample. The index of a Fano variety X is the largest integer l such that $\mathcal{K}_X$ has an lth root.

EXERCISE 29.1.1. *Interpret the case of projective space $\mathbb{P}^r$ (Secs. 26.5.2 and 28.2) as the special case $l=1$ of (i) above (in the projective space of dimension $m=r+1$); interpret differential equation (28.7), i.e.*

$$\left(\hbar\frac{\partial}{\partial t}\right)^{r+1} f^r = e^t f^r,$$

as the Picard–Fuchs equation.

In the case of the quintic threefold, $S_X(t)$ is exactly the right side of Eq. (29.2). The transformation (iii) then specializes to the mirror symmetry prediction, proving Eq. (29.2). The differential equation Eq. (29.3) is a consequence of the quantum differential equation. The results in cases (i) and (ii) also have direct applications to the quantum cohomology ring of the corresponding Fano hypersurfaces.

29.1.1. Overview of Proof. The correlator S_X will be obtained from the fundamental solution of the quantum differential equation for X. We first analyze S_X by torus localization on $\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m, d)$. The localization formula immediately yields an algorithm to compute any given invariant. However, a more subtle result is required here — we will use the graph structure of the localization formulas to find linear recursions relating the correlators.

In the Fano index >1 case (i), $S_X(t)$ and $S_X^*(t)$ satisfy identical recursions, proving this case. In the Fano index 1 case (ii), the recursions are almost identical — they differ by one term. The index 1 case is then proved by studying the effect of this term. In the Calabi–Yau case (iii), the recursions are related but quite different; this case is the most difficult, and will be dealt with in Ch. 30.

29.1.2. The Clemens Conjecture. One might naively hope that for a generic quintic threefold X the following conditions holds

(i) For each curve class β, there are only finitely many irreducible rational curves $C \subset V$ of class β, each of which is non-singular with normal bundle isomorphic to $\mathcal{O}_{\mathbb{P}^1}(-1)\oplus\mathcal{O}_{\mathbb{P}^1}(-1)$.
(ii) As β varies, these curves are disjoint from each other.

This strong assumption is false, however — there exist nodal degree 5 rational curves on a generic quintic X.

If conditions (i) and (ii) are true for a Calabi–Yau threefold X, then the invariant n_β (obtained by correcting the Gromov–Witten invariants, taking into account multiple covers using the Aspinwall–Morrison formula) is a true count of curves in class β. For this reason, the invariants n_β should be viewed as carrying integral content. The study of the invariants n_β will be taken up again in Ch. 34 — many open questions remain in this area.

A Calabi–Yau threefold is said to be *ideal* if conditions (i) and (ii) hold. As there are essentially no known examples of ideal Calabi–Yau threefolds, the definition is useful at present only for theoretical discussion.

For the quintic, the numbers n_d are enumerative, at least for $d \leq 9$.

A weaker condition on rational curves in the quintic may still be true:

CONJECTURE 29.1.1 (Clemens). *Let $V \subset \mathbb{P}^4$ be a generic quintic threefold. Then for each degree $d \geq 1$:*

(i) *There are only finitely many irreducible rational curves $C \subset V$ of degree d.*
(ii) *These curves, as we vary over all degrees, are disjoint from each other.*
(iii) *If $\mathbb{P}^1 \to C$ is the normalization of an irreducible rational curve $C \subset V$, then the normal bundle N_f to $f : \mathbb{P}^1 \to V$ is isomorphic to $\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$.*

Note that the Clemens conjecture and the invariants n_d together are not sufficient to count curves on the general quintic.

29.2. The Correlators $S(t,\hbar)$ and $S_X(t,\hbar)$

Generalizing the example of projective space in Sec. 28.2, define the *correlator* $S_b = \Psi_{0b}$ for a general space X, where Ψ is the matrix corresponding to the *small* quantum differential equation (see Sec. 28.1 — just set

$$t_0 = t_{p+1} = \cdots = t_m = 0$$

in the matrix corresponding to the big differential equation). For convenience, as in Sec. 28.2, let $T_i = H^i$ be the natural basis for the cohomology of $\mathbb{P}^m$, and let $t = t_1$.

Define the correlator $S(t,\hbar)$ by

$$S(t,\hbar) = \sum_{b=0}^{m} S_b H^{m-b} = \sum_{b=0}^{m}\sum_{d} e^{dt}\langle 1, \frac{e^{Ht/\hbar}}{\hbar-\psi}H^b\rangle_d^X H^{m-b}. \tag{29.4}$$

The second equality uses Eq. (28.5). Note that $S(t,\hbar)$ depends on l, but for simplicity of notation, we suppress l in the notation.

Recall that the $d=0$ case is special; the bracket should be interpreted as $\langle 1, \frac{e^{Ht/\hbar}}{\hbar-\psi}H^b, 1\rangle_d^X$ (as in Eq. (28.6)).

For notational convenience throughout this section, define $E_d := \pi_* f^* \mathcal{O}_{\mathbb{P}^m}(l)$. From Sec. 26.1.3, E_d is the bundle used to define the virtual fundamental class of $\overline{\mathcal{M}}_{0,n}(X,d)$ (or more precisely, its push-forward to $\overline{\mathcal{M}}_{0,n}(\mathbb{P}^m,d)$). Notice that there is a natural (surjective) map $E_d \to \mathrm{ev}_2^*(\mathcal{O}_{\mathbb{P}^m}(l))$. (Recall that the fiber of E_d above a point of $\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)$ is canonically a section of $f^*(\mathcal{O}_{\mathbb{P}^m}(l))$ on Σ. The value of this section at the second marked point is naturally an element of the fiber of $\mathrm{ev}_2^*(\mathcal{O}_{\mathbb{P}^m}(l))$.) Let E_d' be the kernel of this morphism, so the sequence

$$0 \to E_d' \to E_d \to \mathrm{ev}_2^*\mathcal{O}_{\mathbb{P}^m}(l) \to 0$$

is exact. The rank of E_d' is $5d$, and

$$\mathrm{e}(E_d) = lH\mathrm{e}(E_d'). \tag{29.5}$$

PROPOSITION 29.2.1.

$$S(t,\hbar) = \sum_{d\geq 0} e^{dt}e^{Ht/\hbar}\mathrm{ev}_{2*}\left(\frac{\mathrm{e}(E_d)}{\hbar-\psi}\right). \tag{29.6}$$

WARNING 29.2.1. Here ev_2 is the evaluation map $\mathrm{ev}_2 : \overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d) \to \mathbb{P}^m$, not $\mathrm{ev}_2 : \overline{\mathcal{M}}_{0,2}(X,d) \to X$.

When $d=0$, $\mathrm{ev}_{2*}\left(\frac{\mathrm{e}(E_d)}{\hbar-\psi}\right)$ should be interpreted as dH.

PROOF. We consider the $d>0$ case here; the special case of checking that the $d=0$ conventions on both sides agree is left as an exercise.

To show two elements of $H^*(\mathbb{P}^m)$ are the same, it suffices to show that the "H^{m-b} component" of both elements are the same, for all $0 \leq b \leq m$. The H^{m-b} component of the left side of Eq. (29.6) is

$$\begin{aligned} e^{dt}\langle 1, \frac{e^{Ht/\hbar}}{\hbar-\psi}H^b\rangle &= e^{dt}\int_{[\overline{\mathcal{M}}_{0,2}(X,d)]^{\mathrm{vir}}} \frac{e^{Ht/\hbar}H^b}{\hbar-\psi} \\ &= e^{dt}\int_{\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)} \frac{\mathrm{e}(E_d)}{\hbar-\psi}e^{Ht/\hbar}H^b \end{aligned}$$

(by Sec. 26.1.3)

$$= e^{dt}\int_{\mathbb{P}^m} \mathrm{ev}_{2*}\left(\frac{\mathrm{e}(E_d)}{\hbar-\psi}\right)e^{Ht/\hbar}H^b$$

(by the projection formula), which is the H^{m-b} part of the right side of Eq. (29.6). □

Note that by Eq. (29.5), $S(t,\hbar)$ is "divisible by lH". Define $S_X(t,\hbar)$ by

$$S_X(t,\hbar) := \frac{1}{lH} S(t,\hbar),$$

or equivalently,

$$S_X(t,\hbar) := \sum_{d\geq 0} e^{dt} e^{Ht/\hbar} \mathrm{ev}_{2*}\left(\frac{\mathbf{e}(E'_d)}{\hbar-\psi}\right).$$

29.2.1. The Correlator $S_X(t,\hbar)$ Encodes Gromov–Witten Invariants. The correlator $S_X(t,\hbar)$ encodes Gromov–Witten invariants (generalizing the remarks at the end of Sec. 28.2). We show this in the special case of the quintic threefold; the general case is similar. In this case, by Proposition 29.2.1,

$$S_X(t,\hbar=1) = \frac{1}{5H}\sum_{d>0} e^{(H+d)t} \mathrm{ev}_{2*}\left(\frac{\mathbf{e}(E_d)}{1-\psi_2}\right).$$

Note that

$$\mathrm{ev}_{2*}\left(\frac{\mathbf{e}(E_d)}{1-\psi_2}\right) = \mathrm{ev}_{2*}\left(\mathbf{e}(E_d)+\mathbf{e}(E_d)\psi_2+\mathbf{e}(E_d)\psi_2^2\right).$$

EXERCISE 29.2.1. *Show that no other terms can appear.* Hint: *Use*

$$\mathrm{vdim}\,\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d) - \mathrm{rank}(E_d) = 2.$$

Also, $\mathrm{ev}_{2*}(\mathbf{e}(E_d)) = 0$, as we can push forward via the composition

$$\overline{\mathcal{M}}_{0,\{1,2\}}(\mathbb{P}^m,d) \to \overline{\mathcal{M}}_{0,\{2\}}(\mathbb{P}^m,d) \overset{\mathrm{ev}_2}{\to} \mathbb{P}^m.$$

Hence (using the same method as in the proof of Proposition 29.2.1):

$$\mathrm{ev}_{2*}\left(\frac{\mathbf{e}(E_d)}{1-\psi_2}\right) = H^3\langle 1\cdot\tau_1(H^1)\rangle_d^X + H^4\langle 1\cdot\tau_2(H^0)\rangle_d^X.$$

EXERCISE 29.2.2. *Use the string, dilaton, and divisor equations (Sec. 26.3) to show that this is*

$$dN_dH^3 - 2N_dH^4.$$

Define $F(t) = \frac{5}{6}t^3 + \sum_{d>0} N_d e^{dt}$. Mirror symmetry can then be interpreted as

$$F(t) = \frac{5}{2}\left(\frac{I_1(t)I_2(t)-I_3(t)I_0(t)}{I_0(t)^2}\right) = \frac{5}{2}\left(\frac{I_1(t)}{I_0(t)}\cdot\frac{I_2(t)}{I_0(t)} - \frac{I_3(t)}{I_0(t)}\right).$$

This is the standard form of the mirror prediction for quintic threefolds.

29.3. The Torus Action

Consider the torus $\mathbb{T}$ and action on $\mathbb{P}^m$ as described in Sec. 27.1. From here onwards, all geometric structures (sheaves, maps, pushforwards, cohomology groups, etc.) will be given their canonical equivariant interpretations. For example, the vector bundle E_d has a natural $\mathbb{T}$-action, as it was defined canonically.

The correlator

$$S(t,\hbar) = \sum_{d\geq 0} e^{(H/\hbar+d)t} \mathrm{ev}_{2*}\left(\frac{\mathbf{e}(E_d)}{\hbar-\psi_2}\right)$$

(see Eq. (29.6)) will be hereafter interpreted equivariantly, that is, as an equivariant cohomology class. The non-equivariant version of $S(t,\hbar)$ can be recovered by setting the α_i to 0.

Recall that $S(t,\hbar)$ is completely determined by its pairings with the ϕ_i (Exercise 27.1.5 (c)). Motivated by this, we compute

$$\begin{aligned}(\phi_i, S(t,\hbar)/lH) &= \frac{e^{\alpha_i t/\hbar}}{l\alpha_i}\sum_{d\geq 0} e^{dt}\int_{\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)} \frac{\mathbf{e}(E'_d) l\alpha_i}{\hbar-\psi_2}\mathrm{ev}_2^*(\phi_i)\\ &= e^{\alpha_i t/\hbar} Z_i(e^t,\hbar)\end{aligned}$$

(using Eq. (29.5), the projection formula, and Exercise 27.1.5 (a)), where

$$Z_i(e^t,\hbar) = 1 + \sum_{d>0} e^{dt}\int_{\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)} \frac{\mathbf{e}(E'_d)}{\hbar-\psi_2}\mathrm{ev}_2^*(\phi_i). \tag{29.7}$$

As the dimension of $\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)$ is $(m+1)d+m-1$, the rank of E'_d is ld, and the codimension of $\mathrm{ev}_2^*(\phi_i)$ is m, the initial terms in the $(1/\hbar)$ expansion of the above expression with ψ_2-degree less than $(m+1-l)d-1$ vanish for dimensional reasons. Hence if $l \leq m$ (cases (i) and (ii)),

(29.8)
$$Z_i(e^t,\hbar) = 1 + \sum_{d>0}\left(\frac{e^t}{\hbar^{m+1-l}}\right)^d \int_{\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)} \frac{\psi_2^{(m+1-l)d-1}}{1-\psi_2/\hbar}\mathbf{e}(E'_d)\mathrm{ev}_2^*(\phi_i).$$

In this case, define $z_i(Q,\hbar) = Z_i(Q\hbar^{m+1-l},\hbar)$.

29.4. Localization

We will use localization to analyze the correlators Z_i. Recall (Sec. 27.3) that the fixed loci of the $\mathbb{T}$-action on $\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)$ correspond to triples

(Γ, μ, d) where Γ is a two-pointed (or two-tailed) tree, and

$$\mu : \mathrm{Vert}(\Gamma) \to \{p_0, \ldots, p_m\} \quad \text{and} \quad d : \mathrm{Edge}(\Gamma) \to \mathbb{Z}^{>0}.$$

As usual, we will let Γ denote the entire decorated graph structure. Let G_d be the set of all possible Γ.

By the localization formula Eq. (4.4), if W is any equivariant cohomology class on $\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m, d)$,

$$\int_{\overline{\mathcal{M}}_{0,2}(\mathbb{P}^m,d)} W \mathrm{ev}_2^*(\phi_i) = \sum_{\Gamma \in G_d} \int_{\overline{\mathcal{M}}_\Gamma} \frac{1}{\mathrm{e}(N_\Gamma)} W \mathrm{ev}_2^*(\phi_i). \tag{29.9}$$

We will use this formula several times. For example, taking

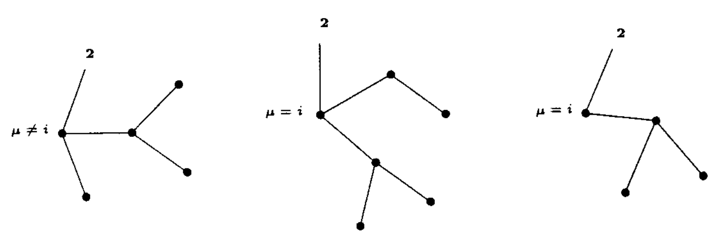
$$W = \frac{\psi_2^{(m+1-l)d-1}}{1 - \psi_2/\hbar} \mathrm{e}(E'_d),$$

we can compute Z_i in the case $l \leq m$ using Eq. (29.8).

Fix an index $0 \leq i \leq m$. We partition G_d into three disjoint subsets G_d^{i*}, G_d^{i0}, and G_d^{i1}.

- The set G_d^{i*} consists of the fixed loci for which the second marked point is *not* mapped to $p_i \in \mathbb{P}^m$. Equivalently, the vertex v containing the second "tail" is not labelled i.
- The set G_d^{i0} consists of loci for which an *irreducible component* of the source curve Σ containing the second marked point is collapsed to p_i. Equivalently, v is labelled i, and has valence at least 3.
- Finally, G_d^{i1} consists of loci for which the second marking is mapped to p_i without lying on a collapsed component. Equivalently, v is labelled i, and has valence 2.

Fig. 1 schematically illustrates the three subsets.



G_d^{i*} $\qquad$ G_d^{i0} $\qquad$ G_d^{i1}

FIGURE 1. The three types of graphs

Let $G_d^i = G_d^{i0} \cup G_d^{i1}$. Let G^{i0} and G^{i1} denote the unions $\bigcup_{d>0} G_d^{i0}$ and $\bigcup_{d>0} G_d^{i1}$ respectively. Notice that these graphs have the following basic properties with respect to integrals of type (29.9):

Type G_d^{i*}. Let $\Gamma \in G_d^{i*}$. As $\mathrm{ev}_2^*(\phi_i)$ vanishes when restricted to $\overline{\mathcal{M}}_\Gamma$, the contribution of Γ to Eq. (29.9) is 0.

Type G_d^{i0}. Let $\Gamma \in G_d^{i0}$. Let $v \in \mathrm{Vert}(\Gamma)$ be the vertex containing the second tail. The restriction of ψ_2 to $\overline{\mathcal{M}}_\Gamma$ carries the trivial $\mathbb{T}$-action. Hence a simple nilpotency result holds:

$$\psi_2^{\dim(v)+1} = 0 \in H_{\mathbb{T}}^*(\overline{\mathcal{M}}_\Gamma) \tag{29.10}$$

where $\dim(v) = \mathrm{val}(v) - 3$ is the dimension of $\overline{\mathcal{M}}_{0,\mathrm{val}(v)}$. As Γ has at most d edges and two tails (and no loops), $\mathrm{val}(v) \leq d+2$; equality holds only for graphs with d edges (each labelled 1) and both markings, all incident to v. The order of nilpotency of ψ_2 is thus at most d.

Type G_d^{i1}. Let $\Gamma \in G_d^{i1}$. Again let $v \in \mathrm{Vert}(\Gamma)$ be the vertex containing the second tail. It is incident to a unique edge e of Γ. Let v' be the other vertex of e, and let $j = \mu(v')$. If $d(e) < d$, let Γ_j be the two-pointed graph obtained by contracting (or "pruning") e: Γ_j is the complete subgraph of Γ not containing v, with the second marking placed at v'. The graph Γ_j is an element of $G_{d-d(e)}^j$. Note also that $|\mathrm{Aut}(\Gamma)| = |\mathrm{Aut}(\Gamma_j)|$. This pruning of graphs will give recursion relations for the correlators.

LEMMA 29.4.1 (Regularity). *The correlators $Z_i(e^t, \hbar)$ are naturally elements of the ring $\mathbb{Q}(\alpha, \hbar)[[e^t]]$:*

$$Z_i(e^t, \hbar) = 1 + \sum_d e^{dt} \zeta_{id}(\alpha, \hbar).$$

The rational functions $\zeta_{id}(\alpha, \hbar)$ are regular at all values $\hbar = \frac{\alpha_i - \alpha_j}{n}$ where $i \neq j$ and $n \geq 1$.

PROOF. Clearly, $\zeta_{id}(\alpha, \hbar) \in \mathbb{Q}[[\hbar^{-1}]]$:

$$\zeta_{id}(\alpha, \hbar) = \sum_{k=0}^{\infty} \int_{\overline{\mathcal{M}}_0(\mathbb{P}^m,d)} \frac{\psi_2^k \mathrm{e}(E'_d)}{\hbar^{k+1}} \mathrm{ev}_2^*(\phi_i).$$

Let $\Gamma \in G_d$. The contribution of Γ to ζ_{id} is :

$$\mathrm{Contr}_\Gamma(\zeta_{id}) = \sum_{k=0}^{\infty} \int_{\overline{\mathcal{M}}_\Gamma} \frac{\psi_2^k \mathrm{e}(E'_d)}{\hbar^{k+1} \mathrm{e}(N_\Gamma)} \mathrm{ev}_2^*(\phi_i).$$

By Type G_d^{i*} vanishing, we obtain:

$$\zeta_{id} = \sum_{\Gamma \in G_d^{i0}} \mathrm{Contr}_\Gamma(\zeta_{id}) + \sum_{\Gamma \in G_d^{i1}} \mathrm{Contr}_\Gamma(\zeta_{id}). \tag{29.11}$$

Let $\Gamma \in G_d^{i0}$. By the Type G_d^{i0} nilpotency condition,

$$\mathrm{Contr}_\Gamma(\zeta_{id}) = \sum_{k=0}^{d-1} \frac{P_{\Gamma,k}(\alpha)}{\hbar^{k+1}} \tag{29.12}$$

where $P_{\Gamma,k}(\alpha) \in \mathbb{Q}(\alpha)$. Let $\Gamma \in G_d^{i1}$. The restriction of ψ_2 to $\overline{\mathcal{M}}_\Gamma$ is topologically trivial with equivariant class $(\alpha_j - \alpha_i)/d(e)$ (Exercise 27.2.2). Hence, the contributions of Γ to the terms $k \geq 0$ form a geometric series, with sum

$$\mathrm{Contr}_\Gamma(\zeta_{id}) = \frac{P_\Gamma(\alpha)}{\left(\hbar + \frac{\alpha_i - \alpha_j}{d(e)}\right)}, \tag{29.13}$$

where $P_\Gamma(\alpha) \in \mathbb{Q}(\alpha)$. By Eqs. (29.11)–(29.13), $\zeta_{id} \in \mathbb{Q}(\alpha, \hbar)$. The explicit forms of Eq. (29.12) and Eq. (29.13) prove the regularity claim at $\hbar = (\alpha_i - \alpha_j)/n$. $\square$

29.4.1. Contributions of G_d^{i0} and G_d^{i1} in the Fano Cases (i) and (ii). If $l \leq m$ (cases (i) and (ii)), the contributions of G_d^{i0} and G_d^{i1} to the integrals in Eq. (29.8) yield linear recursion relations for the correlators $z_i(Q, \hbar)$. The contribution of G_d^{i0} will be the initial part of the relation. This contribution is analyzed first.

LEMMA 29.4.2. *In cases (i) and (ii), if $C_i(Q, \hbar)$ is the contribution of graph type G_d^{i0} to $z_i(Q, \hbar)$, then*

$$C_i(Q,\hbar) = \begin{cases} 0 & \text{if } l < m, \\ -1 + \exp\left(-m!Q + \frac{(m\alpha_i)^m}{\prod_{j\neq i}(\alpha_i - \alpha_j)} Q\right) & \text{if } l = m. \end{cases}$$

PROOF. Suppose $\Gamma \in G_d^{i0}$. If $d > 0$ and $l < m$, $(m+1-l)d - 1 \geq d$. The restriction of the integrand of Eq. (29.8) to $\overline{M}_\Gamma$ vanishes by Eq. (29.10) and the valence bound, so $C_i(Q, \hbar) = 0$.

If $l = m$, the $(1/\hbar)$-expansion of the integrand Eq. (29.8) contains only one possibly non-vanishing term after restriction to $\overline{\mathcal{M}}_\Gamma$: $\psi_2^{d-1} \mathrm{e}(E_d') \mathrm{ev}_2^*(\phi_i)$. This term also vanishes unless the valence of v is $d+2$, in which case (as remarked earlier) the only contributing graphs are those in which v has valence $d+2$ (i.e., v is incident to d edges marked 1, and the two tails). It is then straightforward to explicitly compute the contribution $C_i(Q, \hbar)$ from the combinatorics of these graphs via the localization formula. $\square$

Suppose now that $\Gamma \in G_d^{i1}$. If $d(e) = d$, then the contribution of Γ to z_i is $Q^d C_i^j(d, \hbar)$. Assume $d(e) < d$. Let Γ_j be the pruned graph obtained from $\Gamma \in G_d^{i1}$ as described in Type G_d^{i1} above. The linear recursion will be obtained from

$$\mathrm{Contr}_\Gamma(z_i(Q,\hbar)) = Q^{d(e)} C_i^j(d(e), \hbar) \cdot \mathrm{Contr}_{\Gamma_j}\left(z_j\left(Q, \frac{\alpha_j - \alpha_i}{d(e)}\right)\right), \tag{29.14}$$

where Contr_Γ denotes the contribution of Γ to the argument. The flag (v', e) in the graph Γ corresponds to a node in the domain curve; the normal bundle of $\overline{\mathcal{M}}_\Gamma \subset \overline{\mathcal{M}}_0(\mathbb{P}^m, d)$ has a line bundle quotient obtained from the deformation space of this node. This nodal deformation is absent in the normal bundle contributions for the graph Γ_j, but appears algebraically in the evaluation of the correlator z_j at $\hbar = (\alpha_j - \alpha_i)/d$.

EXERCISE 29.4.1. *Prove Eq. (29.14), using the graph-pruning strategy and the explicit recursions.*

By summing Eq. (29.14) over all graphs, we get linear recursion relations for $z_i(Q, \hbar)$:

$$z_i(Q,\hbar) = 1 + C_i(Q,\hbar) + \sum_{j\neq i} \sum_{d>0} Q^d C_i^j(d,\hbar) z_j\left(Q, \frac{\alpha_j - \alpha_i}{d}\right), \tag{29.15}$$

where the recursion coefficients are

$$C_i^j(d,\hbar) = \frac{1}{\left(\frac{\alpha_i - \alpha_j}{\hbar}\right) + d} \cdot \frac{\prod_{r=1}^{ld} \left(\frac{ld\alpha_i}{\alpha_j - \alpha_i}\right) + r}{\prod_{k=0}^{m} \prod_{r=1,\,(k,r)\neq(j,d)}^{d} \left(\frac{d(\alpha_i - \alpha_k)}{\alpha_j - \alpha_i} + r\right)}. \tag{29.16}$$

The initial term C_i in Eq. (29.15) is the contribution of G^{i0} to z_i, and the double sum is the contribution of G^{i1}.

The substitution $\hbar = (\alpha_j - \alpha_i)/d$ in Eq. (29.15) is well defined by the regularity lemma 29.4.1; it arises from a normal bundle factor in the Localization formula.

EXERCISE 29.4.2. *Show that Eq. (29.15) uniquely determines all z_i.* (Hint: *Use induction on increasing powers of Q.*)

29.4.2. Proof of Cases (i) and (ii). Define the correlators $Z_i^* \in \mathbb{Q}[[\hbar^{-1}, e^t]]$ by

$$Z_i^*(e^t, \hbar) = \sum_{d=0}^{\infty} e^{dt} \frac{\prod_{r=1}^{ld}(l\alpha_i + r\hbar)}{\prod_{j=0}^{m}\prod_{r=1}^{d}(\alpha_i - \alpha_j + r\hbar)}. \tag{29.17}$$

For all $l \leq m+1$, $Z_i^* \in \mathbb{Q}(\alpha, \hbar)[[e^t]]$, and the correlators Z_i^* satisfy the regularity property of the regularity lemma 29.4.1.

When $l \leq m$, let $z_i^*(Q, \hbar) = Z_i^*(Q\hbar^{(m+1-l)}, \hbar)$.

EXERCISE 29.4.3. *If $l < m$, show that the correlators $z_i^*(Q, \hbar)$ satisfy the recursion Eq. (29.15) (This is a direct algebraic computation.) Thus, as the recursions have a unique solution, $z_i^*(Q, \hbar) = z_i(Q, \hbar)$.*

Define the correlators $S^*(t, \hbar) \in H_{\mathbb{T}}^*(\mathbb{P}^m)[[\hbar^{-1}, t, e^t]]$ by

$$S^*(t, \hbar) = \sum_{d \geq 0} \frac{e^{(H/\hbar + d)t}\prod_{r=0}^{ld}(lH + r\hbar)}{\prod_{j=0}^{m}\prod_{r=1}^{d}(H - \alpha_j + r\hbar)}. \tag{29.18}$$

Then

$$\langle \phi_i, S^*(t, \hbar)\rangle = e^{\alpha_i t/\hbar} l\alpha_i Z_i^*(e^t, \hbar). \tag{29.19}$$

If $l < m$, we have in addition

$$\begin{aligned} e^{\alpha_i t/\hbar} l\alpha_i Z_i^*(e^t, \hbar) &= e^{\alpha_i t/\hbar} l\alpha_i Z_i(e^t, \hbar) \\ &= \langle \phi_i, S(t, \hbar)\rangle. \end{aligned}$$

Hence (by Exercise 27.1.5 (c)), $S^*(t, \hbar) = S(t, \hbar)$. As the non-equivariant correlator is recovered by setting the α_i to 0, this proves Case (i) of the Mirror conjecture.

For $l = m$, a direct calculation shows that the slightly modified correlator $e^{-m!Q} z_i^*(Q, \hbar)$ satisfies Eq. (29.15). Thus $e^{-m!Q} z_i^*(Q, \hbar) = z_i(Q, \hbar)$ by uniqueness, and the equality $e^{-m!e^t/\hbar} S^*(t, \hbar) = S(t, \hbar)$ then follows analogously, proving Case (ii) of the Mirror conjecture.

CHAPTER 30

The Mirror Conjecture for Hypersurfaces II: The Calabi–Yau Case

30.1. Correlator Recursions

Write Z_i in partially expanded form:

$$Z_i(e^t, \hbar) = 1 + \sum_{d>0} e^{dt} \left(\sum_{k=0}^{d-1} \hbar^{-k-1} \int_{\overline{\mathcal{M}}_0(\mathbb{P}^m, d)} \psi_2^k \mathrm{e}(E_d')\mathrm{ev}_2^*(\phi_i) \right) \tag{30.1}$$

$$+ \sum_{d>0} \left(\frac{e^t}{\hbar}\right)^d \int_{\overline{\mathcal{M}}_0(\mathbb{P}^m, d)} \frac{\psi_2^d}{\hbar - \psi_2} \mathrm{e}(E_d')\mathrm{ev}_2^*(\phi_i). \tag{30.2}$$

Define $z_i(Q, \hbar) = Z_i(Q\hbar, \hbar)$.

WARNING 30.1.1. This is a *different* z_i from the proofs of Cases (i) and (ii) of the Mirror conjecture, although its role will be similar.

LEMMA 30.1.1 (z-Recursion). *The linear recursions for z_i in the Calabi–Yau case take the form*

$$z_i(Q, \hbar) = 1 + \sum_{d>0} \frac{Q^d}{d!} R_{id} + \sum_{d>0}\sum_{j \neq i} Q^d C_i^j(d, \hbar) z_j\left(Q, \frac{\alpha_j - \alpha_i}{d}\right),$$

where $R_{id} = \sum_{j=0}^{d} R_{id}^j \hbar^{d-j}$ is a polynomial in $\mathbb{Q}[\hbar]$ of $\hbar$-degree (at most) d, and

$$C_i^j(d, \hbar) = \frac{1}{\alpha_i - \alpha_j + d\hbar} \frac{\prod_{r=1}^{(m+1)d}\left((m+1)\alpha_i + r\frac{\alpha_j - \alpha_i}{d}\right)}{d! \prod_{k \neq i}\prod_{r=1, (k,r) \neq (j,d)}^{d}\left(\alpha_i - \alpha_k + r\frac{\alpha_j - \alpha_i}{d}\right)}. \tag{30.3}$$

The proof of this recursion relation is similar (but not identical) to the proof of Eq. (29.15), which was derived by separating the contribution of graph types G_d^{i0} and G_d^{i1}. Here, we instead separate the contributions of the terms (30.1) and (30.2) in the expansion of $Z_i(e^t, \hbar)$.

EXERCISE 30.1.1. *Show that the contribution of the sum (30.1) is $1 + \sum_{d>0} Q^d R_{id}/d!$ for some polynomials R_{id}.*

We next analyze the contribution of (30.2). The G_d^{i0}-type contributions vanish since the power of ψ_2 appearing is too high (see Sec. 29.4). Hence only graphs of type G_d^{i1} contribute to (30.2); once again, we count them recursively by "pruning".

In this case, we include the details of the required localization calculation. Let $\Gamma \in G_d^{i1}$. Two equations, (30.4) and (30.5), are needed.

Suppose first that Γ is the unique graph of type G_d^{i1} with a single edge e connecting fixed points i and $j \neq i$ and satisfying $d(e) = d$. We now prove:

$$\text{(30.4)} \qquad \mathrm{Contr}_\Gamma\left(Q^d \int_{\overline{\mathcal{M}}_0(\mathbb{P}^m,d)} \frac{\psi_2^d}{\hbar - \psi_2} \mathbf{e}(E'_d)\mathrm{ev}_2^*(\phi_i)\right) = Q^d C_i^j(d,\hbar).$$

The Deligne–Mumford stack $\overline{\mathcal{M}}_\Gamma$ is zero-dimensional with automorphism group $\mathbb{A}_\Gamma$ of order d (see Sec. 27.3). Let

$$f : (\Sigma, x_1, x_2) \to \mathbb{P}^m$$

be the fixed map corresponding to Γ. Note that

$$\mathbf{e}(E'_d)|_{\overline{\mathcal{M}}_\Gamma} = \prod_{r=1}^{(m+1)d}\left((m+1)\alpha_i + r\frac{\alpha_j - \alpha_i}{d}\right).$$

From Eq. (27.8), after a short calculation,

$$\begin{aligned}
\mathbf{e}(N_\Gamma) &= \prod_{k=0}^{m} \prod_{r=0,(k,r)\neq(i,0),(j,d)}^{d} \left(\alpha_i - \alpha_k + r\frac{\alpha_j - \alpha_i}{d}\right) \\
&= d!\left(\frac{\alpha_j - \alpha_i}{d}\right)^d \cdot \prod_{k\neq i}(\alpha_i - \alpha_k) \\
&\quad \cdot \prod_{k\neq i} \prod_{r=1,(k,r)\neq(j,d)}^{d} \left(\alpha_i - \alpha_k + r\frac{\alpha_j - \alpha_i}{d}\right).
\end{aligned}$$

The classes ψ_2 and $\mathrm{ev}_2^*(\phi_i)$ restrict to $(\alpha_j - \alpha_i)/d$ and $\prod_{k\neq i}(\alpha_i - \alpha_k)$ respectively. Since

$$\begin{aligned}
&\mathrm{Contr}_\Gamma\left(Q^d \int_{\overline{\mathcal{M}}_0(\mathbb{P}^m,d)} \frac{\psi_2^d}{\hbar - \psi_2} \mathbf{e}(E'_d)\mathrm{ev}_2^*(\phi_i)\right) \\
&= \frac{Q^d}{|\mathbb{A}_\Gamma|} \int_{\overline{\mathcal{M}}_\Gamma} \frac{\psi_2^d}{(\hbar - \psi_2)\mathbf{e}(N_\Gamma)} \mathbf{e}(E'_d)\mathrm{ev}_2^*(\phi_i),
\end{aligned}$$

Eq. (30.4) follows as an algebraic consequence of these weight calculations (pulled back to $H_{\mathbb{T}}^*(\overline{\mathcal{M}}_\Gamma)$).

Suppose next that $\Gamma \in G_d^{i1}$ satisfies $d(e) < d$. We will show that

$$\begin{aligned}
\text{(30.5)} \quad &\mathrm{Contr}_\Gamma\left(Q^d \int_{\overline{\mathcal{M}}_0(\mathbb{P}^m,d)} \frac{\psi_2^d}{\hbar - \psi_2} \mathbf{e}(E'_d)\mathrm{ev}_2^*(\phi_i)\right) = Q^{d(e)} C_i^j(d(e),\hbar) \\
&\cdot \mathrm{Contr}_{\Gamma_j}\left(\left(\frac{\alpha_j - \alpha_i}{d(e)} Q\right)^{d-d(e)} \int_{\overline{\mathcal{M}}_{d-d(e)}} \frac{\mathbf{e}(E'_{d-d(e)})}{\frac{\alpha_j - \alpha_i}{d(e)} - \psi_2} \mathrm{ev}_2^*(\phi_j)\right).
\end{aligned}$$

Standard weight calculations (via a natural restriction sequence of sections of $f^*(\mathcal{O}_{\mathbb{P}^m}(m+1))$ to the component corresponding to edge e) yield:

$$\text{(30.6)} \ \mathbf{e}(E'_d)|_{\overline{\mathcal{M}}_\Gamma} = \mathbf{e}(E'_{d-d(e)})|_{\overline{\mathcal{M}}_{\Gamma_j}} \cdot \prod_{r=1}^{(m+1)d(e)} \left((m+1)\alpha_i + r\left(\frac{\alpha_j - \alpha_i}{d(e)}\right)\right).$$

From our formula (27.8) for the Euler class of the normal bundle to the fixed locus, again after a short calculation,

$$\begin{aligned}
\frac{\mathrm{ev}_2^*(\phi_i)}{\mathbf{e}(N_\Gamma)} &= \frac{1}{\frac{\alpha_j - \alpha_i}{d(e)} - \psi_2} \frac{\mathrm{ev}_2^*(\phi_j)}{\mathbf{e}(N_{\Gamma_j})} \\
&\cdot \frac{1}{d(e)!\left(\frac{\alpha_j - \alpha_i}{d(e)}\right)^{d(e)} \cdot \prod_{k\neq i}\prod_{r=1,(k,r)\neq(j,d(e))}^{d(e)}\left(\alpha_i - \alpha_k + r\frac{\alpha_j - \alpha_i}{d(e)}\right)},
\end{aligned}$$

where the left and right sides are naturally classes on $\overline{\mathcal{M}}_\Gamma$ and $\overline{\mathcal{M}}_{\Gamma_j}$ respectively. The first term on the right is the nodal deformation corresponding to the pruned node.

Finally, $|\mathbb{A}_\Gamma| = d(e)|\mathbb{A}_{\Gamma_j}|$. Eq. (30.5) now follows.

The linear recursions are obtained from Eqs. (30.4) and (30.5) by summing over graphs of type G_d^{i1}. This completes the proof of the z-recursion lemma 30.1.1. □

30.2. Polynomiality

The Calabi–Yau case is difficult for several reasons. The recursion relations for z_i are not yet determined as the functions R_{id} are unknown. It is necessary to find additional conditions satisfied by the correlators z_i. Givental's idea here is to prove a polynomiality constraint satisfied by a related double correlator Φ. Define $\Phi(z, e^t) \in \mathbb{Q}(\alpha,\hbar)[[z, e^t]]$ by

$$\text{(30.7)} \qquad \Phi(z, e^t) = \sum_{i=0}^{m} \frac{(m+1)\alpha_i}{\prod_{j\neq i}(\alpha_i - \alpha_j)} e^{\alpha_i z} Z_i(e^{t+z\hbar}, \hbar) Z_i(e^t, -\hbar).$$

We will find a "polynomiality" constraint on $\Phi(z, e^t)$ that can be interpreted as a further condition on the correlators z_i.

A geometric construction is needed for the polynomiality constraint. Consider a new one-dimensional torus $\mathbb{C}^*$. Let $\mathbb{Q}[\hbar]$ be the equivariant cohomology ring of $\mathbb{C}^*$. (Again, $\hbar$ is the first Chern class of the dual of the standard representation of $\mathbb{C}^*$.) Let $\mathbb{C}^*$ act on the vector space $V = \mathbb{C}^2$ via the exponential weights $(0, -1)$, and let y_1, y_2 be the respective fixed points for the induced action on $\mathbb{P}^1 = \mathbb{P}V$. The equivariant Chern classes of the tangent representations at the fixed points are $\hbar, -\hbar$ respectively. There are naturally induced $(\mathbb{C}^* \times \mathbb{T})$-actions on $\mathbb{P}V \times \mathbb{P}^m$ and $\overline{\mathcal{M}}_{0,2}(\mathbb{P}V \times \mathbb{P}^m, (1, d))$, induced by the $\mathbb{C}^*$ action on V, and the $\mathbb{T}$-action on $\mathbb{C}^{m+1}$ of Eq. (27.1). The space of interest to us will be

$$L_d = \mathrm{ev}_1^{-1}(\{y_1\} \times \mathbb{P}^m) \cap \mathrm{ev}_2^{-1}(\{y_2\} \times \mathbb{P}^m) \subset \overline{\mathcal{M}}_{0,2}(\mathbb{P}V \times \mathbb{P}^m, (1, d)).$$

L_d is easily seen to be a non-singular, $(\mathbb{C}^* \times \mathbb{T})$-equivariant substack.

Let L'_d denote the polynomial space $\mathbb{P}(\mathbb{C}^{m+1} \otimes \mathrm{Sym}^d(V^*))$ with the canonical $(\mathbb{C}^* \times \mathbb{T})$-representation. A degree d algebraic map $\mathbb{P}V \to \mathbb{P}^m$ canonically yields a point in L'_d. There is a natural $(\mathbb{C}^* \times \mathbb{T})$-equivariant morphism

$$f : \mathcal{M}_{0,2}(\mathbb{P}V \times \mathbb{P}^m, (1, d)) \to L'_d$$

obtained by identifying an element of the left moduli space with the graph of a uniquely determined map $\mathbb{P}V \to \mathbb{P}^m$. It may be shown that f extends to a $(\mathbb{C}^* \times \mathbb{T})$-equivariant morphism from the Deligne–Mumford stack $\overline{\mathcal{M}}_{0,2}(\mathbb{P}V \times \mathbb{P}^m, (1, d))$. Let $f : L_d \to L'_d$ be the induced map and let $P \in H^*_{\mathbb{C}^* \times \mathbb{T}}(L'_d)$ be the equivariant first Chern class of $\mathcal{O}_{L'_d}(1)$. Let E_d be the equivariant bundle on L_d with fiber over a stable map $[(f_{\mathbb{P}V} \times f_{\mathbb{P}^m}) : \Sigma \to \mathbb{P}V \times \mathbb{P}^m]$ equal to $H^0(\Sigma, f^*_{\mathbb{P}^m}(\mathcal{O}_{\mathbb{P}^m}(m+1)))$.

LEMMA 30.2.1.

$$\Phi(z, e^t) = \sum_{d \geq 0} e^{dt} \int_{L_d} e^{f^*(P) \cdot z} \mathbf{e}(E_d), \tag{30.8}$$

where the integral on the right is the $(\mathbb{C}^* \times \mathbb{T})$-equivariant push-forward to a point.

PROOF. The remarkable feature of this equality is the following. On the left side of Eq. (30.8), $\hbar$ is a formal parameter. On the right side, it is an element of equivariant cohomology. As L_d is a non-singular Deligne–Mumford stack, the $(\mathbb{C}^* \times \mathbb{T})$-localization formula yields an explicit graph summation form for the integral on the right that is directly matched with Eq. (30.7).

The first step is to identify the graph types of the fixed loci of L_d. Recall the definitions of G^{i0}_d and G^{i1}_d from Sec. 29.4. Let $G^i_d = G^{i0}_d \cup G^{i1}_d \cup \{\mathrm{Triv}(i)\}$ where $\mathrm{Triv}(i)$ is the edgeless two pointed graph with a single vertex v satisfying $f(v) = p_i$. Let $\deg(\mathrm{Triv(i)}) = 0$. The components of $L_d^{\mathbb{C}^* \times \mathbb{T}}$ are in bijective correspondence to triples (i, Γ_1, Γ_2) where $0 \leq i \leq m$ and $\Gamma_1, \Gamma_2 \in G^i_d$ satisfy $\deg(\Gamma_1) + \deg(\Gamma_2) = d$. The graphs Γ_1, Γ_2 describe the configurations lying over the points $y_1, y_2 \in \mathbb{P}V$ respectively. A fixed map

$$f : (\Sigma, x_1, x_2) \to \mathbb{P}V \times \mathbb{P}^m$$

in the corresponding component satisfies the following properties. The domain is a union of three subcurves $\Sigma = \Sigma_1 \cup \Sigma_m \cup \Sigma_2$. The curve Σ_m is mapped isomorphically by f to $\mathbb{P}V \times \{p_i\}$. Σ_1 and Σ_2 contain x_1 and x_2 and lie over y_1 and y_2 respectively. Lemma 30.2.1 will follow from the calculation of the contribution of (i, Γ_1, Γ_2) to the integral in Eq. (30.8).

Let $\mathbf{\Gamma} = (i, \Gamma_1, \Gamma_2)$. Let d_1, d_2 equal $\deg(\Gamma_1), \deg(\Gamma_2)$ respectively. We treat the generic case: $d_1, d_2 > 0$. The degenerate cases in which either Γ_1 or Γ_2 equals $\mathrm{Triv}(i)$ are computed analogously. The "contribution equation" is:

$$\begin{aligned}\mathrm{Cont}_{\mathbf{\Gamma}}\left(e^{dt} \int_{L_d} e^{f^*(P) \cdot z} \mathbf{e}(E_d)\right) &= \frac{(m+1)\alpha_i}{\prod_{j \neq i}(\alpha_i - \alpha_j)} e^{\alpha_i z} \\ &\cdot (e^{t + z\hbar})^{d_1} \cdot \mathrm{Cont}_{\Gamma_1}\left(\int_{\overline{\mathcal{M}}_{d_1}} \frac{\mathbf{e}(E'_{d_1})}{\hbar - \psi_2} \mathrm{ev}_2^*(\phi_i)\right) \\ &\cdot e^{d_2 t} \cdot \mathrm{Cont}_{\Gamma_2}\left(\int_{\overline{\mathcal{M}}_{d_2}} \frac{\mathbf{e}(E'_{d_2})}{-\hbar - \psi_2} \mathrm{ev}_2^*(\phi_i)\right).\end{aligned}$$

The "contribution equation" in the degenerate cases is identical (with the convention $\mathrm{Cont}_{\mathrm{Triv(i)}} = 1$).

The equation is proven by expanding the localization formula for the left side. Note first that the fixed stack $\overline{\mathcal{M}}_{\mathbf{\Gamma}} \subset L_d$ is naturally isomorphic to $\overline{\mathcal{M}}_{\Gamma_1} \times \overline{\mathcal{M}}_{\Gamma_2}$. As

$$f(\overline{\mathcal{M}}_{\mathbf{\Gamma}}) = [\Sigma_i \otimes [(y_1^*)^{d_2}(y_2^*)^{d_1}]],$$

the class $f^*(P)$ is pure weight equal to $\alpha_i + d_1\hbar$. The class $\mathbf{e}(E_d)|_{\overline{\mathcal{M}}_\Gamma}$ is pure weight and factors as

$$(m+1)\alpha_i \cdot \mathbf{e}(E'_{d_1})|_{\overline{\mathcal{M}}_{\Gamma_1}} \cdot \mathbf{e}(E'_{d_2})|_{\overline{\mathcal{M}}_{\Gamma_2}}$$

by the restriction sequence to Σ_m. Similarly,

$$\frac{\prod_{j\neq i}(\alpha_i - \alpha_j)}{\mathbf{e}(N_\Gamma)}$$

is computed to equal the product of $\mathrm{ev}_2^*(\phi_i)/((\hbar - \psi_2)\mathbf{e}(N_{\Gamma_1}))$ from $\overline{\mathcal{M}}_{\Gamma_1}$ with $\mathrm{ev}_2^*(\phi_i)/((-\hbar - \psi_2)\mathbf{e}(N_{\Gamma_2}))$ from $\overline{\mathcal{M}}_{\Gamma_2}$. This normal bundle expression is obtained by the restriction sequence of tangent sections to Σ_m and an accounting of nodal deformations. As N_Γ is the normal bundle in L_d, only tangent sections of $H^0(\Sigma, f^*T_{\mathbb{P}V})$ vanishing at the markings x_1 and x_2 appear in the normal bundle expression. The "contribution equation" now follows directly.

We finally obtain Eq. (30.8) from the "contribution equation", the definition of $Z_i(e^t, \hbar)$, and a sum over graphs. □

By Lemma 30.2.1, $\Phi(z, e^t)$ may be rewritten as

$$\Phi(z, e^t) = \sum_{d\geq 0} e^{dt} \int_{L'_d} e^{Pz} f_*(\mathbf{e}(E_d)). \tag{30.9}$$

The group $\mathbb{C}^* \times \mathbb{T}$ acts with $(m+1)(d+1)$ isolated fixed points on L'_d. A weight calculation of the representation $\mathbb{C}^{m+1} \otimes \mathrm{Sym}^d(V^*)$ yields the standard presentation

$$H^*_{\mathbb{C}^*\times\mathbb{T}}(L'_d) = \mathbb{Q}[P, \alpha, \hbar] / \left(\prod_{j=0}^{m} \prod_{r=0}^{d} (P - \alpha_j - r\hbar) \right).$$

As $f_*(\mathbf{e}(E_d)) \in H^{(m+1)d+1}_{\mathbb{C}^*\times\mathbb{T}}(L'_d)$, there is a unique polynomial

$$E^Z_d(P, \hbar, \alpha) \in \mathbb{Q}[P, \alpha, \hbar]$$

of homogeneous degree $(m+1)d + 1$ satisfying $f_*(\mathbf{e}(E_d)) = E^Z_d(P, \alpha, \hbar)$ in $H^*_{\mathbb{C}^*\times\mathbb{T}}(L'_d)$. The localization formula Eq. (4.4) for the integral in Eq. (30.9) then yields

$$\Phi(z, e^t) = \frac{1}{2\pi i} \oint e^{Pz} \sum_{d\geq 0} \frac{e^{dt} E^Z_d(P, \alpha, \hbar)}{\prod_{j=0}^{m} \prod_{r=0}^{d} (P - \alpha_j - r\hbar)} dP. \tag{30.10}$$

Givental's polynomiality constraint is the following: $\Phi(z, e^t)$ is expressible as a residue integral of the form Eq. (30.10) where $E^Z_d(P, \alpha, \hbar) \in \mathbb{Q}[P, \alpha, \hbar]$ is of P-degree at most $(m+1)d + m$.

30.3. Correlators of Class $\mathcal{P}$

Let $\{Y_i(e^t, \hbar)\}_{i=0}^m \subset \mathbb{Q}[[\hbar^{-1}, e^t]]$ be a set of functions (called correlators). Assume the correlators Y_i satisfy the rationality and regularity conditions of the regularity lemma 29.4.1: $Y_i \in \mathbb{Q}(\alpha, \hbar)[[e^t]]$ with no poles at $\hbar = (\alpha_i - \alpha_j)/n$ (for all $j \neq i$ and $n \geq 1$). Let $y_i(Q, \hbar) = Y_i(Q\hbar, \hbar)$, and suppose y_i satisfies the recursion relation

$$y_i(Q, \hbar) = 1 + \sum_{d>0} \frac{Q^d}{d!} I_{id} + \sum_{d>0} \sum_{j\neq i} Q^d C^j_i(d, \hbar) y_j \left(Q, \frac{\alpha_j - \alpha_i}{d}\right), \tag{30.11}$$

where $I_{id} = \sum_{j=0}^{d} I^j_{id} \hbar^{d-j} \in \mathbb{Q}(\alpha)[\hbar]$ is an element of $\hbar$-degree at most d. The recursions Eq. (30.11) determine y_i uniquely from the initial data I_{id} (by the same method as in Exercise 29.4.2). A direct algebraic consequence of Eq. (30.11) is the existence of a unique expression,

$$y_i(Q, \hbar) = \sum_{d\geq 0} Q^d \frac{N_{id}}{d! \prod_{j\neq i} \prod_{r=1}^{d} (\alpha_i - \alpha_j + r\hbar)}, \tag{30.12}$$

where $N_{id} \in \mathbb{Q}(\alpha)[\hbar]$ is a polynomial of $\hbar$-degree at most $(m+1)d$, and $N_{i0} = 1$. We may also consider the double correlator $\Phi^Y \in \mathbb{Q}(\alpha, \hbar)[[z, e^t]]$:

$$\Phi^Y(z, e^t) = \sum_{i=0}^{m} \frac{(m+1)\alpha_i}{\prod_{j\neq i}(\alpha_i - \alpha_j)} e^{\alpha_i z} Y_i(e^{t+z\hbar}, \hbar) Y_i(e^t, -\hbar). \tag{30.13}$$

After the substitution of Eq. (30.12) in Eq. (30.13), a straightforward algebraic computation shows that

$$\Phi^Y(z, e^t) = \frac{1}{2\pi i} \oint e^{Pz} \sum_{d\geq 0} \frac{e^{dt} E^Y_d(P, \alpha, \hbar)}{\prod_{j=0}^{m} \prod_{r=0}^{d} (P - \alpha_j - r\hbar)} dP, \tag{30.14}$$

where $E^Y_d = \sum_{k=0}^{(m+1)d+m} f_k(\alpha, \hbar) P^k$ is the unique function of P-degree at most $(m+1)d+m$ determined by the values at the $(m+1)(d+1)$ evaluations $P = \alpha_i + r\hbar$ $(0 \leq i \leq m,\ 0 \leq r \leq d)$:

$$E^Y_d(\alpha_i + r\hbar) = (m+1)\alpha_i N_{ir}(\hbar) N_{i(d-r)}(-\hbar). \tag{30.15}$$

In general, the coefficients $f_k(\alpha, \hbar) \in \mathbb{Q}(\alpha, \hbar)$ will be rational functions. The correlators Y_i satisfy Givental's polynomiality condition if $E^Y_d \in \mathbb{Q}[P, \alpha, \hbar]$.

LEMMA 30.3.1. *The correlators Y_i satisfy Givental's polynomiality condition if and only if $\Phi^Y(z, e^t) \in \mathbb{Q}[\alpha, \hbar][[z, e^t]]$.*

PROOF. By the localization formula Eq. (4.4), the integral

$$\frac{1}{2\pi i} \oint \sum_{d \geq 0} \frac{P^k}{\prod_{j=0}^m \prod_{r=0}^d (P - \alpha_j - r\hbar)} dP \tag{30.16}$$

computes the $(\mathbb{C}^* \times \mathbb{T})$-equivariant push-forward to a point of the class $P^k \in H^*_{\mathbb{C}^* \times \mathbb{T}}(L'_d)$. We therefore see:

(a) for $k < (m+1)d + m$, (30.16) vanishes,

(b) for $k = (m+1)d + m$, (30.16) equals 1,

(c) for $k > (m+1)d + m$, (30.16) is an element of $\mathbb{Q}[\alpha, \hbar]$.

Expand the integrand of (30.14) in power series by $e^{Pz} = \sum_{k=0}^{\infty} (Pz)^k / k!$. Properties (a)–(c) then prove that the polynomiality of the coefficients of $E_d^Y = \sum_{k=0}^{(m+1)d+m} f_k(\alpha, \hbar) P^k$ is equivalent to the polynomiality of all coefficients of the terms $\{z^k e^{dt}\}_{k=0}^{\infty}$ in $\Phi^Y(z, e^t)$. □

A set of correlators $Y_i \in \mathbb{Q}[[\hbar^{-1}, e^t]]$ is defined to be of class $\mathcal{P}$ if the following three conditions are satisfied.

I. The rationality and regularity conditions hold.

II. The correlators y_i satisfy relations of the form Eq. (30.11).

III. Givental's polynomiality condition is met.

A suitable interpretation of II actually implies I, but we separate these conditions for clarity.

The most important property of class $\mathcal{P}$ is Givental's uniqueness result.

LEMMA 30.3.2 (Uniqueness). *Let $Y_i, \overline{Y}_i \in \mathbb{Q}[[\hbar^{-1}, e^t]]$ be two sets of correlators of class $\mathcal{P}$. If for all i,*

$$Y_i \equiv \overline{Y}_i \pmod{\hbar^{-2}}, \tag{30.17}$$

then the sets of correlators agree identically: $Y_i = \overline{Y}_i$.

The class $\mathcal{P}$ assumption is essential.

PROOF. Let I_{id} and $\overline{I}_{id}$ be the respective initial data in the associated recursions Eq. (30.11). By the recursion formula Eq. (30.11) and the coefficient formula Eq. (30.3), we obtain the equality

$$Y_i \equiv \sum_{d \geq 0} e^{dt} \left(I_{id}^0 + \frac{I_{id}^1}{\hbar} \right) \pmod{\hbar^{-2}} \tag{30.18}$$

(and analogously for $\overline{Y}_i$). Assumption Eq. (30.17) therefore implies $I_{id}^0 = \overline{I}_{id}^0$ and $I_{id}^1 = \overline{I}_{id}^1$ for all i and d. In particular, $I_{i1} = \overline{I}_{i1}$.

To establish the lemma, we prove that $I_{id} = \overline{I}_{id}$ by induction. Assume $I_{ik} = \overline{I}_{ik}$ for all $0 \leq i \leq m$ and $k < d$. The equality $N_{ik} = \overline{N}_{ik}$ for $k < d$ then follows from the recursions. By Eq. (30.15), $\delta E_d = E_d^Y - E_d^{\overline{Y}}$ vanishes at $P = \alpha_i + r\hbar$ for all i and $1 \leq r \leq d-1$. Hence, the polynomial δE_d is divisible by $\prod_{j=0}^m \prod_{r=1}^{d-1} (P - \alpha_j - r\hbar)$. By Eq. (30.15) and the recursion Eq. (30.11), a computation shows that

$$\delta E_d(P = \alpha_i + d\hbar) = (m+1)\alpha_i \prod_{j \neq i} \prod_{r=1}^{d} (\alpha_i - \alpha_j + r\hbar)(I_{id} - \overline{I}_{id}).$$

By the polynomiality condition $\delta E_d \in \mathbb{Q}[P, \alpha, \hbar]$ and the above divisibility, we find $\hbar^{d-1}$ divides $I_{id} - \overline{I}_{id}$. Therefore the initial data is allowed to differ only in the $\hbar^d$ and $\hbar^{d-1}$ coefficients. However, these coefficients are precisely the two appearing in Eq. (30.18) which agree by assumption Eq. (30.17). We have proven the equality $I_{id} = \overline{I}_{id}$, completing the induction. □

By the results of Sec. 30.2, the correlators $Z_i(e^t, \hbar)$ are of class $\mathcal{P}$. Recall the hypergeometric correlators $Z_i^*(e^t, \hbar)$ defined by Eq. (29.17). A straightforward exercise in algebra shows the correlators Z_i^* also to be of class $\mathcal{P}$. The polynomials $E_d^{Z^*}(P, \alpha, \hbar)$ associated to the correlators Z_i^* are

$$E_d^{Z^*} = \prod_{r=0}^{(m+1)d} ((m+1)P - r\hbar).$$

The two sets of correlators Z_i, Z_i^* do not agree modulo $\hbar^{-2}$. The expansions modulo $\hbar^{-2}$ may be explicitly evaluated. From expression Eq. (30.1), the $\hbar^0$ term in Z_i is 1. The $\hbar^{-1}$ term in Eq. (30.1) vanishes since the classes in the relevant integrals over $\overline{\mathcal{M}}_d$ are pulled back via the map forgetting the first marking. Hence, $Z_i \equiv 1 \pmod{\hbar^{-2}}$. A direct computation yields

$$Z_i^* \equiv F(e^t) + \frac{\alpha_i(m+1)(G_{m+1}(e^t) - G_1(e^t)) + G_1(e^t)\sum_{j=0}^m \alpha_j}{\hbar} \pmod{\hbar^{-2}},$$

where the functions $F(e^t)$ and $G_l(e^t)$ are defined by

$$F(e^t) = \sum_{d=0}^{\infty} e^{dt} \frac{((m+1)d)!}{(d!)^{m+1}}, \quad G_l(e^t) = \sum_{d=1}^{\infty} e^{dt} \frac{((m+1)d)!}{(d!)^{m+1}} \left(\sum_{r=1}^{ld} \frac{1}{r} \right).$$

The last step in the proof of the Calabi–Yau case (iii) is the following. An explicit transformation $\overline{Z}_i$ of the correlator Z_i is found that satisfies

(1) $\overline{Z}_i$ is of class $\mathcal{P}$,
(2) $\overline{Z}_i \equiv Z_i^*$ (mod $\hbar^{-2}$).

Then, by the uniqueness lemma 30.3.2, $\overline{Z}_i = Z_i^*$. This transformation will yield the mirror prediction in the quintic threefold case.

30.4. Transformations

Let Y_i be a set of correlators of class $\mathcal{P}$. We define three transformations:

(a) $\overline{Y}_i(e^t, \hbar) = f(e^t)Y_i(e^t, \hbar)$,
(b) $\overline{Y}_i(e^t, \hbar) = \exp(\alpha_i g(e^t)/\hbar)Y_i(qe^{g(e^t)}, \hbar)$,
(c) $\overline{Y}_i(e^t, \hbar) = \exp(Cg(e^t)/\hbar)Y_i(e^t, \hbar)$,

where $f(e^t), g(e^t) \in \mathbb{Q}[[e^t]]$ satisfy $f(0) = 1$ and $g(0) = 0$, and C is a homogeneous linear function of the α's.

LEMMA 30.4.1. *In each case (a)–(c), $\overline{Y}_i$ is a set of correlators of class $\mathcal{P}$.*

PROOF. Since rational functions in $\alpha, \hbar$ satisfying the regularity condition of the regularity lemma 29.4.1 form a subring, the correlators $\overline{Y}_i$ clearly satisfy condition I of class $\mathcal{P}$. A direct algebraic check shows the correlators $\overline{y}_i$ satisfy recursion relations of the form Eq. (30.11). The initial terms $\overline{I}_{id}$ change, but remain in $\mathbb{Q}(\alpha)[\hbar]$ of $\hbar$-degree at most d. The values $f(0) = 1$ and $g(0) = 0$ are needed for this verification. Condition II therefore holds for $\overline{Y}_i$.

Condition III of class $\mathcal{P}$ is checked using Lemma 30.3.1. The transformations (a)–(c) have the following effect on the double correlator:

(a) $\Phi^{\overline{Y}}(z, e^t) = f(e^{t+z\hbar})f(e^t) \cdot \Phi^Y(z, e^t)$,
(b) $\Phi^{\overline{Y}}(z, e^t) = \Phi^Y(z + (g(e^{t+z\hbar}) - g(e^t))/\hbar, e^{t+g(e^t)})$,
(c) $\Phi^{\overline{Y}}(z, e^t) = \exp(C \cdot (g(e^{t+z\hbar}) - g(e^t))/\hbar) \cdot \Phi^Y(z, e^t)$.

In each case, $\Phi^{\overline{Y}}$ is easily seen to remain in $\mathbb{Q}[\alpha, \hbar][[z, e^t]]$. Case (a) is clear. Since

$$\frac{g(e^{t+z\hbar}) - g(e^t)}{\hbar} \in \mathbb{Q}[\alpha, \hbar][[z, e^t]],$$

the change of variables in case (b) and multiplication in case (c) preserve membership in $\mathbb{Q}[\alpha, \hbar][[z, e^t]]$. □

The transformation from $Z_i(e^t, \hbar)$ to $Z_i^*(e^t, \hbar)$ can now be established. Define the correlators $\overline{Z}_i$ by

$$\overline{Z}_i(e^t, \hbar) = F(e^t)\cdot\exp\left(\frac{(m+1)\alpha_i(G_{m+1}(e^t) - G_1(e^t)) + G_1(e^t)\sum_{j=0}^m \alpha_j}{\hbar F(e^t)}\right)$$
$$\cdot Z_i\left(e^t \exp\left(\frac{(m+1)(G_{m+1}(e^t) - G_1(e^t))}{F(e^t)}\right), \hbar\right).$$

By a composition of transformations established in Lemma 30.4.1, the correlators $\overline{Z}_i$ are of class $\mathcal{P}$. An explicit calculation using the results of Sec. 30.3 shows $\overline{Z}_i(e^t, \hbar) \equiv Z_i^*(e^t, \hbar)$ (mod $\hbar^{-2}$). By the uniqueness lemma 30.3.2, $\overline{Z}_i(e^t, \hbar) = Z_i^*(e^t, \hbar)$.

Consider the change of variables defined by

$$T = t + \frac{(m+1)(G_{m+1}(e^t) - G_1(e^t))}{F(e^t)}. \tag{30.19}$$

Exponentiating Eq. (30.19) yields

$$e^T = e^t \cdot \exp\left(\frac{(m+1)(G_{m+1}(e^t) - G_1(e^t))}{F(e^t)}\right).$$

Together these two formulas define a change of variables from formal series in T, e^T to formal series in t, e^t. This transformation is easily seen to be invertible.

Let $S(T, \hbar) \in H_{\mathbb{T}}^*(\mathbb{P}^m)[[\hbar^{-1}, T, e^T]]$ be the equivariant correlator (29.6) in the variable T. Let the correlator $\overline{S}(t, \hbar) \in H_{\mathbb{T}}^*(\mathbb{P}^m)[[\hbar^{-1}, t, e^t]]$ be obtained from $S(T, \hbar)$ by the change of variables Eq. (30.19) followed by multiplication by the function

$$F(e^t) \cdot \exp\left(\frac{G_1(e^t)\sum_{j=0}^m \alpha_j}{\hbar F(e^t)}\right).$$

By Eq. (29.19) and the definition of $\overline{Z}_i$, we find

$$\langle \phi_i, \overline{S}(t, \hbar)\rangle = e^{\alpha_i t/\hbar} l\alpha_i \overline{Z}_i(e^t, \hbar).$$

Consider the correlator $S^*(t, \hbar) \in H_{\mathbb{T}}^*(\mathbb{P}^m)[[\hbar^{-1}, t, e^t]]$ defined by Eq. (29.18). By Eq. (29.19), the equality $\overline{Z}_i(e^t, \hbar) = Z_i^*(e^t, \hbar)$, and Exercise 27.1.5 (c), we conclude $\overline{S}(t, \hbar) = S^*(t, \hbar)$.

After passing from equivariant to standard cohomology (that is, setting $\alpha_i = 0$) and setting $\hbar = 1$, we obtain case (iii) of the Mirror conjecture. The series $S_X^*(t, \hbar = 1), S_X(t, \hbar = 1) \in H^*(\mathbb{P}^m)[t][[e^t]]$ are determined by

$$\begin{aligned} S_X^*(t, \hbar = 1) &= \frac{1}{(m+1)H} S_{\mathbb{T}}^*(t, \hbar)|_{\alpha_i = 0, \hbar = 1} \\ &= \sum_{i=0}^{m-1} I_i(t) H^i, \\ S_X(t, \hbar = 1) &= \frac{1}{(m+1)H} S_{\mathbb{T}}(T, \hbar)|_{\alpha_i = 0,\ \hbar = 1}, \end{aligned}$$

where $I_i(t) \in \mathbb{Q}[t][[e^t]]$. The following equalities hold:

$$I_0(t) = F(e^t), \quad I_1(t)/I_0(t) = t + \frac{(m+1)(G_{m+1}(e^t) - G_1(e^t))}{F(e^t)}.$$

We have shown $S_X(t, \hbar = 1)$ is obtained from $\sum_{i=0}^{m-1} I_i(t)/I_0(t)$ by the change of variables $T = I_1(t)/I_0(t)$. The proof of this explicit transformation between $S_X^*(t, \hbar = 1)$ and $S_X(t, \hbar = 1)$ completes case (iii) of the Mirror conjecture.

Part 5

Advanced Topics

CHAPTER 31

Topological Strings

Our discussion of two-dimensional QFTs so far has assumed a fixed worldsheet geometry, i.e., a fixed Riemann surface. It is natural to ask whether there is a modification of the theory that allows us to integrate over worldsheet geometries. This idea is the starting point of string theory. In particular, maps from the worldsheet geometries to target space are interpreted as "Feynman diagrams" for string theory. In this context, integrating over the shapes of Feynman diagrams is the same as integrating over the complex structure of the Riemann surface (i.e., metric variation up to conformal equivalence). Integrating over metrics, i.e., including the metric tensor as one of the fields one integrates over, is what is referred to as "quantum gravity." Thus string theory perturbation can be viewed as studying certain quantum gravity theories on the worldsheet. There are some distinct types of string theories depending on which precise quantum gravity one considers on the worldsheet (in particular distinguished by the number of supersymmetries on the worldsheet). We will not discuss this vast topic here. Instead we will concentrate on one such class of string theory, known as topological strings (which can be viewed as a special type of "bosonic string").

31.1. Quantum Field Theory of Topological Strings

In topological strings one would like to couple the topological sigma models to worldsheet gravity; in other words we would like to define what it means to integrate over worldsheet geometries in the context of topological sigma models.

In Ch. 16 we discussed the twisting of $\mathcal{N} = 2$ supersymmetric two-dimensional theories to get topological field theories. We found that there were, up to conjugation, two inequivalent twistings — the A-twist and the B-twist — depending, respectively, on a conserved vector or axial $U(1)$ charge. Landau–Ginzburg (LG) theories could be B-twisted and non-linear sigma models (NLSMs) on Kähler target spaces could be A-twisted. Calabi–Yau sigma models preserve both $U(1)$ charges and admit both kinds of twist. Here we will discuss the A-twisted sigma model with Kähler target M and we will see that the case where M is a Calabi–Yau threefold is special. The case of the B-twisted theory with a Calabi–Yau target or an LG theory can be similarly defined, where again dimension 3 and quasi-homogeneity of the LG superpotential make it more special.

Recall that the A-twisting amounts to the replacement of the worldsheet holonomy (corresponding to the canonical $U(1)_E$ bundle) with $U(1)_E \to \mathrm{diag}(U(1)_E \times U(1)_V)$. The supercharge $Q = Q_- + \overline{Q}_+$ becomes a nilpotent scalar (spin 0) supercharge under this twisting. The (a, c) ring elements are the physical operators (Q-cohomology representatives) of this topological field theory and can be identified with differential forms on M via

$$\begin{aligned}(31.1)\quad \omega &= \omega_{i_1\cdots i_p\bar{j}_1\cdots\bar{j}_q}(z,\bar{z})dz^{i_1}\cdots d\bar{z}^{\bar{j}_q}\\ &\to \mathcal{O}_\omega = \omega_{i_1\cdots i_p\bar{j}_1\cdots\bar{j}_q}(z,\bar{z})\chi^{i_1}\cdots\chi^{i_p}\chi^{\bar{j}_1}\cdots\chi^{\bar{j}_q}.\end{aligned}$$

The Q-cohomology is realized, as discussed before, as the de Rham cohomology of M, and the path-integral localizes to holomorphic maps ($\phi : \Sigma \to M$ such that $\overline{\partial}\phi = 0$).

Counting of the fermion zero modes led to a selection rule for correlation functions in the topological theory — we found that

$$\begin{aligned}(31.2)\quad \#(\chi^i \text{ zero modes})) - \#(\chi^{\bar{i}} \text{ zero modes}))\\ = \int_\Sigma \phi^*(c_1(M)) + \dim_{\mathbb{C}}(M)(1-g)\end{aligned}$$

As discussed in Ch. 10, this is the index of the R-twisted Dolbeault complex on Σ computed by the Hirzebruch–Riemann–Roch theorem:

$$\begin{aligned}\dim H^0(\phi^*T_M) - \dim H^1(\phi^*T_M) &= \int_\Sigma \mathrm{ch}\,(\phi^*T_M)\cdot \mathrm{td}\,(T_\Sigma)\\ &= (1-g)\dim_C(M) - \deg(\phi^*K_M)\end{aligned}$$

and this is the (virtual) complex dimension of the moduli space $\mathcal{M}$ of maps in this class. The fermion zero modes give rise to a net violation of axial fermion number which gives the selection rule:

(31.3)
$$\langle\mathcal{O}_1\cdots\mathcal{O}_s\rangle = 0 \quad \text{unless} \quad \sum_{i=1}^s p_i = \sum_{i=1}^s q_i = c_1(M)[\beta] + \dim_{\mathbb{C}}(M)(1-g)$$

where $\beta \in H_2(M, \mathbb{Z})$ is the homology class of the image of Σ, and p_i, q_i are the holomorphic and anti-holomorphic degrees of the differential form on M corresponding to the operator $\mathcal{O}_i$.

Notice that if $c_1(M) > 0$, then a given correlator at fixed genus g gets contributions only from a finite set of classes of maps. For this reason the calculation is more trivial for manifolds with $c_1 > 0$. We already saw an example of this in the context of the sigma model on $\mathbb{CP}^1$, in Sec. 15.1.1. The case where $c_1 = 0$ is much more interesting because arbitrary classes of holomorphic maps can contribute to a given correlator at a fixed genus: the quantum cohomology ring can get contributions from holomorphic maps of any degree. It is also more natural to consider the case where $c_1 = 0$, because in this case the two-dimensional quantum field theory is believed to flow to a conformal theory, i.e., one which only depends on the complex structure of the Riemann surface. From this point on we will focus on this case, although similar aspects can be also developed for the case of Kähler manifolds with $c_1 > 0$.

Let us focus on the Calabi–Yau case: $c_1(M) = 0$. Let us first ignore coupling to gravity and consider just the ordinary topological sigma model we have already discussed. For this case, the vector symmetry is conserved, but we have seen in Ch. 16 that the axial fermion number gets violated by an amount equal to $2\dim_{\mathbb{C}}(M)(g-1)$. When $g = 0$, the selection rule tells us that the non-vanishing correlators are $\langle\prod \mathcal{O}_{(p_i,q_i)}\rangle$ where $\sum_i p_i = \sum_i q_i = \dim_{\mathbb{C}}(M)$. Geometrically, this corresponds to the fact that the integral $\int_M \omega_{(p_1,q_1)} \wedge \cdots \wedge \omega_{(p_r,q_r)}$ makes sense when the form degrees add up to the dimension of the manifold, and this classical fact remains true in the quantum theory for the Calabi–Yau case. In fact the path-integral reduces to precisely this computation when one considers constant maps (degree 0 holomorphic maps). But there are also contributions from higher-degree maps. The nonzero correlators at genus 0 on a Calabi–Yau threefold are of the form $\langle\mathcal{O}^i_{(1,1)}\mathcal{O}^j_{(1,1)}\mathcal{O}^k_{(1,1)}\rangle$. The degree 0 contribution is the classical intersection number of D_i, D_j, D_k (the divisors Poincaré dual to the forms $\omega^i, \omega^j, \omega^k$) — i.e., constant maps to the intersection points. The higher-degree contributions count maps of a $\mathbb{CP}^1$ into M whose image intersects the divisors D_i, D_j, D_k. In this way the classical cohomology ring gets deformed into a quantum cohomology ring, as discussed before.

Are there any other correlators that have a chance of being nonzero in genus 0? Recall the topological descendants that were discussed in Ch. 16; given a topological observable $\mathcal{O}$, one could consider the corresponding two-form observable $\mathcal{O}^{(2)}$ integrated over the worldsheet: $\int_\Sigma \mathcal{O}^{(2)}$. This operator has axial charge 0, and so arbitrary numbers of these can be inserted in the path-integral without changing the axial charge (consistent with the selection rule). Notice that these operators are precisely those that can be used to deform the action: $\delta S = \sum_i t_i \int_\Sigma \mathcal{O}_i^{(2)}$. An insertion of $\int_\Sigma \mathcal{O}_i^{(2)}$ therefore corresponds exactly to varying the original correlator with respect to t_i (the corresponding Kähler parameter):

$$\frac{\delta}{\delta t_i}\langle \mathcal{O}^j \mathcal{O}^k \mathcal{O}^l\rangle = \langle \mathcal{O}^j \mathcal{O}^k \mathcal{O}^l \int_\Sigma \mathcal{O}^{i(2)}\rangle. \tag{31.4}$$

Hence these correlators carry no extra information (beyond that contained in the $c_{ijk}(t)$). By conformal invariance it does not matter whether we fix the three points corresponding to operators j, k, l and integrate over the one corresponding to i or exchange i with any of the other three operators. This simple statement implies that

$$\partial_l C_{ijk}(t) = \partial_i C_{ljk}(t). \tag{31.5}$$

This equation is known as the WDVV (Witten–Dijkgraaf–Verlinde–Verlinde) equation. Together with the symmetry of C_{ijk} under permutation of its labels, it follows that we can define a function (called the genus 0 free energy) $F_0(t)$ with the property that

$$C_{ijk}(t) = \partial_i \partial_j \partial_k F_0(t). \tag{31.6}$$

Actually this presupposes a particular choice of coordinates (known as topological flat coordinates). In a more invariant form one has to replace the above derivatives with covariant derivatives. We will discuss this further below. F_0 thus defined is known as the genus 0 topological string partition function. Note that for three-point functions, given the rigidity of $\mathbb{P}^1$ with three punctures, there is no integration over moduli of a Riemann surface to do. This is consistent with the notion one wants of a topological string as integrating over all worldsheet geometries for this case.

EXERCISE 31.1.1. *Show that the n-point functions of topological strings can be defined as nth partial derivatives of F_0 and that the integration is automatically over the correct moduli space of $\mathbb{P}^1$ with n punctures.*

For genus $g = 1$, the axial anomaly vanishes, and in this case no operators need to be inserted in the path-integral to keep it from vanishing and the path-integral computes the partition function, which is the Euler characteristic of M. However one would also want, not a fixed genus 1 curve, but a sum over all possible complex structures. This has to be incorporated in some way. Similarly, when $g > 1$, the axial anomaly is equal to $6(g-1) > 0$, so no insertions of topological observables (which have positive axial charge) can absorb the required number of fermion zero modes and all correlators vanish — the higher genus correlators are all trivial. Mathematically, this corresponds to the fact that for a fixed Riemann surface, there are generically no holomorphic higher genus maps to M. However, we could consider allowing the metric (complex structure) on Σ to vary and integrate over all metrics – physically speaking, this corresponds to coupling the topological sigma model to topological gravity. We will see below how to define this theory for $g \geq 1$, extending the definition for $g = 0$.

We will now discuss why the case of the Calabi–Yau *three*fold is especially nice for coupling to topological gravity. The complex dimension of the moduli space of maps for a fixed Riemann surface ($g \geq 2$) in these cases is formally negative: $\dim_{\mathbb{C}} \mathcal{M} = (\dim M)(1-g) = 3(1-g) < 0$. On the other hand, we know that the dimension of the moduli space of metrics up to conformal transformations on a genus $g \geq 2$ surface has dimension $3(g-1)$. This suggests that if we consider integrating over the moduli space of Riemann surfaces we should "formally" have isolated points where there are holomorphic maps. In other words the formal dimension of this moduli space problem, where *we do not fix the complex structure of the worldsheet*, is zero. So in this case allowing integration over the complex structure of the Riemann surface exactly soaks up the positive violation of the axial charge. Three is the critical dimension of Calabi–Yau for this to happen, because the complex dimension of the moduli space of Riemann surfaces is $3(g-1)$. This is why the case of the Calabi–Yau threefold is so special.

In order to define topological string theory for Calabi–Yau threefolds we need to recall a few ingredients. We saw in Ch. 16 that for the twisted theory, we have

$$T_{\mu\nu} = \{Q, G_{\mu\nu}\} \tag{31.7}$$

where T is the stress-energy tensor, defined as

$$T_{\mu\nu} = \frac{1}{\sqrt{g}}\delta S/\delta g^{\mu\nu}. \tag{31.8}$$

Since we started with a conformal theory (Calabi–Yau sigma model), the trace of the stress tensor $T^\mu_\mu = 0$ (i.e., action is invariant under changes in the metric by rescaling). Therefore the only nonzero components of T are $T_{zz} := T_{++}$ and $T_{\overline{z}\overline{z}} := T_{--}$. So we see that

$$T_{++}(z,\overline{z}) = \{Q, G_{++}(z,\overline{z})\}, \qquad T_{--}(z,\overline{z}) = \{Q, G_{--}(z,\overline{z})\}, \tag{31.9}$$

where G_{++}, G_{--} are the currents that correspond to the charges $Q_+, \overline{Q}_-$. Since T has axial charge 0 and Q has axial charge 1, the G's have axial charge -1 (in the left and right sectors, respectively). We wish to use the G's to define a measure on the moduli space $\mathcal{M}_g$ of Riemann surfaces of genus g. In other words, for a given *point* $\Sigma \in \mathcal{M}_g$ and a choice of $6g-6$ tangent vectors we wish to get a number (with the appropriate multilinearity property). The tangent to $\mathcal{M}_g$ at a point Σ corresponds to a choice of Beltrami differential on the Riemann surface Σ. That is, $T\mathcal{M}_g|_\Sigma = H^1(T\Sigma)$, which we think of as holomorphic vector-valued $(0,1)$-forms. Let μ_i denote $3g-3$ such Beltrami differentials which span the complex tangent space to $\mathcal{M}_g$ at the point Σ.

The measure on $\mathcal{M}_g$ is defined by

$$\left\langle \prod_{i=1}^{3g-3} G_{++}(\mu_i) \prod_{i=1}^{3g-3} G_{--}(\overline{\mu}_{\overline{i}}) \right\rangle \tag{31.10}$$

where

$$G_{++}(\mu_i) := \int G_{zz}\mu_{\overline{z}}^{z} d^2 z \tag{31.11}$$

with a similar definition for $G_{--}(\overline{\mu}_{\overline{i}})$. Here μ_i, $i = 1, \ldots, 3g-3$, are the Beltrami differentials. Since the G's each have axial charge -1, the product has charge $(3-3g, 3-3g)$ which cancels the axial charge anomaly, hence the measure is, a priori, nonzero. So the genus g topological string amplitude (for $g > 1$) is defined by

$$F_g = \int_{\mathcal{M}_g} \prod_{i=1}^{3g-3} dm_i d\overline{m}_{\overline{i}} \langle \prod_{i=1}^{3g-3} G_{++}(\mu_i) \prod_{i=1}^{3g-3} G_{--}(\overline{\mu}_{\overline{i}})\rangle, \tag{31.12}$$

where dm_i are the dual one-forms to the μ_i. For the case of $g = 1$ we define F_1 by considering the one-point function. This is related to the fact that an elliptic curve has isometries and we thus have to fix a point on it to make it rigid. So if we insert one observable in the Kähler class of axial charge (1,1) at a point, this is exactly what will cancel the $(-1,-1)$ axial charge from the insertion of the pair (G^{++}, G^{--}) (corresponding to having a one-dimensional complex moduli space of elliptic curves). In other words we define

$$\partial_i F_1 = \int_{\mathcal{M}_{1,1}} dm d\overline{m} \langle G_{++}(\mu) G_{--}(\overline{\mu}) \mathcal{O}_i \rangle. \tag{31.13}$$

It is not difficult to show that this can be integrated in the form

$$F_1 = \frac{1}{2}\int \frac{d^2\tau}{\tau_2} \mathrm{Tr}(-1)^F F_L F_R q^{H_L} \overline{q}^{H_R} \tag{31.14}$$

where the integral is over the fundamental domain of the moduli space of elliptic curves and τ is the standard parametrization of the Teichmüller space. Here $q = \exp(2\pi i \tau)$. This definition is well defined up to addition of a constant to F_1.

EXERCISE 31.1.2. *Derive the above result for F_1. Hint: you will need the explicit realization of the F_L and F_R currents, and the equivalence of the operator and path-integral representations.*

An important question to ask is whether this definition of topological string amplitudes is consistent with the topological symmetry, Q? In other words we would like to demonstrate that the correlation functions involving Q-trivial fields vanish. So consider

$$\langle Q\Lambda \rangle \tag{31.15}$$

where by $\langle \cdots \rangle$ we mean the measure defined in Eq. (31.12), in addition to the insertion of $Q\Lambda$. Were it not for the insertion of G's in the correlation functions, this would have vanished by the Q-symmetry of the path-integral. Thus the only potential non-vanishing terms will come from the Q-variation of G_{++} and G_{--} insertions in Eq. (31.12). But according to Eq. (31.9), these variations give T_{++} and T_{--} folded in with the Beltrami differentials. However, from the definition of the energy-momentum tensor, Eq. (31.8), this is simply the derivative of the action along the variation of the corresponding moduli of the Riemann surface. Thus we obtain in this way a total derivative of a lower degree form on the moduli space of Riemann surfaces, $\mathcal{M}_g$, and so barring contributions from boundaries of moduli space this gives zero. As we will discuss later, for most choices of Λ there are no contributions from infinity and this shows that topological symmetry is respected by this extension of the theory to include integration over moduli of Riemann surfaces. But for some choices of Λ there are contributions from the boundaries of $\mathcal{M}_g$ to Q-trivial fields. These will be topological anomalies (also known as holomorphic anomalies), and the appreciation of their structure is crucial to the computation of higher-genus topological string amplitudes.[1] We shall discuss these anomalies in the next section.

There is one important point we have to take into account. F_g is *not to be viewed as a function on the moduli space of Calabi–Yau manifolds, but as a section of a bundle over it.* Let $\mathcal{L}$ denote the line bundle corresponding to the lowest dimension (a,c) state (or (c,c) field in the case of B-twisted topological theory). In the case of A-model twisting this state corresponds to the cohomology element $H^0(M)$ (and in the B-model case it corresponds to the holomorphic three-form). We will now argue that

$$F_g \in \Gamma(\mathcal{L}^{2g-2}) \qquad (31.16)$$

i.e., F_g is a section of the line bundle $\mathcal{L}^{2g-2}$. From the path-integral definition of F_g, this arises because of choices in the Grassmann integration over the fermionic modes due to the axial violation of charge in the path-integral. For example, as discussed in the context of the B-model twisting in Ch. 16, F_0 is a section of $\mathcal{L}^{-2}$, where $\mathcal{L}$ is the line bundle on the moduli space of complex structures of a Riemann surface corresponding to a holomorphic three-form on the Calabi–Yau. This is reflected in the fact that

$$C_{ijk} = \int_X \Omega \wedge \partial_i \partial_j \partial_k \Omega \qquad (31.17)$$

(see Eq. (31.6)), where X denotes the threefold and Ω is a choice for the holomorphic three-form on the Calabi–Yau. In other words, if we rescale Ω by a factor λ, then C_{ijk} is rescaled by a factor of λ^2. Thus F_0 takes values

[1] The definition of topological string is modeled after the definition of bosonic strings, and it should be viewed as providing simple examples of bosonic strings. In fact in some cases the relation is clearer. For example, the non-critical bosonic string on a target circle at the self-dual radius is equivalent to the B-model topological string at the conifold. The identification between fields in bosonic strings and those in the topological strings is as follows: $Q \equiv Q_{BRST}, T \equiv T, G_{++} \equiv b, G_{--} \equiv \bar{b}$, and ghost number $\equiv$ fermion number. The ghost number violation by $(3g-3)$ in the bosonic string on a genus g surface mirrors the axial charge violation by the same amount in the topological string.

in the dual bundle to $\mathcal{L}^2$. This result, of course, is related to the fermion number violation of genus 0 which in the B-model is $(-3,3)$; now Ω, as a state, has charge $(3/2,-3/2)$, so putting two insertions of it will soak up the zero modes. This is also related to the fact that on the hemisphere (as discussed in Ch. 17 in the context of tt^* equations), the topologically twisted theory gives the lowest charge ground state. The result for higher genus follows by similar reasoning, using the axial charge violation. Again it is most easily understood in the context of B-model topological twisting. Namely at genus g we have left-right fermion number violation of $(3g-3,-3g+3)$. Ω has charge $(3/2,-3/2)$ and so $2g-2$ insertions of it neutralizes it. Thus F_g is a section of $\mathcal{L}^{2g-2}$.

31.2. Holomorphic Anomaly

In this section we would like to discuss how the decoupling of some of the topologically trivial terms fails in the context of topological strings. Recall that for Calabi–Yau threefolds there are two, up to complex conjugation, topological twistings, A and B. Let us also denote their complex conjugate twisting by $\overline{\mathrm{A}}$ and $\overline{\mathrm{B}}$. Topological symmetry suggests that if we are considering say the A topological theory, then the $\overline{\mathrm{A}}, \mathrm{B}, \overline{\mathrm{B}}$ observables are trivial (and similarly the D-terms). Here we will discuss how this fails for the $\overline{\mathrm{A}}$ observables, but continues to be true for the B and $\overline{\mathrm{B}}$ observables. Similarly if we consider the B topological theory, the partition function will depend on the $\overline{\mathrm{B}}$ observable, but not on A or $\overline{\mathrm{A}}$.

To be concrete, let us focus on the A model. Let us recall the argument that led to localization of the path-integral to holomorphic maps: the sigma model action is given by

$$S \sim \int d^2z [t g_{i\bar{j}} \partial\phi^i \bar{\partial}\phi^{\bar{j}} + \bar{t} g_{i\bar{j}} \bar{\partial}\phi^i \partial\phi^{\bar{j}} + \text{fermionic terms}]. \qquad (31.18)$$

Here t is a complex variable that parametrizes the complexified Kähler class. In general there are many Kähler classes parametrized by complex variables t_i, and here we have exhibited only one of them explicitly. Since $\bar{t}$ variations are Q-trivial, i.e., belong to $\overline{\mathrm{A}}$ observables, if we have Q-invariance we can take $\bar{t} \to \infty$ without affecting the computation of the partition function. In this limit the path-integral receives contributions only from the holomorphic maps. In other words, we consider the asymmetric limit of fixing t but taking

$\bar{t} \to \infty$. We have

$$F_g(t,\bar{t}) = \lim_{\bar{t}\to\infty} F_g(t,\bar{t}) = \sum_d F_{g,d} e^{-dt}. \tag{31.19}$$

Here $F_{g,d}$ define the Gromov–Witten invariants in terms of the genus g partition function of the topological string.

However, we will see that the holomorphic anomaly (or the topological anomaly) implies that $F_g(t,\bar{t})$ is not quite independent of $\bar{t}$. (Even though F_g depends on $\bar{t}$ it still does not depend on the complex structure deformation parameters.)

We now discuss how the $\bar{t}$ dependence arises. We will be rather brief here, and just discuss the main features that arise. We will first summarize the results. The genus 0 correlators have no holomorphic anomaly (i.e., the quantum cohomology ring and its B-model version are purely holomorphic). For genus 1 we have

$$\bar{\partial}_i(\partial_j F_1) = \frac{1}{2}\mathrm{tr}\, C_j\overline{C_i} - \frac{\mathrm{Tr}(-1)^F}{24} G_{j\bar{i}} \tag{31.20}$$

where C_i and $\overline{C}_j$ were defined in Eq. 17.15 of Ch. 17, and correspond to the action of chiral and anti-chiral fields on the ground states. $G_{i\bar{j}}$ denotes the metric on the moduli space of Calabi–Yau manifolds (see Ch. 17), which we discussed in the context of the tt^* equations. In general, for genus $g > 2$,

$$\bar{\partial}_{\bar{i}} F_g(t_j, t_{\bar{j}}) = \frac{1}{2} C_{\bar{i}}^{jk} \left(\sum_{r=1}^{g-1} D_j F_{g-r}\, D_k F_r \,+\, D_j D_k F_{g-1} \right) \tag{31.21}$$

The covariant derivatives on the RHS reflect the fact that the free energies F_g are not numbers — they are sections of bundles ($F_g \in \mathcal{L}^{2g-2}$), as well as the fact that there is a metric connection on the moduli space (in taking the second derivative). The fact that the covariant derivative does not appear on the LHS reflects the fact that the topological theory chooses a holomorphic section of the bundle so that the anti-holomorphic part of the connection is zero, as was discussed in our derivation of the tt^* equations. Also

$$C_{\bar{i}}^{jk} \equiv C_{\overline{ijk}} g^{j\bar{j}} g^{k\bar{k}},$$

where g denotes the metric on the ground states and is related to $G_{i\bar{j}} = e^K g_{i\bar{j}}$ where $g_{0\bar{0}} = e^{-K}$.

The general structure of contributions to the anomaly can be understood in the following way for $g \geq 2$. We argued that the holomorphic anomaly

would come from the boundaries of the moduli space $\mathcal{M}_g$. There are two classes of terms on the right-hand side. The first class of terms on the right-hand side reflects the contribution coming from the boundary component of the moduli space corresponding to the degeneration of a genus g curve to a curve of genus r and another of genus $(g-r)$ with one added puncture on each curve, connected by a tube with an *anti-chiral* field $\overline{\phi}_{\bar{i}}$ inserted. See Fig. 1.

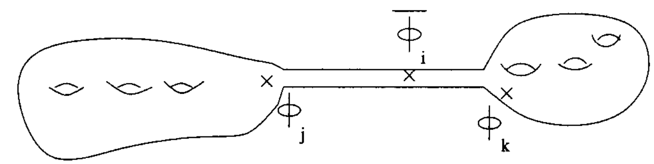


FIGURE 1. Degeneration of a genus g curve to two disconnected curves

The second term comes from the boundary component corresponding to curves of genus $(g-1)$ with two extra punctures, again with the $\overline{\phi}_{\bar{i}}$ inserted on a tube connecting the two punctures. See Fig. 2.

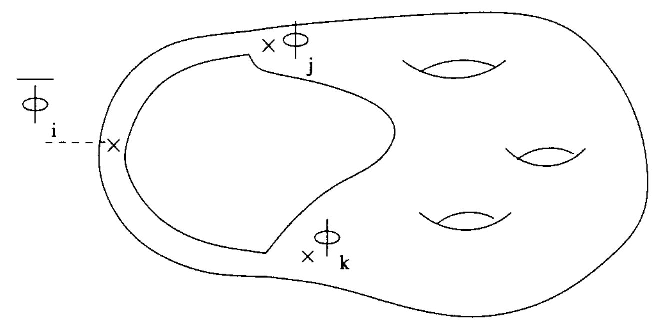


FIGURE 2. Degeneration of a genus g curve to a genus $(g-1)$ curve

The arguments leading to this result are very much in the spirit of the derivation of the tt^* equations, which we discussed in detail in Sec. 17.1, and so here we will be brief in the presentation of the proof. Differentiating $\bar{\partial}_i F_g$ is equivalent to the additional insertion of

$$\int_\Sigma Q_- \overline{Q}_+ \overline{\phi}_i$$

in the correlation function defining the topological string. Using both Q_- and $\overline{Q}_+$ and the discussion after Eq. (31.15) we obtain

$$\overline{\partial}_i F_g = \int_{\mathcal{M}_g} \partial\overline{\partial}\omega \tag{31.22}$$

where ω is a $(3g-4, 3g-4)$-form on $\mathcal{M}_g$. ω is defined through the correlation function involving the insertion of $3g-4$ pairs of G_{++} and G_{--} (folded in with the corresponding Beltrami differentials) and in addition the insertion of the integral of $\int \overline{\phi}_i$ on the Riemann surface. Integrating the total derivative in Eq. (31.22), we are led to consider the boundaries of $\mathcal{M}_g$ with the further derivative in the normal direction to this boundary, i.e., we have

$$\int_{\mathcal{M}_g} \partial\overline{\partial}\omega = \int_{\partial\mathcal{M}_g} \partial_n\omega. \tag{31.23}$$

As discussed before there are two kinds of boundaries, each involving a long tube. The normal derivative in Eq. (31.23) corresponds to changing the length T of the tube. There are two possibilities for where the $\int \overline{\phi}_i$ is placed relative to the tube: on or off it. The parts where the integral is off the tube give zero, because the infinitely long tube will project to ground states, and then $\partial_T \exp(-TH) \to 0$. The part corresponding to $\int \overline{\phi}_i$ being inserted on the tube gives a factor of T corresponding to integration of $\overline{\phi}_i$ and the ∂_T simply gets rid of it. So we end up with $\overline{\phi}_i$ (integrated over a circle) on an infinite tube. The infinite tube projects the states to the ground states. The insertion of $\overline{\phi}_i$ on these states is represented by the matrix $\overline{C}_i$ acting on the ground states as in our discussion of the derivation of the tt^* equations. It is not difficult to see that the only relevant ground states on either side of the tube will correspond to charge $(1,1)$ states (other states are annihilated by the action of G's on either end of the tube). Let us label the two states by j and k on either end of the tube. The two points of the connection of the tube to the rest of the Riemann surface are integrated over and thus correspond to the insertion of $\int Q_+\overline{Q}_-\phi_i$ and $\int Q_+\overline{Q}_-\phi_j$ on the two ends. This is equivalent to taking derivatives of the corresponding partition functions with respect to t_i and t_j. This explains the above anomaly formula for $g > 1$, Eq. (31.21) (the factor of $1/2$ comes from the symmetry factor in this decomposition and avoids a double counting). It is not too difficult to extend this to the observables of the other two topological theories and show that they would not lead to any anomalies.

EXERCISE 31.2.1. *Demonstrate the above statement.*

The derivation of the anomaly equation for genus 1, Eq. (31.20), follows a similar reasoning. The only additional subtlety comes from the fact that we have already one field inserted (the chiral field corresponding to ϕ_j) so there is an additional boundary corresponding to when $\overline{\phi}_i$ is near ϕ_j and this gives rise to the second term in Eq. (31.20) involving $G_{j\bar{i}}$.

As we see, the anomaly is associated with integration over moduli of Riemann surfaces, which is why the genus 0 theory, being rigid, does not suffer from holomorphic anomalies.[2]

Taking into account the holomorphic anomaly turns out to be essential in computing F_g for $g > 0$, especially in the context of B-model topological twisting. The basic idea is that if F_g had no anomalies it would have been a holomorphic section, and using the global geometry of the moduli space of the Calabi–Yau and physical insight into the behaviour of the partition function near the boundaries of the moduli space of the Calabi–Yau, one could write down F_g, up to a finite number of undetermined constants. F_g is not holomorphic, but the anomaly equation tells exactly how it fails to be holomorphic. So one first constructs *any* F_g' that satisfies the holomorphic anomaly equation. Then $F_g - F_g'$ is a purely holomorphic section which can be determined up to a finite number of constants parametrizing the finite-dimensional space of holomorphic sections with fixed order poles at various singularities. This will be discussed in more detail in Ch. 35. We will also discuss there the connection between the genus 1 holomorphic anomaly Eq. (31.20) in the context of the B-model topological twisting and the holomorphic anomaly corresponding to the curvature of the determinant line bundle of certain operators on the Calabi–Yau, known as the Quillen anomaly.

[2]If a sufficient number of fields are inserted on the Riemann surface then there is an anomaly even at genus 0 coming from the boundary of the moduli space of spheres with many punctures. In this case, the anomaly is equivalent to the relation between properties of the metric on the moduli space and the the chiral ring matrices.

CHAPTER 32

Topological Strings and Target Space Physics

In this chapter we will discuss a reinterpretation of topological string amplitudes from the target space viewpoint. Thus far we have been talking about two-dimensional field theories on Riemann surfaces as probes of target space geometry. In the context of string theory, one is interested in the spectrum of particles and their interactions in the target space. In other words, worldsheets of strings correspond to Feynman diagrams of a theory in target space. In order to introduce the target space viewpoint, we need a little background in string theory.

32.1. Aspects of Target Space Physics

There are two kinds of superstring theories of closed strings with left-moving and right-moving supersymmetry on the worldsheet: Type IIA and Type IIB (the nomenclature A,B is not unrelated to that of the A and B models that we have talked about on the worldsheet). Consistency of these theories[1] imposes the condition that the target space should be ten-dimensional. One is then led to consider maps $\Sigma \to M^{(10)}$ of Riemann surfaces to some ten-dimensional space-time.

In "first-quantized" string theory, there is a perturbative expansion, in terms of genus of the worldsheet, that parallels the loop expansion of Feynman diagrams in field theory. However, as we know, there is a richer structure to QFT than is seen at the level of the perturbation expansions (e.g., instantons), and the same goes for string theory. One viewpoint on perturbative QFT (parallel to the worldsheet point of view) is to view it as the quantum mechanics system of maps $\mathbb{R} \to M$, where $\mathbb{R}$ represents the wordline of a particle. As we know, the target space interpretation of this is as a QFT on M. In string theory, there is a similar interpretation — a

[1] "Consistency" can mean worldsheet conformal invariance, which allows us to integrate over worldsheet metrics with a finite-dimensional moduli space integral, or it can mean ten-dimensional Poincaré invariance.

fixed time picture of it looks like a map from a circle, known as the string, to a spatial slice of M. Thus the study of the worldsheet theory induces a QFT in the ten-dimensional space-time. The states in the Hilbert space of the two-dimensional theory will correspond to particle states from the viewpoint of the target space. There are infinitely many such states, corresponding to particles of increasing mass, as seen from the target space. One of these states is a massless spin 2 particle in space-time, interpreted as the graviton describing the fluctuations of the metric in the target space — this was one of the original motivations for studying string theory: the fact that it seemed to produce gravity in the target space automatically.

In a field theory, one usually starts with a "vacuum configuration," one that solves the equations of motion at the classical level. One of the equations of motion in the target space is the Ricci-flatness condition on the metric, a condition obeyed by Calabi–Yau manifolds. As we saw in Sec. 14.2.4, Ricci flatness is also the condition of conformality of the worldsheet sigma model at one loop. Conformality of the worldsheet theory thus corresponds to the classical equations of motion in space-time being satisfied. Since we are interested in Ricci-flat target spaces as classical string vacua, one starting point is $\mathbb{R}^{9,1}$, ten-dimensional Minkowski space. This theory in target space also enjoys supersymmetry with 32 supercharges.

A physically more interesting starting point is $\mathbb{R}^{3,1} \times X^{(6)}$, the product of a four-dimensional Minkowski space with a six-dimensional compact ("internal") manifold, which we take to be a Calabi–Yau threefold.[2] Calabi–Yau threefolds, in addition to enabling us to use the classical equations, allow the preservation of some supersymmetry in four dimensions. If the threefold has trivial holonomy, i.e., a six-dimensional flat torus, then the spinor bundle is trivial and all 32 supercharges are preserved: this corresponds to $N = 8$ supersymmetry in four dimensions. If the Calabi–Yau threefold has an $SU(2)$ holonomy, for example $X^{(6)} = K3 \times T^2$, half the supersymmetries are preserved and we get an $\mathcal{N} = 4$ theory in four dimensions. If we consider a generic Calabi–Yau threefold with $SU(3)$ holonomy, only a quarter of the supersymmetries are preserved, and that corresponds to $\mathcal{N} = 2$ supersymmetric theories in four dimensions (with eight supercharges).

When $X = X^{(6)}$ is sufficiently small, the space-time looks macroscopically four-dimensional. One is typically interested in the spectrum of light

[2]"Three" refers to the complex dimension.

particles and their interactions in four dimensions, and all this data depends on the data of the Calabi-Yau. The effective theory we obtain in four dimensions is called the Kaluza–Klein reduction of the ten-dimensional theory on the Calabi–Yau manifold. Typically the light modes in four dimensions are related to zero modes of an appropriate Laplace operator acting on forms on the Calabi–Yau. For example, consider a massless scalar field Φ in ten dimensions satisfying Laplace's equation: $\Delta_{(10)}\Phi(x^\mu, y^i) = 0$, where x^μ are coordinates on four-dimensional space-time and y^i are coordinates on the internal space. Decomposing the Laplacian operator into a four-dimensional piece and a six-dimensional piece, with the ansatz $\Phi(x, y) = \phi(x)\psi(y)$, we see that $(\Delta_{(4)} + m^2)\phi(x) = 0$ if ψ is an eigenfunction of the Laplacian on the internal space with eigenvalue m^2 — i.e., from the four-dimensional viewpoint we have a particle with mass m. Since the eigenvalues of the Laplacian on the internal space scale with its size as $1/R^2$, where R is the "radius" of the Calabi–Yau, we see that, if the internal space is very small, most of the modes living on it will be very massive and we will end up with a field theory in four dimensions, by dimensional reduction, with some finite number of massless or light degrees of freedom (i.e., the very massive modes can be "integrated out"). Massless degrees of freedom in the low-energy theory will therefore correspond to zero modes of the appropriate differential operator (the Laplacian, for a scalar field) on the internal space. For example, variations of the metric of the Calabi–Yau (the complex and Kähler deformations), which are zero modes of certain operators, correspond to massless fields in four dimensions. In other words, if t parametrizes the moduli space of the Calabi–Yau, then $t(x)$ is a four-dimensional field, where x denotes a point in four-dimensional space-time. Geometrically, this corresponds to varying the Calabi–Yau metric over space-time.

32.2. Target Space Interpretation of Topological String Amplitudes

We are now ready to discuss what topological strings compute in connection with target space physics questions. Let us focus on X being a generic Calabi–Yau threefold. If we consider superstring compactification on X, in the left-over four-dimensional space we then obtain an $\mathcal{N} = 2$ theory (with eight supercharges). It turns out that the superspace properties of this four-dimensional theory are analogous to the $\mathcal{N} = 2$ superspace aspects in two dimensions. Namely we have, up to complex conjugation, two different kinds of F-terms, which correspond to integrating over $\int d^4\theta$ and $\int d^4\tilde{\theta}$. Fields that can appear in one F-term cannot appear in the other three. For example, the Kähler moduli of the Calabi–Yau, which correspond to scalars, will appear in one F-term superspace, whereas the complex structure moduli fields will appear in the other. This is surprisingly similar to the situation we found in the two-dimensional case, where the Kähler deformation and complex deformation of the sigma model corresponded to F-terms in the (c, c) or (c, a) superspace. Moreover, generic deformations of the four-dimensional, $\mathcal{N} = 2$ theory will involve integration over $d^8\theta$ and are analogous to the D-terms we discussed in the two-dimensional theory with $\mathcal{N} = 2$.

Since we saw in the context of two-dimensional theories that F-terms are particularly nice objects to compute, it is natural to ask if we can compute the corresponding F-terms in the $\mathcal{N} = 2$ theory in four dimensions. It turns out that this question is related to computations in the topological string theory.

It can be shown that for certain superpotential terms in the four-dimensional effective theory, the string theory integral over the space of all maps from Riemann surfaces to the target space manifold reduces to (or localizes on) the topological string computations. In particular, if $F_g(t_i)$ denotes the genus g topological amplitude, then in the four-dimensional action there is a term generated which looks like

$$\int d^4x d^4\theta \mathcal{W}^{2g} F_g(t_i) = \int d^4x F_g(t_i) R_+^2 F_+^{2g-2} + \cdots . \tag{32.1}$$

Here $\mathcal{W}$ is the superfield multiplet in the four-dimensional space-time that contains as its top component the field strength of the gravitational multiplet (the self-dual part of the curvature) and as its lowest component the graviphoton field strength (the $U(1)$ gauge field in the same supersymmetry multiplet as the graviton). In particular R_+^2 is a contraction of the self-dual part of the Riemann tensor with itself, while $F_+ = F + *F$ is the self-dual part of the field strength of the graviphoton field strength. Let us now concentrate on Type IIA superstrings on a Calabi–Yau. F_g will denote the A-model partition function on X. The t_i parametrize the (complexified) Kähler classes of the Calabi–Yau, and corresponding to each one we get a $U(1)$ vector multiplet in four dimensions whose scalar is $t_i(x)$ (the bosonic field content of a vector multiplet in four dimensions is a gauge field and a

complex scalar). In the Type IIB superstring, we get the same statement for the B-model topological string partition function, where in that case t_i parametrize complex moduli of the Calabi–Yau. Mirror symmetry relates the A-model to the corresponding B-model on the *mirror* Calabi–Yau. In particular, the theory of Type IIA superstrings on a Calabi–Yau is equivalent to type IIB superstrings on the *mirror* Calabi–Yau.

A special case is $g = 0$ where (after we integrate over the superspace in Eq. (32.1)) we get a term in the effective action of the form

$$\int d^4x (\partial_i \partial_j F_0) F_i^+ \wedge F_j^+$$

where F_i^+ is the self-dual part of the field strength of the gauge field in the ith vector multiplet (in the same multiplet as the scalar t_i). F_0 is the genus 0 prepotential and

$$\partial_i \partial_j F_0 \equiv \tau_{ij}$$

is the two-point function in genus 0, which is interpreted as the gauge coupling in four dimensions. So the genus 0 prepotential tells us the gauge coupling of the $U(1)$ gauge fields.

As in two dimensions, there is a twisting that makes the vector multiplet fields of the A-model dynamical while decoupling the other sectors. Since the string coupling constant lies in a hypermultiplet which can thus appear in a different superspace, and since as discussed before the different types of F-terms decouple in the four-dimensional effective theory, we learn that the A-model topological string cannot depend on the string coupling constant and thus actually captures *exact* (non-perturbative in the string coupling constant) information about the effective action in four dimensions.[3]

It is natural to think of giving a vacuum expectation value to the self-dual part of the graviphoton field strength, i.e., setting

$$F_+ = \lambda. \tag{32.2}$$

[3]Studying the four-dimensional topological theory on four-manifolds (Witten's reformulation of Donaldson invariants) can be related to studying string theories propagating on special non-compact Calabi–Yau threefolds, which is an example of "geometric engineering of quantum field theories." (See Ch. 36,) In this context, the geometry of the $U(1)$ gauge coupling constant captured by second derivatives of the prepotential F_0 is a key ingredient.

Then the full (nonperturbatively exact) result for these F-terms is

$$\int d^4x R_+^2 F(t_i) \tag{32.3}$$

where $F(t_i)$ is the full partition function of the topological string theory (notational warning: $F(t_i)$ is *not* a field strength):

$$F(t_i) = \sum_g \lambda^{2g-2} F_g(t_i). \tag{32.4}$$

As we have seen, the topological string suffers in general from a holomorphic anomaly, so that F depends not only on the t_i but also the $\bar{t}_i$. However, we can still expand the expression for $F(t_i, \bar{t}_i)$ around the basepoint $\bar{t}_i \to \infty$. $F(t_i)$ represents the leading term in that expansion.

In summary, topological string amplitudes on $\mathbb{R}^{3,1} \times X$, which are constructed from integrals over Riemann surfaces of varying moduli and all genera, compute certain terms in the effective four-dimensional action. These terms have a particularly simple form (Eq. (32.3)) against the background of a constant self-dual graviphoton field strength.

32.2.1. Target Space Viewpoint of the Generation of F-term. Now we ask the following question: How does target space physics view the generation of the term in Eq. (32.3)? The answer turns out to be that there are certain hidden degrees of freedom (solitons in the form of minimal D2-branes wrapped over two-cycles of the Calabi–Yau) that have been integrated out and have led to the effective action of Eq. (32.3). To explain this, we will need to digress a little and go back to the early days of quantum field theory when Schwinger did a computation of the effect of integrating out a charged scalar field coupled to a constant $U(1)$ field strength. In other words, the question is the following: How is the existence of a charged scalar field reflected in the properties of the $U(1)$ gauge theory alone?

Consider the two-dimensional version of this: a charged scalar field ϕ is coupled to a constant (background) field strength $F_{\mu\nu} = \epsilon_{\mu\nu} F$. The path-integral over the complex field ϕ will depend on F and is naturally interpreted as the effect on the $U(1)$ gauge theory of integrating out a charged field.

$$e^{-S} = \int \mathcal{D}\phi \exp\left(- \int |(\partial_\mu - eA_\mu)\phi|^2 + m^2|\phi|^2 \right). \tag{32.5}$$

As discussed in Sec. 10.1.1, this computation involves a simple Gaussian integral involving the determinant of an operator. We have

$$S = \ln\det[\Delta + m^2] = \mathrm{Tr}\ln(\Delta + m^2) = \int_\epsilon^\infty \frac{ds}{s} \mathrm{Tr}\, e^{-s(\Delta+m^2)}. \tag{32.6}$$

In order to evaluate this integral, we note that

$$\Delta = D_1^2 + D_2^2$$

and

$$[D_1, D_2] = eF.$$

This algebra in the 2d computation is the same algebra as that of the harmonic oscillator we studied in the context of quantum mechanics in Ch. 10. We can thus evaluate the above evolution operator as in Eqs. 10.36 and 10.47 and find

$$S = \int_\epsilon^\infty \frac{ds}{s} \frac{e^{-sm^2}}{2\sin(seF/2)}. \tag{32.7}$$

In the four-dimensional case, we are interested in the self-dual part of F coupled to a charged scalar field. The relevant part of the computation involves the determinant of the four-dimensional Δ operator. The computation splits into two parts, one for each two-dimensional subspace of the four-dimensional space, and gives (where $F_{12} = F_{34} = F$)

$$S = \int_\epsilon^\infty \frac{ds}{s} \frac{e^{-sm^2}}{(2\sin(seF/2))^2}. \tag{32.8}$$

Here we have been considering charged scalar fields coupled to a $U(1)$ gauge field in four dimensions. If, instead, we integrate out charged fields of mass m, transforming in a non-trivial representation of the four-dimensional Lorentz group, the computation above is easily modified, and the main additional term comes from the fact that the relevant Laplacian Δ has an additional term

$$\Delta \to \Delta + 2e\sigma_R^{\mu\nu} F_{\mu\nu},$$

where $\sigma_R^{\mu\nu}$ denotes the Lie-algebra representation of the $SO(4)$ Lorentz group acting on the field which is in the representation R of the Lorentz group. Note that at the Lie-algebra level $SO(4) = SU(2)_L \times SU(2)_R$ and for the the self-dual field strength configuration, where $F_{12} = F_{34}$, only the $SU(2)_L$ content of the representation R will enter the above formula.

We thus find, taking into account that fermions and bosons have opposite powers of determinant (leading to the $(-1)^F$ insertion below):

$$S = \ln \det(\Delta + m^2 + \sigma_L F) = \int_\epsilon^\infty \frac{ds}{s} \frac{\mathrm{Tr}(-1)^F e^{-sm^2} e^{-2se\sigma_L F}}{(2\sin(seF/2))^2}, \tag{32.9}$$

where σ_L denotes the Cartan element of $SU(2)_L$. For later convenience, we rescale $s \to s\lambda/e$ and rewrite Eq. 32.9 as

$$S = \int_\epsilon^\infty \frac{ds}{s} \frac{\mathrm{Tr}(-1)^F e^{-sm^2\lambda/e} e^{-2s\sigma_L \lambda F}}{(2\sin(s\lambda F/2))^2}. \tag{32.10}$$

With this technology in hand, we can try to understand the generation of these F-terms in the context of integrating out charged degrees of freedom from the target space viewpoint. Then comparing to the topological string partition function will give us a reinterpretation of the topological string.

32.3. Counting of D-branes and Topological String Amplitudes

As already noted, the basic idea is that the terms in Eq. (32.3) arise from integrating out some degrees of freedom — in this case, D2-branes wrapped around two-cycles of the Calabi–Yau threefold. These are minimally wrapped branes and are BPS states. They are similar to the kind of equations we studied in the context of BPS states in two dimensions in Ch. 18, namely the Landau–Ginzburg solitons, which preserved half of the supersymmetry. In fact, the situation is very similar to the case of the Schwinger computation: a wrapped D2-brane in class $Q \in H_2(X, \mathbb{Z})$ corresponds to a charged particle in four dimensions with charge Q whose mass is given by

$$m = \frac{1}{\lambda}\int_Q k = \frac{1}{\lambda} t_Q$$

where λ is the string coupling constant and where k denotes the Kähler form on X. Moreover the particle is charged with respect to the graviphoton field, with its charge equal to the mass, i.e.,

$$e = m.$$

There is one more difference between the case at hand and the Schwinger case. Here we have more supersymmetry than the original problem studied by Schwinger. This means that the $SO(4) = SU(2)_L \times SU(2)_R$ representation content of states preserving half of the supersymmetry is of the form

$$[(1/2, 0) + 2(0, 0)] \otimes \mathcal{R}, \tag{32.11}$$

where $\mathcal{R}$ is some representation of $SO(4)$. However, if for this particle content we compute the correction induced in the action of the form $\int R_+^2$, it can be shown that the computation becomes equivalent to that of Schwinger for the *non-supersymmetric* computation above with the represention $\mathcal{R}$. Thus the extra R_+^2 insertion has "absorbed" the extra representation in front of Eq. (32.11). In other words, for the computations involving corrections to $\int R_+^2$ for integrating out particles with representation given by Eq. (32.11), we might just as well consider the equivalent Schwinger computation with representation $\mathcal{R}$. Substituting $e = m$ in Eq. (32.10), using $m = t_Q/\lambda$, and absorbing the field strength F into λ (i.e., replacing λF by λ), we find

$$S = \int F(t,\lambda) R_+^2 \quad \text{where} \quad F(t,\lambda) = \int_\epsilon^\infty \frac{ds}{s} \frac{\mathrm{Tr}(-1)^F e^{-st_Q} e^{-2s\sigma_L \lambda}}{(2\sin(s\lambda/2))^2}. \tag{32.12}$$

In evaluating the above contribution we need to know the degeneracy of the wrapped D2-branes, as well as the $SO(4)$ content of the field they correspond to. First of all, we have to explain how a massive particle, which in four dimensions will transform according to some represention of the $SO(3)$ rotation group, gives rise to an $SO(4)$ representation. Secondly, it turns out that for each BPS state corresponding to a wrapped D2-brane, there is an infinite number of them labelled by an integer[4] with the replacement

$$t_Q \to t_Q + 2\pi i n \tag{32.13}$$

in Eq. (32.12). To explain these observations and to gain some further insight into the structure of these BPS states it turns out to be useful to use a duality that relates Type IIA superstrings in ten dimensions to a theory known as M-theory in 11 dimensions. The basic relation between the two is that the Type IIA superstring on the manifold $\mathbb{R}^4 \times X$ is equivalent to M-theory on $\mathbb{R}^4 \times X \times S^1$, where the radius R of the S^1 becomes large in the limit that the coupling constant λ of the Type IIA superstring becomes large ($R^3 = \lambda^2$).

M-theory has M2-branes (membranes) which are somewhat like the D2-branes of Type IIA. In the limit of small radius for the S^1, (eleven-dimensional) M-theory reduces to (ten-dimensional) Type IIA. Under this dimensional reduction, the M2-brane with one dimension wrapping the S^1

[4]This integer can be identified with the first Chern class of the line bundle on the D2-brane.

corresponds to the string of Type IIA, and the unwrapped membrane corresponds to the D2-brane. Thus in a sense M-theory unifies these two objects into one!

Since, as already mentioned, the F-terms are independent of the string coupling under suitable normalizations (for example $F \to F/\lambda$ in the above expressions), we can take $\lambda \to \infty$, in which case we have a new description of the Schwinger type of compuation in terms of M theory on $\mathbb{R}^5 \times X$.

In the resulting five-dimensional effective theory, the particles will form representations of the spatial rotation group $SO(4) \sim SU(2)_L \times SU(2)_R$. We can label these representations in terms of left- and right- $SU(2)$ spins j_L and j_R, which are (integers or) half-integers. The relevant BPS states are M2-branes wrapped on two-cycles in the Calabi–Yau X. Upon reduction on an extra circle, they correspond to the wrapped D2-branes of Type IIA superstrings we have been discussing. In fact each such M2-brane may have an additional momentum of n units around the extra circle. Thus for each $M2$ brane we get an infinity of wrapped D2-branes indexed by n, whose masses are proportional to $|t_Q + 2\pi i n|$. This in fact explains Eq. (32.13). We also see how to assign $SO(4)$ representations for wrapped D2-branes: The M-theory rotation group in one higher dimension gives rise to an unambiguous $SO(4)$ content for each particle.

Let us introduce the numbers $n^Q_{(j_L,j_R)}$ which count the number of BPS M2-branes (i.e., minimally-embedded curves) in the Calabi–Yau in the class Q, which give rise to particles that form the representation of the five-dimensional rotation group $SO(4)$ given by

$$[(1/2,0)+2(0,0)]\otimes(j_L,j_R).$$

It turns out that the numbers $n^Q_{(j_L,j_R)}$ are very sensitive to the data of the Calabi–Yau; they change as we change the complex structure, for instance. However the numbers

$$n^Q_{j_L} = \sum_{j_R}(-1)^{2j_R}(2j_R+1)n^Q_{(j_L,j_R)}$$

are invariant under smooth deformations of the theory. This is very similar to the phenomenon we discussed in Ch. 10 in the context of lifting of the ground states in the supersymmetric theories. It turns out that this combination of BPS states is protected by a similar supersymmetry argument. In other words, the $n^Q_{(j_L,j_R)}$ change because pairs of them can join to form a

non-BPS multiplet. This is the analogue of two ground states with opposite fermion numbers joining and both becoming massive. The fact that only $n^Q_{j_L}$ is invariant is quite satisfactory as precisely this kind of combination appears in Eq. (32.12). In particular only the $SU(2)_L$ content of the $SO(4)$ representation is relevant for the topological string computations.

Now we wish to return to Eq. (32.12) and compute the effect of all wrapped D2-branes. For our purposes, it is convenient to choose a non-standard basis for the $SU(2)_L$ representations given by

$$I_r \equiv I_1^{\otimes r} \equiv [(1/2)+2(0)]^{\otimes r},$$

with I_0 defined to be the trivial representation. A relevant fact that is easy to establish is that the I_r ($r=0,\ldots,\infty$) form a basis for $SU(2)$ representations with integer coefficients. So we have

$$\sum_{r=0}^{\infty} n^Q_r I_r = \sum_{j_L} n^Q_{j_L}[j_L], \tag{32.14}$$

where we have just defined a new set of numbers n^Q_r (not necessarily positive) and we have denoted the spin-j_L representation as $[j_L]$ to avoid confusion. Note that

$$\begin{aligned} \mathrm{Tr}_{I_1}(-1)^F e^{-2s\sigma_L\lambda} &= [2\,\sin(s\lambda/2)]^2,\\ \mathrm{Tr}_{I_r}(-1)^F e^{-2s\sigma_L\lambda} &= [\mathrm{Tr}_{I_1}(-1)^F e^{-s\sigma_L\lambda}]^r = [2\sin(s\lambda/2)]^{2r}. \end{aligned} \tag{32.15}$$

EXERCISE 32.3.1. *Verify the above statement.*

In other words each BPS particle with $SU(2)_L$ spin content I_r (in addition to the overall I_1) will contribute $[2\sin(s\lambda/2)]^{2r}$ to Eq. (32.12).

As noted before, each wrapped D2-brane gives rise to infinitely many BPS states, labeled by n, the momentum around the extra circle in the M-theory. Thus using Eq. (32.13) for each wrapped D2-brane in the charge class Q in the I_r representation the contribution to $F(t)$ is given by

$$\sum_n \int \frac{ds}{s} e^{-s(t_Q+2\pi i n)}[2\sin(s\lambda/2)]^{2r-2}. \tag{32.16}$$

This expression can be recast by first summing over n and using

$$\sum_n \exp(-2\pi i n s) = \sum_m \delta(s-m)$$

and then integrating over s. The answer is

$$\sum_{m\geq 0} \frac{1}{m} e^{-mt_Q} [2\sin(m\lambda/2)]^{2r-2}. \tag{32.17}$$

There are also contributions to the $F(t)$ from the unwrapped D2-branes, i.e., in the class $Q=0$. These would be constant additions to F_g independent of t. For $g=0,1$ the constant part of F_g is ambiguous (recall that we need to fix three points for genus 0 and 1 point for genus 1 to define the amplitude, which corresponds to taking derivatives with respect to t), so let us consider the contribution that unwrapped D-branes make for $g\geq 2$. These would have mass zero, were it not for the momentum n in Eq. (32.13). In other words they give rise to $m = 2\pi i n$. However here we restrict to $n>0$ because negative n corresponds to an anti-particle state. There are effectively $-\chi$ such particles where χ is the Euler characteristic of the Calabi–Yau.[5] Now, using

$$\int \frac{ds}{s} \frac{e^{-sZ}}{[2\sin(s\lambda/2)]^2} = \sum \left(\frac{\lambda}{Z}\right)^{2g-2} \chi_g$$

where $\chi_g = (-1)^{g-1} B_g / 2g(2g-2)$ (which is also the Euler characteristic of moduli space of genus g curves) and B_g is the gth Bernoulli number, we find that the contribution of unwrapped D-branes is given by

(32.18)
$$\frac{-\chi}{2}\lambda^{2g-2}\chi_g \sum_{n\neq 0} (2\pi i n)^{2-2g} = (-1)^{g-1}\chi_g \chi \frac{\zeta(2g-2)}{(2\pi)^{2g-2}} = \frac{-\chi B_g B_{g-1}}{4g(2g-2)(2g-2)!}.$$

As noted before these only contribute to $g>1$. For genus 0 and genus 1, there are additional t-dependent terms that are present even if we have no wrapped D2-branes, and that can be viewed as coming from the massless modes, which we have to take into account in order to get the full answer for F_0 and F_1. For genus 0 we have the leading term of F_0 which comes

[5]The way this arises is that the moduli space of D0-branes in the Calabi–Yau is itself the Calabi–Yau, and the particles are in one-to-one correspondence with the ground states of the corresponding supersymmetric theory, i.e., as discussed in Ch. 10, the cohomology elements of the Calabi–Yau. Moreover, in this case the Lefschetz $SU(2)$ action on the cohomology of the Calabi–Yau coincides with the $SU(2)_R$ quantum number of the $SO(4)$ rotation group. In other words the particle content will correspond to the representation of $[(0,3/2)] + (h^{1,1}-1)[(0,1/2)] + (2h^{2,1}+2)[(0,0)]$. Thus the $SU(2)_L$ content of the particles is just minus the Euler characteristic of the Calabi–Yau.

from the classical triple intersection and is of the form

$$\int_X k^3 + P_2(t) = \frac{1}{6} C_{ijk} t^i t^j t^k + P_2(t), \tag{32.19}$$

where C_{ijk} denotes the classical intersection and $P_2(t)$ is ambiguous,[6] in the sense that strictly speaking the topological string at genus 0 is only defined through $\partial^3/\partial t^3$. Similarly, for genus 1, one can show that there is a leading term of the form

$$\frac{-1}{24}\int_X k \wedge c_2 = \frac{-1}{24} c_2^i t_i + \text{const.}, \tag{32.20}$$

where the const. reflects the ambiguity of addition of a constant to F_1. Here c_2 denotes the second Chern class of X. Putting together the contribution of wrapped and unwrapped D2-branes to the topological string amplitude from Eq. (32.17) and Eq. (32.18), as well as the massless mode contributions, we find

$$\begin{aligned} F(t,\lambda) = {} & \frac{1}{\lambda^2}\left[\frac{1}{6}C_{ijk}t^i t^j t^k + P_2(t)\right] + \frac{-1}{24} c_2^i t_i + \text{const.} \\ & + \sum_{g>1} \frac{-\chi B_g B_{g-1}}{4g(2g-2)(2g-2)!}\lambda^{2g-2} + \sum_{m,Q,r} \frac{e^{-mt_Q}}{m[2\sin(m\lambda/2)]^{2-2r}} n_r^Q, \end{aligned} \tag{32.21}$$

where the sum on the RHS is over all $m>0$, $r\geq 0$, $Q\in H_2(X,\mathbb{Z})$ and the n_r^Q are all integers and are known as Gopakumar–Vafa invariants. We have thus been able to rewrite the topological string partition function in terms of integral quantities related to spectrum of wrapped BPS D2-branes. This implies that Gromov–Witten invariants at all genera, which are not in general integral, can be captured by integral Gopakumar–Vafa invariants.

In order to obtain various F_g from Eq. (32.21) we need to expand the above expression in powers of λ and collect the terms with power λ^{2g-2}. For example, the genus 0 term comes from the n_r^Q with $r=0$ and by replacing $\sin(m\lambda/2)$ by $m\lambda/2$ in the last terms of the above expression.

$$F_0(t) = \left[\frac{1}{6}C_{ijk}t^i t^j t^k + P_2(t)\right] + \sum_{m>1,Q} \frac{n_0^Q}{m^3} e^{-mt_Q}. \tag{32.22}$$

The rational GW invariants at genus 0 are thus captured by the integral GV invariants n_0^Q. This integrality prediction is in line with the result obtained thus far from all computations of GW invariants at genus zero. This is

[6]However, there seems to be a structure of the form $P_2(t) = \frac{-\chi}{2}\zeta(3) - \frac{\pi^2}{6}\int_X c_2 \wedge k + A_{ij}t^i t^j$, where A_{ij} is not caonically defined.

also consistent with the contributions coming from isolated genus 0 rational curves, as discussed in Ch. 33. Notice that all representations I_r ($r > 0$) have equal numbers of bosons and fermions, so n_0^Q computes the net number (weighted by $(-1)^F$) of BPS states.

From the structure of Eq. (32.21) it is clear that for a given F_g only the n_r^Q with $r \leq g$ contribute. This allows us to compute n_r^Q recursively, if we know all F_g with $g \leq r$. In fact this is one reason for the organization of the BPS degeneracies in the representations given by I_r. As we have briefly mentioned, the holomorphic anomaly equation is ideally applied to the computation of F_g in an inductive fashion; in other words starting from genus 0 and working up towards higher g. Aspects of this will be discussed in Ch. 33. [7]

32.4. Black Hole Interpretation

Wrapped BPS M2-branes are particles, and if they have sufficiently large mass they can be viewed as five-dimensional black holes in the context of M-theory compactification on Calabi–Yau threefolds. Then the numbers of BPS states n_r^d are related to computations of the entropy (the logarithm of the number of quantum states) of spinning black holes of "charge" d. For large degrees d, these black holes are very massive, and the appropriate description in that regime is gravity. We can use this description to make predictions about the growth of these numbers n_r^d based on our understanding of black holes (due to the so-called Bekenstein–Hawking entropy). The prediction is that

$$\sum_r n_r^d \binom{2r+2}{r+1+m} \sim \exp(\sqrt{d^3 - m^2}), \tag{32.23}$$

in the limit $d \gg 1$, $d \gg m$ where d is the charge of the black hole and m is its spin. There is no known way to prove in general the validity of these predictions (though they have been verified in special cases).

[7]How should one think about this reformulation of GW invariants in terms of GV invariants? It might be useful to draw an analogy to Donaldson invariants, which are related to intersection theory on the moduli space of instantons on four-manifolds. Seiberg and Witten reformulated the problem in terms of a simpler $U(1)$ gauge theory and this led to a reformulation of the Donaldson invariants in terms of some simpler invariants — the Seiberg-Witten invariants. This is roughly analogous to what we are seeing here.

We now turn to the question of how one goes about computing the GV invariants n_r^Q.

CHAPTER 33

Mathematical Formulation of Gopakumar–Vafa Invariants

We have just noted in Ch. 32 that the GV invariants completely determine the topological amplitudes. Here we would like to discuss a mathematical framework that can be used to compute them directly. Related mathematical issues, especially integrality issues in Gromov–Witten theory, will be discussed in Ch. 34.

Recall from Ch. 32 that n_r^Q capture the $SU(2)_L$ content of the number of wrapped BPS D2-branes with charge $Q \in H_2(X, \mathbb{Z})$ in a particular basis for the $SU(2)_L$ representation ring.

As we noted in our discussion of D2-branes, there is a flat vector bundle living on the brane which pulls back to the the boundary of the worldsheet. Now we have to consider the condition of being a BPS state. This essentially amounts to the condition that the D2-brane wraps a holomorphic curve in X and that the bundle is stable.[1] In other words, the mathematical notion of stability is the same as the physical notion of stability, namely the notion of being a ground state. Thus, roughly speaking, we are studying components $\widehat{\mathcal{M}}$ of the moduli space $\mathcal{M}^Q$ of stable sheaves F in the CY where the class of the sheaf is given by $c_1(F) = Q$. In order to count the degeneracy of D2-branes we have to study the supersymmetric quantum mechanics problem whose target manifold is the moduli space $\widehat{\mathcal{M}}$ of stable sheaves under consideration. The spectrum of BPS states is therefore given by the cohomology of $\widehat{\mathcal{M}}$. We also have to give the $SU(2)_L$ content of the cohomology of this moduli space.

In order to see how to read off the n_r^Q, we have to understand how the $SU(2)_L$ and $SU(2)_R$ act on the cohomology of $\widehat{\mathcal{M}}$. Kähler manifolds admit an $SU(2)$ action on their cohomology, where the $SU(2)$ raising operator J_+

[1] The condition is more subtle, as the curves can have several components, the components can have multiplicities, and the bundle needs to be replaced by sheaves more generally. We will see this in later examples.

corresponds to wedging with the Kähler class, J_- corresponds to its adjoint and J_3 acting on $H^{p,q}$ has an eigenvalue given by

$$(p+q-\dim_{\mathbb{C}}\widehat{\mathcal{M}})/2. \tag{33.1}$$

For example, the cohomology of $\mathbb{CP}^n$ forms a spin $n/2$ representation of $SU(2)$, which we will denote by $(\mathbf{n/2})$.

Since $\widehat{\mathcal{M}}$ is a Kähler manifold, there is a natural Lefschetz $SU(2)$ action on its cohomology. It turns out that this $SU(2)$ corresponds to the diagonal embedding, i.e., $SU(2) = \mathrm{diag}(SU(2)_L \times SU(2)_R)$. In order to identify the further $SU(2)_L$ and $SU(2)_R$ decomposition of the moduli space we need further data which turns out to come from the fibration structure of $\widehat{\mathcal{M}}$.

Let $\mathcal{M}$ denote the moduli space of the embedded curve, i.e., forgetting the bundle data in $\widehat{\mathcal{M}}$. Then there is a bundle structure

$$\widehat{\mathcal{M}} \to \mathcal{M} \tag{33.2}$$

where the fibers correspond to the moduli space of the flat bundle for a fixed embedded curve. In the case of a genus g curve Σ and line bundles, the fiber is generically isomorphic to the Jacobian $\mathrm{Jac}(\Sigma)$, which is the g complex dimensional torus parametrizing moduli of flat $U(1)$ bundles over Σ. Roughly speaking, the $SU(2)_L$ is identified with the fiber Lefschetz action and the $SU(2)_R$ with that of the base. Since we are only interested in the $SU(2)_L$ action, let us discuss that in more detail. What this action means is as follows: there is a $(1,1)$-form k_f corresponding to the Kähler class of the fiber of $\widehat{\mathcal{M}}$. [2] This is expected to be true even if there are degenerate fibers in the fibration Eq. (33.2) (consider for example elliptic surfaces). We can consider the action of k_f on the cohomology of $\widehat{\mathcal{M}}$ and deduce from that the $SU(2)_L$ content of the cohomology of $\widehat{\mathcal{M}}$.

There are two easy consequences of the above statement:

$$n_0^Q = (-1)^{\dim_{\mathbb{C}}(\widehat{\mathcal{M}})}\chi(\widehat{\mathcal{M}}) \qquad n_g^Q = (-1)^{\dim_{\mathbb{C}}(\mathcal{M})}\chi(\mathcal{M}) \tag{33.3}$$

The first statement follows from what we already noted before, that n_0 counts the total number of BPS states alternating between bosons and fermions.

The statement about n_g^Q follows by noting that the cohomologies corresponding to highest $SU(2)_L$ spin $g/2$ is given by $(1, k_f, ..., k_f^g)$ times the

[2] This is to be understood as a heuristic formulation, since we can add the pull-back of any Kähler class on the base to k_f without changing the Kähler class of the fiber.

cohomology elements of $\mathcal{M}$. Said differently, the $SU(2)_L \times SU(2)_R$ representation contains a summand of the form

$$\left(\frac{\mathbf{g}}{\mathbf{2}}\right) \otimes H^*(\mathcal{M}), \tag{33.4}$$

where the $(\mathbf{g/2})$ denotes an $SU(2)_L$ representation and $H^*(\mathcal{M})$ is an $SU(2)_R$ representation via the Lefschetz action on the base, and furthermore, this is the total content of the $(\mathbf{g/2})$ of $SU(2)_L$. Since the $SU(2)_R$ measures the cohomology with respect to the base, but n_r^Q are insensitive to them (except for the parity $(-1)^{F_R}$), we obtain the above formula.

Let us formulate the problem in a mathematical setting. We will see in Ch. 34 that, for a general X, we have

$$F_g = \sum_{\beta \in H_2(X,\mathbb{Z})} N_\beta^g q^\beta$$

where $N_\beta^g \in \mathbb{Q}$ are GW-invariants. When X is a Calabi–Yau threefold, the GW-invariants are given by the degree of the virtual fundamental class

$$N_\beta^g = \deg\left[\overline{\mathcal{M}}_{g,0}(X,\beta)\right]^{virt}. \tag{33.5}$$

Despite the precise definition Eq. (33.5), there are relatively few situations where the virtual fundamental class can be computed explicitly. A general situation where the virtual fundamental class can be computed arises when the computation can be translated into a computation on a space admitting a torus action. When $g = 0$, this is the case for hypersurfaces and complete intersections in projective spaces, and the virtual fundamental class can be expressed in terms of the Euler class of a bundle on $\overline{\mathcal{M}}_{0,0}(\mathbb{CP}^r, d)$ for the appropriate r and d. These ideas have been discussed extensively in Part 4. This formula holds even without assuming the existence of a torus action, but is only useful if there is a torus action. No such formulas exist for $g > 0$.

The data of the N_β^g is entirely encoded in the BPS invariants n_β^r [3] which are related by

$$F = \sum \lambda^{2g-2} F_g = \sum n_\beta^r \frac{1}{m}\left(2\sin\frac{m\lambda}{2}\right)^{2r-2} q^{m\beta} \tag{33.6}$$

[3] There is an unfortunate mismatch between standard notation in the mathematics and physics literature. The notation n_β^r is standard in the mathematics literature and also matches usage in the next chapter. Furthermore, the genus 0 invariants n_d introduced in Part 4 could be expressed as n_d^0 in this notation. But n_β^r would have been written as n_r^β in the previous section. When we want to make a direct reference to the previous section we will use the explicit symbol Q in place of β and will revert to the old notation n_r^Q.

as in Ch. 32. This can be thought of physically as a formula for the F_g in terms of the BPS states, or as a mathematical definition of the n^g_β which is made rigorous through Eq. (33.5). In this form, it is not clear that the n^g_β are even integers. Verifying the integrality is an important mathematical question, and is the subject of investigation.

A more modest goal is to attempt to define the n^g_β geometrically, so that their integrality is manifest. Then Eq. (33.6) becomes a conjecture that needs to be proven. Our main goal is to understand how to define these invariants.

Let us look at some simple examples:

EXAMPLE 33.0.1. *Consider the local geometry in the Calabi–Yau given by*

$$\mathcal{O}(-1) + \mathcal{O}(-1) \to \mathbb{CP}^1.$$

Choose a class $Q = d[\mathbb{CP}^1]$ *and look at* $\widehat{\mathcal{M}}$. *In the* $d = 1$ *case, there are no normal deformations for the D2-brane wrapped on the* $\mathbb{CP}^1$, *so the base* $\mathcal{M}$ *is trivial. There are also no flat connections on the* $\mathbb{CP}^1$, *so the fiber is trivial as well. The moduli space* $\widehat{\mathcal{M}}$ *is thus a point. Therefore* $n^0_1 = 1$. *Also* $n^0_d = 0$ *because the moduli space of stable rank d bundles on* $\mathbb{CP}^1$ *is the empty set (this is the statement that there are no D2-brane bound states on* $\mathbb{CP}^1$*). Also* $n^r_d = 0$ *for* $r > 0$ *as there are no higher genus embedded curves in this example (recall that we are not considering maps to the target space, but embedded curves). We thus obtain for the contribution of non-trivial holomorphic maps in the GW theory,*

$$F(t,\lambda) = \sum_m \frac{e^{-mt}}{m(2\sin(m\lambda/2))^2} \tag{33.7}$$

We see that even this trivial example captures a lot of non-trivial information.

In Example 33.0.2 below, we will consider the local geometry

$$\mathcal{O}(-3) \to \mathbb{CP}^2,$$

and study D2-branes wrapping curves in $\mathbb{CP}^2$ of degrees $d = 1, 2$. For ease of exposition, we pause to give some formulas for plane curves of arbitrary degree d. A degree d curve in $\mathbb{CP}^2$ is the zero-locus of an equation of the form

$$\sum_{i+j+k=d} a_{ijk} x_0^i x_1^j x_2^k = 0, \tag{33.8}$$

where (x_0, x_1, x_2) are coordinates on $\mathbb{CP}^2$. There are $(d^2 + 3d + 2)/2$ coefficients in this equation, and scalar multiplication of the equation does not alter the curve, so

$$\mathcal{M} = \mathbb{CP}^{(d^2+3d)/2}. \tag{33.9}$$

The genus of a smooth plane curve of degree d is

$$g = \frac{(d-1)(d-2)}{2}. \tag{33.10}$$

In particular, curves in $\mathbb{CP}^2$ with $d = 1, 2$ have $g = 0$.

EXAMPLE 33.0.2. *We now study this local geometry* $\mathcal{O}(-3) \to \mathbb{CP}^2$, *and D2-branes wrapping curves with* $d = 1, 2$. *Since* $g = 0$ *in each case, the bundles have no deformations, and* $\widehat{\mathcal{M}} = \mathcal{M}$.

For the case $d = 1$, *we let* $Q = [H]$ *be the hyperplane class. Now* $\mathcal{M}$ *is itself a* $\mathbb{CP}^2$, *by Eq. (33.9), whose cohomology is in the* $(\mathbf{1})$ *representation. Since* $g = 0$, *Eq. (33.4) says that BPS states contain the* $(\mathbf{0}, \mathbf{1})$ *representation of* $SU(2)_L \times SU(2)_R$. *Since* $\widehat{\mathcal{M}} = \mathcal{M}$, *this is the entire BPS spectrum. This gives us the element* $3(\mathbf{0}) = 3I_0$ *of the* $SU(2)_L$ *representation ring. Using* $\sum_r n^Q_r I_r = \sum n^Q_{j_L,j_R}(-1)^{2j_R}(2j_R+1)[j_L]$, *we see that* $n^0_1 = 3$ *and* $n^r_1 = 0$ *for* $r > 0$.

Next, we consider $d = 2$, *so we put* $Q = 2[H]$. *Again* $g = 0$ *and there are no flat bundles, so* $\widehat{\mathcal{M}} = \mathcal{M}$, *which is* $\mathbb{CP}^5$ *by Eq. (33.9). We deduce as above that the BPS states form the representation* $(\mathbf{0}, \mathbf{5/2})$ *of* $SU(2)_L \times SU(2)_R$, *and are led to the element* $-6(\mathbf{0}) = -6I_0$ *of the representation ring of* $SU(2)_L$ *(note the minus sign, since the* $(\mathbf{5/2})$ *representation is fermionic). We conclude that* $n^0_2 = -6$ *and* $n^r_2 = 0$ *for* $r > 0$. *These numbers are in agreement with low-genus computations of Gromov–Witten invariants.*

EXAMPLE 33.0.3. *Now for a case where the base is trivial, i.e.,* $\mathcal{M}$ *is a point (no deformations of the curve) but the fibre is nontrivial (flat connections exist). This is the case of a rigid curve* Σ_g *of genus* $g > 0$ *in a* CY. *Here* $\widehat{\mathcal{M}} = \mathrm{Jac}(\Sigma_g) = T^{2g}$. *Now we have a non-trivial* $SU(2)_L$ *action on the cohomology, and in fact the* $SU(2)_L$ *raising operator is given by wedging* *with the Kähler class θ of the theta divisor on $\mathrm{Jac}(\Sigma_g)$. It is easy to see that the BPS states are in the left-representation I_g. To see this, notice first that the cohomology of T^2 forms the representation $(\frac{1}{2}) + 2(\mathbf{0}) = I_1$ and that the cohomology of T^{2g} forms the g^{th} tensor product of this representation, which is the I_g representation. In fact, this is the reason for the definition of I_g. So*

$$n^g_{[\Sigma_g]} = 1$$

and the rest of $n^r_{[\Sigma_g]}$ vanish.

What about higher degrees ($Q = d[\Sigma_g]$)? Let us look at $\Sigma_g = T^2$ for simplicity. In this case the moduli space is once again T^2 (the moduli space of stable rank d bundles on T^2 is itself a T^2, for all d), so we find that $n^1_{d[T^2]} = 1$ and $n^r_{d[T^2]} = 0$ for $r > 1$, this being true for all d. For Σ a higher genus curve, the higher-degree cases are harder because the moduli space has singularities. Though we start out with a rigid curve (one with no normal deformations), there are some loci on the moduli space of flat connections on the curve from which there emerge branches of the moduli space corresponding to normal deformations. This would be interesting to study in greater detail.

EXAMPLE 33.0.4. *We return to $\mathcal{O}(-3)$ over $\mathbb{CP}^2$, and now consider $d=3$. We have in this case $g = 1$ by Eq. (33.10), and by Eq. (33.9) we have $\mathcal{M} = \mathbb{CP}^9$, which has Lefschetz decomposition $(\mathbf{9/2})$.*

By Eq. (33.4) we get

$$H^*(\widehat{\mathcal{M}}) = \left(\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{9}}{\mathbf{2}}\right) \oplus (\mathbf{0}, R) \tag{33.11}$$

for some $SU(2)_R$ representation R (not necessarily irreducible). To find R, recall that the diagonally embedded $SU(2) \subset SU(2)_L \times SU(2)_R$ gets identified with the Lefschetz decomposition of $H^*(\widehat{\mathcal{M}})$. Restricted to the diagonal,

$$H^*(\widehat{\mathcal{M}}) = \left(\frac{\mathbf{1}}{\mathbf{2}}\right) \otimes \left(\frac{\mathbf{9}}{\mathbf{2}}\right) \oplus R = (\mathbf{4}) \oplus (\mathbf{5}) \oplus R. \tag{33.12}$$

This is as far as we can get with the techniques of the previous examples, but we can do better fairly easily. We want to study $\widehat{\mathcal{M}} = \{(C, \mathcal{L})\}$, where C is a cubic plane curve, $\mathcal{L}$ is a line bundle, where we are free to fix a degree. If C is smooth, then degree 1 bundles are given by $\mathcal{L} \simeq \mathcal{O}(p)$ for some $p \in C$. Therefore,

$$\widehat{\mathcal{M}} \simeq \mathcal{C} = \{(C, p) : p \in C\},$$

the universal curve over $\mathcal{M}$.[4] *We now turn to the computation of the cohomology of $\mathcal{C}$ and its Lefschetz decomposition. We have the diagram*

$$\begin{array}{ccccc} \mathbb{CP}^8 & \to & \mathcal{C} & \subset & \mathbb{CP}^2 \times \mathbb{CP}^9 \\ \downarrow & & \downarrow \pi & & \\ p & \in & \mathbb{CP}^2 & & \end{array} \tag{33.13}$$

where the fiber of π over $p \in \mathbb{CP}^2$ is isomorphic to $\mathbb{CP}^8$, as indicated in Eq. (33.13). To see this, note that $\pi^{-1}(p) = \{C : p \in C\} \subset \mathbb{CP}^9$. If C is described by Eq. (33.8), then the condition $p \in C$ is a non-trivial linear equation in the coefficients a_{ijk}. So $\pi^{-1}(p)$ is a hyperplane in $\mathbb{CP}^9$, justifying our assertion. Thus $\mathcal{C}$ is a $\mathbb{CP}^8$ fibration over $\mathbb{CP}^2$.

This is all that we need. We get $e(\mathcal{C}) = e(\mathbb{CP}^2)e(\mathbb{CP}^8) = 27$ for the Euler characteristic, but we can see the Lefschetz decomposition just as easily. Since $H^*(\mathbb{CP}^2) = (\mathbf{1})$ and $H^*(\mathbb{CP}^8) = (\mathbf{4})$ as $SU(2)$ Lefschetz representations, we conclude that

$$H^*(\mathcal{C}) = (\mathbf{1}) \otimes (\mathbf{4}) = (\mathbf{3}) \oplus (\mathbf{4}) \oplus (\mathbf{5}). \tag{33.14}$$

Comparing Eq. (33.14) with Eq. (33.12), we conclude that $R = (\mathbf{3})$. Then substituting into Eq. (33.11), we get

$$H^*(\widehat{\mathcal{M}}) = \left(\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{9}}{\mathbf{2}}\right) \oplus (\mathbf{0}, \mathbf{3}). \tag{33.15}$$

Restricting Eq. (33.15) to $SU(2)_L$ and taking the sign into account, we get

$$H^*(\widehat{\mathcal{M}}) = -10\left(\frac{\mathbf{1}}{\mathbf{2}}\right) + 7\,(\mathbf{0}) = -10 I_1 + 27 I_0.$$

Since $H^*(\mathcal{C}) = \sum n^r_3 I_r$, we have $n^0_3 = 27$ and $n^1_3 = -10$, while all other n^g_3 are zero.

As a check, note that $n^0_3 = (-1)^{\dim \widehat{\mathcal{M}}} e(\widehat{\mathcal{M}}) = e(\mathcal{C}) = 27$ and $n^1_3 = (-1)^{\dim \mathcal{M}} e(\mathcal{M}) = -e(\mathbb{CP}^9) = -10$, so that Eq. (33.3) holds.

We now proceed to a more general situation. Our starting point is to explain a relationship between the $SU(2)$ weights of the cohomology of the Jacobian of a smooth curve C and the cohomology of its symmetric product.

[4] When we investigate the situation more carefully below to include singular curves, we will see that we actually need to consider bundles of degree -1, which are of the form $\mathcal{L} \simeq \mathcal{O}(-p)$ for some $p \in C$ if C is smooth. Since the moduli space of the $\mathcal{O}(p)$ is the same as the moduli space of the $\mathcal{O}(-p)$, we can and will ignore this subtlety for the moment.

The symmetric products arise naturally since an element $\{z_1, \ldots, z_p\}$ of the symmetric product gives rise to a sheaf $\mathcal{O}_C(\sum z_i)$ on C.

We start with the simplest case and relate $H^*(C)$ to $H^*(\mathrm{Jac}(C))$.

For illustration purposes, we expand the first few $I_g = I_1^g$ in terms of the irreducible representations of $SU(2)$.

$$\begin{aligned} I_1 &= (\tfrac{1}{2}) + 2(\mathbf{0}), \\ I_2 &= (\mathbf{1}) + 4\,(\tfrac{1}{2}) + 5(\mathbf{0}), \\ I_3 &= (\tfrac{3}{2}) + 6(\mathbf{1}) + 14\,(\tfrac{1}{2}) + 14(\mathbf{0}). \end{aligned} \tag{33.16}$$

Applying θ^{g-1} for $g = 1, 2, 3$ to Eq. (33.16), we get

$$\begin{aligned} I_1 &= (\tfrac{1}{2}) + 2(\mathbf{0}), \\ \theta I_2 &= (\tfrac{1}{2}) + 4\,(\mathbf{0}), \\ \theta^2 I_3 &= (\tfrac{1}{2}) + 6(\mathbf{0}). \end{aligned} \tag{33.17}$$

Note that the left-hand side of Eq. (33.17) is not an $SU(2)$ representation, but does have well-defined $SU(2)$ weights. So the right-hand side of Eq. (33.17) should be interpreted in the same fashion. A quantity $(\mathbf{a})$ now denotes a vector space with weights equal to the weights of the $(\mathbf{a})$ representation, possibly shifted. It is with this modified meaning that the formulas below involving θ and its powers are to be understood.

The generalization of Eq. (33.17) is clearly

$$\theta^{g-1}\mathrm{Jac}(C) = \left(\frac{\mathbf{1}}{\mathbf{2}}\right) \oplus 2g\,(\mathbf{0}) \tag{33.18}$$

as is easily checked by induction. Note that we have identified I_g with $J(C)$.

Note that $H^*(C) = (\mathbf{1/2}) \oplus 2g(\mathbf{0})$. Comparing with Eq. (33.18), we conclude that

$$H^*(C) = \theta^{g-1}\mathrm{Jac}(C), \tag{33.19}$$

which is the promised relationship.

Now, let $\mathrm{Sym}^2(C)$ be the symmetric product of C, consisting of unordered pairs of points of C (possibly identical). We will exhibit a relationship between the cohomology of $\mathrm{Jac}(C)$ and the cohomology of $\mathrm{Sym}^2(C)$. We consider $H^*(C) = (\mathbf{1/2}) \oplus 2g(\mathbf{0})$ as a superalgebra where $(\mathbf{1/2})$ is even and $2g(\mathbf{0})$ is odd (since they represent even and odd cohomology, respectively). Then $H^*(\mathrm{Sym}^2(C)) \simeq \mathrm{Sym}^2 H^*(C)$, where the symmetric product on the right is taken in the sense of superalgebras. The result is

$$H^*(\mathrm{Sym}^2 C) = (\mathbf{1}) \oplus 2g\left(\frac{\mathbf{1}}{\mathbf{2}}\right) \oplus \begin{pmatrix} 2g \\ 2 \end{pmatrix} (\mathbf{0}). \tag{33.20}$$

We now apply θ^{g-2} to Eq. (33.16) for $g = 2, 3$ to get

$$\begin{aligned} I_2 &= (\mathbf{1}) \oplus 4\,(\tfrac{1}{2}) \oplus 5(\mathbf{0}), \\ \theta I_3 &= (\mathbf{1}) \oplus 6(\tfrac{1}{2}) \oplus 14\,(\mathbf{0}). \end{aligned} \tag{33.21}$$

The generalization of Eq. (33.21) is

$$\theta^{g-2} I_g = (\mathbf{1}) \oplus 2g\left(\frac{\mathbf{1}}{\mathbf{2}}\right) \oplus (2g^2 - g - 1)(\mathbf{0}) \tag{33.22}$$

with the same interpretation as in Eq. (33.17). Conventionally putting $\mathrm{Sym}^0(C)$ equal to a point, we have $H^*(\mathrm{Sym}^0 C) = (\mathbf{0})$. Comparing Eq. (33.22) and Eq. (33.20), we obtain

$$H^*(\mathrm{Sym}^2 C) = \theta^{g-2} I_g \oplus H^*(\mathrm{Sym}^0 C). \tag{33.23}$$

The techniques leading to Eq. (33.19) and Eq. (33.23) generalize. The result is

$$H^*(\mathrm{Sym}^p C) = \theta^{g-p} H^*(\mathrm{Jac}(C)) \oplus H^*(\mathrm{Sym}^{p-2} C). \tag{33.24}$$

In generalizations, we will need to consider singular curves. As we will see below, it will be possible to associate sheaves to elements of the Hilbert scheme $\mathrm{Hilb}^p(C)$. The Hilbert scheme coincides with the symmetric product $\mathrm{Sym}^p(C)$ if C is smooth, but differs from the symmetric product in general. An element of $\mathrm{Hilb}^p(C)$ consists of a subscheme $Z \subset C$ of length p. This means that Z is the zero-locus of equations forming an ideal $I_Z \subset \mathcal{O}_C$, and that the quotient $\mathcal{O}_Z = \mathcal{O}_C / I_Z$ is a vector space of dimension p. A typical example is a set $Z = \{z_1, \ldots, z_p\}$ of distinct points of C, but fewer points with some points having multiplicity greater than 1 can occur.

We now let C vary in a family $\mathcal{M}$ as in the beginning of this section. Since $\mathcal{M}$ typically contain singular curves, we will need to consider the Hilbert schemes of these curves. We actually consider the *relative Hilbert scheme*

$$\mathcal{C}^{(p)} = \{(C, Z) \mid C \in \mathcal{M},\ Z \in \mathrm{Hilb}^p(C)\}. \tag{33.25}$$

Looking at Eq. (33.24), we see that since C varies as the fibers of the family $\mathcal{C} = \mathcal{C}^{(1)}$ over $\mathcal{M}$, then $\mathrm{Sym}^p(C)$ varies as the fibers of the family $\mathcal{C}^{(p)}$ over $\mathcal{M}$. Furthermore, $\mathrm{Jac}(C)$ varies as the fibers of the family $\widehat{\mathcal{M}}$ over $\mathcal{M}$.

Our natural proposal is that Eq. (33.24) generalizes to

$$H^*(\mathcal{C}^{(p)}) = \theta^{g-p} H^*(\widehat{\mathcal{M}}) \oplus H^*(\mathcal{C}^{(p-2)}). \tag{33.26}$$

The strategy is to use Eq. (33.26), the geometry of the $\mathcal{C}^{(p)}$, and the implicit definition of the $n^r_{[C]}$ as

$$H^*(\widehat{\mathcal{M}}) = \sum n^r_{[C]} I_r$$

to compute the $n^r_{[C]}$ in explicit cases.

Before we turn to examples, it will be helpful to first write down some consequences of Eq. (33.26). In the case $p = 0$, Eq. (33.26) is simply

$$H^*(\mathcal{M}) = \theta^g H^*(\widehat{\mathcal{M}}) \tag{33.27}$$

since $\mathcal{C}^{(0)} = \mathcal{M}$. We apply $\mathrm{Tr}(-1)^F$ to Eq. (33.27) and obtain

$$(-1)^{\dim \mathcal{M}} e(\mathcal{M}) = n^g_{[C]} \tag{33.28}$$

and we have recovered the second formula in Eq. (33.3).

The sign on the left of Eq. (33.28) arises from a statement for $\mathcal{M}$ analogous to Eq. (33.1).

In the case $p = 1$, Eq. (33.26) reads

$$\begin{aligned} H^*(\mathcal{C}) &= \theta^{g-1} H^*(\widehat{\mathcal{M}}) \\ &= \theta^{g-1}(n^g_{[C]} I_g + n^{g-1}_{[C]} I_{g-1} + \cdots) \\ &= n^{g-1}_{[C]}(\mathbf{0}) \oplus n^g_{[C]} \left((\tfrac{1}{2}) \oplus 2g(\mathbf{0}) \right). \end{aligned} \tag{33.29}$$

Applying $\mathrm{Tr}(-1)^F$ to Eq. (33.29), we get

$$(-1)^{\dim \mathcal{M}+1} e(\mathcal{C}) = (2g-2) n^g_{[C]} + n^{g-1}_{[C]}, \tag{33.30}$$

where we have used $\dim \mathcal{C} = \dim \mathcal{M} + 1$.

EXAMPLE 33.0.5. *We continue with $X = \mathcal{O}_{\mathbb{CP}^2}(-3)$ and now choose $d = 4$, so that $g = 3$ by Eq. (33.10) and $\mathcal{M} = \mathbb{CP}^{14}$ by Eq. (33.9). By Eq. (33.28), we have $n^3_4 = e(\mathcal{M}) = e(\mathbb{CP}^{14}) = 15$.*

We now compute n^2_4 using Eq. (33.30), which reads

$$-e(\mathcal{C}) = n^2_4 + 4 n^3_4 \tag{33.31}$$

in this case. The method of Example 33.0.4 applies, and in particular Eq. (33.13) gets replaced by

$$\begin{array}{ccl} \mathbb{CP}^{13} \to & \mathcal{C} \subset \mathbb{CP}^2 \times \mathbb{CP}^{14} \\ & \downarrow \\ & \mathbb{CP}^2 \end{array}$$

whereby $\mathcal{C}$ is exhibited as a $\mathbb{CP}^{13}$ fibration over $\mathbb{CP}^2$. We therefore compute that $e(\mathcal{C}) = 3 \cdot 14 = 42$. Substituting the known values into Eq. (33.31), we solve for n^2_4 and obtain $n^2_4 = -102$. To find n^1_4, we need to use higher Hilbert schemes. The generalization of Eq. (33.28) and Eq. (33.30) to $p = 2$ is easily computed using Eq. (33.18), Eq. (33.22), and Eq. (33.26) to be

$$(-1)^{\dim \mathcal{M}} e(\mathcal{C}^{(2)}) = (2g^2 - 5g + 3) n^g_{[C]} + (2g-4) n^{g-1}_{[C]} + n^{g-2}_{[C]}. \tag{33.32}$$

To compute n^1_4, all we need is the Euler characteristic of the relative Hilbert scheme $\mathcal{C}^{(2)}$. The fiber of the natural map $\mathcal{C}^{(2)} \to \mathrm{Hilb}^2(\mathbb{CP}^2)$ over $\{p, q\} \in \mathrm{Hilb}^2(\mathbb{CP}^2)$ over a degree 4 curve C is $\{C \mid p, q \in C\} \subset \mathbb{CP}^{14}$, which is isomorphic to $\mathbb{CP}^{12}$. To see this, note that the conditions $p, q \in C$ become two independent linear equations in the a_{ijk} when C is represented by Eq. (33.8). The fiber can be seen to still be isomorphic to $\mathbb{CP}^{12}$ even when $Z \in \mathrm{Hilb}^2(\mathbb{CP}^2)$ is a single point with multiplicity 2. Thus $\mathcal{C}^{(2)}$ is a $\mathbb{CP}^{12}$ fibration over $\mathrm{Hilb}^2(\mathbb{CP}^2)$. It follows that $e(\mathcal{C}^{(2)}) = 9 \cdot 13 = 117$. Then substituting the known quantities in Eq. (33.32), we get $n^1_4 = 231$.

The formulas Eq. (33.28), Eq. (33.30), and Eq. (33.32) which compute $n^g_{[C]}$, $n^{g-1}_{[C]}$, and $n^{g-2}_{[C]}$ can be extended to compute more general $n^{g-\delta}_{[C]}$. We will give a heuristic mathematical argument for this generalization in Eq. (33.35) below which differs from the arguments given previously in the literature.

Our formulas for the $n^{g-\delta}_{[C]}$ can also be arrived at by an extension of the familiar techniques used in the K3 situation. Let X be a Calabi–Yau threefold, and imagine that we have a complete family $\mathcal{C}$ of Riemann surfaces of genus g in X that is parametrized by a 1-dimensional base $\mathcal{M}$. In addition, suppose that the only singular curves in the family have a single node for their singularity. We assume that the general curve is smooth, so that there are only finitely many of these singular curves. We will usually think of $\mathcal{C}$ as a fibration over $\mathcal{M}$ with singular fibers, as depicted in Fig. 1.

Each of the singular curves is the image of a holomorphic map from a Riemann surface of genus $g-1$, as shown in Fig. 1. Each of these maps is the

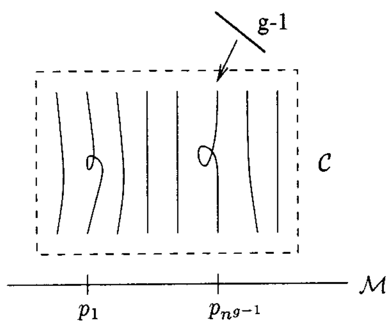


FIGURE 1. A genus g fibration with degenerate fibers

normalization map, which glues together two points of the Riemann surface to form the node. There are $n^{g-1}_{[C]}$ of these maps by the Gromov–Witten interpretation of $n^{g-1}_{[C]}$. So there are $n^{g-1}_{[C]}$ singular fibers in the family $\mathcal{C}$. We denote the discriminant locus of the family by $\{p_1, \ldots, p_{n^{g-1}}\} \subset \mathcal{M}$.

The Euler characteristic of a generic fiber is $2-2g$, and the Euler characteristic of a singular fiber is $3-2g$, since a singular fiber is obtained by pinching a one-cycle of a smooth genus g Riemann surface. Note that by Eq. (33.28) we have $e(\mathcal{M}) = -n^g_{[C]}$. We then compute

$$\begin{aligned} e(\mathcal{C}) &= e(\mathcal{M} - \{p_1, \ldots, p_{n^{g-1}}\})(2-2g) + n^{g-1}_{[C]}(3-2g) \\ &= (-n^g_{[C]} - n^{g-1}_{[C]})(2-2g) + n^{g-1}_{[C]}(3-2g) \\ &= (2g-2)n^g_{[C]} + n^{g-1}_{[C]}, \end{aligned}$$

in complete argreement with Eq. (33.30). This argument can be modified to apply to higher dimensional families $\mathcal{M}$, again reproducing Eq. (33.30).

Similarly, the Euler characteristic of $\mathcal{C}^{(p)}$ can be computed by fibering it over $\mathcal{M}$. We also need to compute the Euler characteristic of the Hilbert scheme of a nodal curve to complete the computation. We arrive at a simple formula relating $e(\mathcal{M})$, various $e(\mathcal{C}^{(k)})$, and $n^{g-p}_{[C]}$. This formula is seen to be equivalent to the formulae for the $n^{g-\delta}_{[C]}$ which has appeared in the literature and was alluded to above.

There is much evidence that justifies the computations which we have described above, but there is much work that needs to be done to make this mathematically rigorous. We also do not have a complete proposal for a direct mathematical definition of the $n^g_{[C]}$ (only the indirect definition Eq. (33.6)). The goals are to give a complete and intrinsic mathematical definition of the n^g, and then to prove that Eq. (33.6) holds using this definition. We concentrate on the first goal in the remainder of this chapter.

We begin our discussion by casting our problem in the context of deformation theory. Let us recall one way to approach Gromov–Witten theory. We deform the complex structure of the Calabi–Yau manifold X to an almost complex structure, and see which holomorphic curves deform to pseudo-holomorphic curves (i.e., curves satisfying the perturbed Cauchy-Riemann equation) in the almost complex structure. We expect finitely many holomorphic curves to deform to curves that are pseudo-holomorphic relative to the nearby almost complex structure.

It is necessary to consider almost complex structures, since there are many situations in which there exist infinitely many curves in a given homology class for a *general* complex structure. The case of $\mathcal{O}(-3)$ over $\mathbb{CP}^2$ is a perfect example, as the $\mathbb{CP}^2$ is preserved under any small deformation of complex structure, and there are infinitely many curves of any degree $d > 0$.

If all curves are isolated, we just count the curves and address the issue of multiple covers. Since this ideal situation does not always occur in complex geometry, we must pass to almost complex geometry in the hopes of achieving an analogous situation. Unfortunately, it is not known whether or not this good situation can be realized, not even for $g = 0$.

A deformation is depicted in Fig. 2 for $d = 4$ in $\mathcal{O}(-3)$. We see n^g_4 pseudo-holomorphic curves of genus $g = 0,1,2,3$ approaching degree 4 curves with $\delta = 3-g$ nodes as the almost complex structure approaches the complex structure that we started with.

Note that the family of all $d = 4$ curves contains all of the information that is needed to compute the n^g_4 for $0 \le g \le 3$. We view this as analogous to the family of all $d = 4$ stable maps, which contains all of the information needed to compute the Gromov–Witten invariants N^g_4, $0 \le g \le 3$. This latter computation is carried out (at least in principle) by formulating a tangent-obstruction theory as in Ch. 26 and computing the associated virtual fundamental class. We expect there to be a similar tangent-obstruction

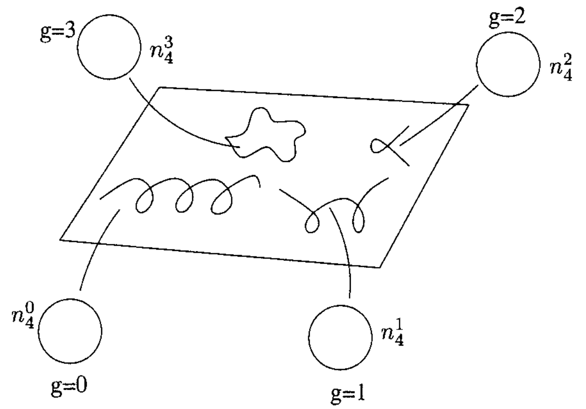


FIGURE 2. Curves of varying genus mapping to degree 3 plane curves

theory associated to the family of curves and sheaves on these curves, whose virtual fundamental class will compute the sought-after n^g_β. These numbers are expected to agree with the counts of pseudo-holomorphic curves discussed above. We will need a different tangent-obstruction theory for each genus.

We will not describe here the precise meaning of a tangent-obstruction theory or any of the definitions of the virtual fundamental class, but will make some general comments. The "tangent" essentially refers to the tangent space to moduli, and contains needed information about infinitesimal deformations. The "obstructions" reflect the fact that there may be obstructions to extending infinitesimal deformations to higher order. In nice situations, the obstruction theory defines an obstruction bundle over the moduli space in question, and the virtual fundamental class is the Euler class of that obstruction bundle. A deformation of complex structure or a C^∞ deformation of almost complex structure yields a section of the obstruction bundle, whose zero-locus coincides with the limits of the pseudo-holomorphic curves that we have described above. Since the zero-locus of a

regular section of a bundle represents its Euler class in general, we see that the virtual fundamental class captures the information that it was designed to compute.

We now consider the simplest case. Let C be a curve in a Calabi–Yau threefold X and let $N = N_{C/X}$ be the normal bundle of C in X. There is a tangent-obstruction theory on the moduli space $\mathcal{M}$ of C.

The deformation space of C is well known to be $H^0(C,N)$. One way to understand this is to think of a deformation of C as a "motion" of C inside X. As C moves inside X infinitesimally, we can think of each point $p \in C$ as moving infinitesimally in X, giving rise to a tangent vector to X at p. Since C is not deformed by vector fields that point in the direction of C itself, we must mod out by these fields and conclude that infinitesimal deformations are in fact described by normal vector fields, as claimed.

The obstruction space turns out to be $H^1(C,N)$. In good cases, we can understand the obstruction space more intrinsically. First note that $H^0(C,N) = T_{\mathcal{M},C}$, the tangent space of $\mathcal{M}$ at C. If $\mathcal{M}$ is smooth, then $H^0(C,N)$ globalizes to give the tangent bundle of $\mathcal{M}$. We now assert that $H^1(C,N)$ globalizes to the cotangent bundle $\Omega^1_{\mathcal{M}}$. To see this, note that there is a perfect pairing

$$H^0(C,N) \otimes H^1(C,N) \to H^1(C,\Lambda^2 N) = H^1(C,K_C) = \mathbf{C}, \tag{33.33}$$

where the first equality follows from the fact that X is Calabi–Yau. This perfect pairing between $H^0(C,N)$ and $H^1(C,N)$ sets up a global duality between the tangent and obstruction bundles, showing that $H^1(C,N_C)$ globalizes to $\Omega^1_{\mathcal{M}}$ as claimed. The virtual fundamental class is then given by $e(\Omega^1_{\mathcal{M}}) = (-1)^{\dim \mathcal{M}} e(\mathcal{M}) = n^g_{[C]}$. Note that this coincides with Eq. (33.28).

Continuing to assume that $\mathcal{M}$ is smooth, we can see directly how a deformation gives rise to a section of the obstruction bundle. Suppose that we have a family of curves in X. We want to deform the complex structure of X and see if these curves stay holomorphic. Recall that an infinitesimal deformation of the complex structure on X is given by an element $\rho \in H^1(T_X)$, the Kodaira-Spencer class. The natural map $T_X \to N = T_X|_C/T_C$ induces a map $H^1(T_X) \to H^1(C,N)$, and the image of ρ gives an element of $H^1(C,N)$, depending on C. As C varies, we get a section s of the obstruction bundle, and computation shows that $C \subset X$ stays holomorphic under this deformation (at least to first-order) if and only if $s(C) = 0$. If there are finitely many C that remain holomorphic, then s has finitely many zeros, and the zero-locus of s represents the Euler class of the obstruction bundle.

In the case of $\mathcal{O}(-3)$, the obstruction bundle has no sections. Nevertheless, we can compute its Euler class. For degree d curves, we see from Eq. (33.9) and (33.28) that this Euler class is $(-1)^{(d^2+3d)/2}(d+1)(d+2)/2$.

The desired tangent-obstruction theory is more subtle in general, and we will not attempt to describe it here. Instead, we describe a related tangent-obstruction theory, and show how it should yield Eq. (33.28), Eq. (33.30), Eq. (33.32), and their generalizations.

Let C be a curve of arithmetic genus g in a Calabi–Yau threefold X. Let $\mathcal{M}$ be the connected component of the Hilbert scheme of C in X. To compute the contribution of $\mathcal{M}$ to $n^{g-k}_{[C]}$ we will need to consider certain sheaves on the curves $C \in \mathcal{M}$. We study the ideal sheaves $I_Z \subset \mathcal{O}_C$ of length k subschemes $Z \in \mathrm{Hilb}^k(C)$. These can be viewed as sheaves of $\mathcal{O}_X$-modules if one wishes to view them in the context of homological mirror symmetry. These sheaves are parametrized by the relative Hilbert scheme $\mathcal{C}^{(k)}$ that we have already considered.

If $\mathcal{C}^{(k)}$ is smooth, there is a natural tangent-obstruction theory on $\mathcal{C}^{(k)}$ where the obstruction space is the cotangent space, just as we were able to deduce from the perfect pairing Eq. (33.33) in the special case considered above. The virtual fundamental class has degree equal to the Euler class of the cotangent bundle, which is just the Euler characteristic of $\mathcal{C}^{(k)}$ with a sign determined by the parity of the dimension of $\mathcal{C}^{(k)}$.

As we mentioned above, this tangent-obstruction theory is *not* the one that we ultimately want, but it is enough to calculate what we need. Suppose that we can deform the complex structure (or almost complex structure) so that $\mathcal{M}$ splits up into $n^{g-\delta}_{[C]}$ (pseudo-)holomorphic curves of genus $g-\delta$, for each $\delta = 0, \ldots, k$. Consider one of these curves with the deformed structure, say $C_{g-\delta}$ of genus $g-\delta$. It is not difficult to see that the family of sheaves I_Z for $Z \in \mathcal{C}^{(k)}$ deforms to the union over all δ and curves $C_{g-\delta}$ of the families of ideal sheaves $I_{Z_{k-\delta}}$, where

$$Z_{k-\delta} \in \mathrm{Hilb}^{k-\delta}(C_{g-\delta}).$$

A weaker statement, which is easier to check and suggests the above assertion, is the equality of Chern characters $\mathrm{ch}(I_Z) = \mathrm{ch}(I_{Z_{k-\delta}})$.

The $I_{Z_{k-\delta}}$ are parametrized by the $(k-\delta)^{th}$ symmetric product of $C_{g-\delta}$. This symmetric product has Euler characteristic

$$\binom{2-2g+k+\delta-1}{k-\delta}. \tag{33.34}$$

It turns out that the virtual fundamental classes considered here are invariant under deformation. Equating the virtual fundamental classes before and after the deformation, using Eq. (33.34), and noting that the $(k-\delta)^{th}$ symmetric product of $C_{g-\delta}$ has dimension $k-\delta$, we obtain the formula

$$(-1)^{\dim \mathcal{C}^{(k)}} e(\mathcal{C}^{(k)}) = \sum_{\delta=0}^{k} (-1)^{k-\delta} n^{g-\delta}_{[C]} \binom{2-2g+k+\delta-1}{k-\delta}. \tag{33.35}$$

Note that for $k = 0, 1, 2$, Eq. (33.35) becomes respectively Eq. (33.28), Eq. (33.30), and Eq. (33.32).

Aside from Example 33.0.3, we have been silent about higher rank bundles that arise in D-branes. We give a general context to illustrate some of the subtleties that can arise. Other interesting subtleties will be discussed in Ch. 34.

Consider a contractible curve $C \simeq \mathbb{CP}^1$ in a Calabi–Yau threefold

$$\begin{array}{ccc} C \simeq \mathbb{CP}^1 & \subset & X \\ \downarrow & & \downarrow \pi \\ p & \in & Y \end{array} \tag{33.36}$$

where $\pi : X \to Y$ is the map that contracts C to a point p. The contributions of multiples of C to F_g were computed for all g. When Eq. (33.6) is then used to compute the $n^g_d = n^g_{d[C]}$ for all d and g, the result is that the n^g_d are always integers, as expected. In particular, it is found that $n^g_{d[C]} = 0$ if $g > 0$ or if $d > 6$. However, nonzero n^0_d can and do arise, and if $d > 1$ these invariants correspond to higher rank bundles (actually sheaves). These sheaves define scheme structures on C with multiplicity d, and n^0_d counts the number of isolated curves that each of these schemes splits up into under a general deformation of complex structure. The integer d can take any value up to and including the *length* ℓ of C, which is defined to be the multiplicity of C in the scheme $\pi^{-1}(p)$. In particular, the multiplicity ℓ scheme alluded to above is just $\pi^{-1}(p)$ itself.

We illustrate with an example.

EXAMPLE 33.0.6. *Suppose that* $N_{C/X} = \mathcal{O} \oplus \mathcal{O}(-2)$. *We can be very explicit in this case.*

We can construct X locally as the union of two $\mathbb{C}^3$ patches, $X = \mathbb{C}^3 \cup \mathbb{C}^3$. We choose coordinates (x, y_1, y_2) for the first $\mathbb{C}^3$, and (w, z_1, z_2) for the second $\mathbb{C}^3$. The two patches are glued together via

$$z_1 = x^2 y_1 + x y_2^n, \qquad z_2 = y_2, \qquad w = x^{-1}. \tag{33.37}$$

The curve C is obtained by gluing together the curve $\{y_1 = y_2 = 0\}$ in the first patch and the curve $\{z_1 = z_2 = 0\}$ in the second patch, and it is immediate to see that $C \simeq \mathbb{CP}^1$.

The integer n occurs in Eq. (33.37) because the normal bundle $N = \mathcal{O} \oplus \mathcal{O}(-2)$ is not enough information to fix $C \subset X$ locally near C. The extra data can be completely described by this single integer $n \geq 2$. If n were taken to be 1 in Eq. (33.37), then the normal bundle of C in X would have been $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$.

Let us try to deform C inside X. To first order, the deformation must be in the $\mathcal{O}$ direction in the normal bundle, as $\mathcal{O}(-2)$ has no global holomorphic sections. A glance at the form of Eq. (33.37) shows that this direction corresponds to the variable y_2 in the first patch, and z_2 in the second patch. We can choose the deformation to be given in the first patch by

$$y_1 = 0, \qquad y_2 = \epsilon.$$

¿From Eq. (33.37), we get in the second patch the deformation

$$z_1 = \epsilon^n x, \qquad z_2 = \epsilon.$$

Since $x = w^{-1}$ is not holomorphic in the (w, z_1, z_2) patch, we are forced to formally require $\epsilon^n = 0$. This says that our moduli space is therefore a non-reduced point, described by a single variable ϵ with $\epsilon^n = 0$. The multiplicity is $n^0_{[C]} = n$.

To see the n deformations, we can for instance deform the complex structure of X, replacing Eq. (33.37) with, for instance,

$$z_1 = x^2 y_1 + x y_2^n + t x y_2, \qquad z_2 = y_2, \qquad w = x^{-1} \tag{33.38}$$

depending on a deformation parameter t. Then C defined by $y_1 = y_2 = 0$ and $z_1 = z_2 = 0$ remains holomorphic for all t, but there are $n-1$ additional curves defined by

$$y_1 = 0,\ y_2 = (-t)^{1/(n-1)}, \qquad z_1 = 0,\ z_2 = (-t)^{1/(n-1)}.$$

Hence C deforms to n curves, as claimed.

In general, with notation as in Eq. (33.36), it is known that $p \in Y$ is a *cDV* singularity, which means that a general surface section $S \subset Y$ containing p has a rational double point at p. These generic surface singularity types have been classified. The main result is that these are determined by length ℓ as shown in the following table.

ℓ	S
1	A_1
2	D_4
3	E_6
4	E_7
5	E_8
6	E_8

In general, one gets thickenings of C up to multiplicity ℓ, i.e. multiplicity d schemes C_d for each $d \leq \ell$. It can be checked that the Chern character $\mathrm{ch}(\mathcal{O}_{C_d})$ of $\mathcal{O}_{C_d}$ coincides with the Chern character of a degree d multiple cover of the original curve C. When the complex structure of X is deformed, the result is that C_d splits up into n^0_d curves in the class $d[C]$.

We anticipate that the GV invariants can be defined more generally in terms of sheaves. It is straightforward to define a natural tangent-obstruction theory in the context of the immediately preceding discussion for which the virtual fundamental class has degree n^0_d, as required. Work is underway to extend the applicability of this construction.

CHAPTER 34

Multiple Covers, Integrality, and Gopakumar–Vafa Invariants

In this chapter, we discuss the integrality conjectures of Gopakumar and Vafa, and state a generalized integrality conjecture for Gromov–Witten invariants of threefolds.[1]

Let X be a Calabi–Yau threefold. Consider the genus 0 Gromov–Witten potential of non-constant maps:

$$\tilde{F}^0(q) = \sum_{0 \neq \beta \in H_2(X,\mathbb{Z})} N^0_\beta q^\beta, \tag{34.1}$$

where N^0_β is the genus 0 Gromov–Witten invariant of X in the curve class β. The potential $\tilde{F}_0$ differs from the full genus 0 potential by the constant map contributions ($\beta = 0$). Define the invariants n^0_β for each nonzero curve class $\beta \in H_2(X, \mathbb{Z})$ by the formula:

$$\tilde{F}^0(q) = \sum_{\beta \neq 0} n^0_\beta \sum_{d>0} \frac{1}{d^3} q^{d\beta}. \tag{34.2}$$

Eq. (34.2) uniquely determines the invariants n^0_β from $\tilde{F}^0(q)$. If X is an *ideal* Calabi–Yau threefold, the invariants n^0_β enumerate rational curves in X of class β — in this case, Eq. (34.2) relates Gromov–Witten theory to enumerative geometry via the Aspinwall–Morrison formula (see Section 29.1.2).

Let X be a quintic Calabi–Yau threefold. By the existence of Vainsencher's curves, X is not ideal (see Sec. 29.1.2). The Gromov–Witten invariants of X are related to hypergeometric series by the Mirror conjecture. The invariants n^0_β may then be recursively determined from Eq. (34.2). In every case computed, the invariant n^0_β has been found to be an integer. The integrality here is a subtle issue — a priori, only the rationality of n^0_β is expected once the ideal conditions on rational curves in X fail.

[1]Readers should refer to Part 4 for notation and details regarding Gromov–Witten theory.

As integrality holds in the ideal case and (experimentally) holds in the quintic case, it is very natural to conjecture the integrality of n^0_β for all Calabi–Yau threefolds.

CONJECTURE 34.0.1. *The invariants n^0_β are integers for all Calabi–Yau threefolds X and curve classes $\beta \neq 0$.*

Assuming Conjecture 34.0.1 is true, it is natural to seek a geometric interpretation of n^0_β. A striking proposal to answer this question has been made by Gopakumar and Vafa.

Let X be a Calabi–Yau threefold. We will now consider the Gromov–Witten invariants of X in all genera:

$$\tilde{F}(q,\lambda) = \sum_{g\geq 0} \lambda^{2g-2} \tilde{F}^g(q,\lambda),$$

$$\tilde{F}^g(q,\lambda) = \sum_{0\neq\beta\in H_2(X,\mathbb{Z})} N^g_\beta q^\beta,$$

where N^g_β is the genus g Gromov–Witten invariant of X in the curve class β. Gopakumar and Vafa first proposed the following generalization of Eq. (34.2).

DEFINITION 34.0.7. *Define the* Gopakumar–Vafa invariants n^g_β *for each genus g and each nonzero curve class $\beta \in H_2(X,\mathbb{Z})$ by the formula*

$$\tilde{F}(q,\lambda) = \sum_{\substack{g\\ \beta\neq 0}} n^g_\beta \lambda^{2g-2} \sum_{d>0} \frac{1}{d}\left(\frac{\sin(d\lambda/2)}{\lambda/2}\right)^{2g-2} q^{d\beta}. \tag{34.3}$$

CONJECTURE 34.0.2 (Gopakumar–Vafa). *The invariants n^g_β are integers for all Calabi-Yau threefolds X, genera $g \geq 0$, and curve classes $\beta \neq 0$.*

The invariants n^g_β arise as BPS state counts in a study of Type IIA string theory on X via M-theory. In the context of physics, the integrality of n^g_β is shown by this identification.

By Eq. (34.3), a single genus 0 BPS state counted by n^0_β contributes

$$\frac{\lambda^{-2}}{d}\left(\frac{\lambda/2}{\sin(d\lambda/2)}\right)^2 q^{d\beta}$$

to the Gromov–Witten theory of X. This is precisely the Gromov–Witten contribution obtained from a rigid curve $\mathbb{P}^1 \subset X$ of class β by Theorem 27.7.1. However, Eq. (34.3) is not a multiple cover formula in Gromov–Witten theory for arbitrary rigid curves $C \subset X$ — the BPS states $n^{g(C)}_\beta$ affect Gromov–Witten invariants in degrees and genera where multiple cover contributions of C vanish. There should be a relationship between Eq. (34.3) and multiple cover contributions.

Gopakumar and Vafa further proposed a mathematical construction of n^g_β using a moduli space of sheaves on X. The invariants n^g_β arise as multiplicities of special representations of $\mathfrak{sl}_2$ in the cohomology of the moduli space of sheaves. A full mathematical development of this approach has not yet been completed. However, there is compelling evidence in several cases.

We will explain here a conjecture for integrality constraints for all threefolds that generalizes Conjecture 34.0.2.

34.1. The Gromov–Witten Theory of Threefolds

Let X be a non-singular, projective, algebraic threefold. As before, let $T_0,\ldots,T_m$ be a basis of $H^*(X,\mathbb{Z})$ modulo torsion. For simplicity we will assume the classes T_i are even — no essential difference occurs in the general case. Let T_0 denote the fundamental class, let $T_1,\ldots,T_p$ denote the (real) degree 2 classes, and let $T_{p+1},\ldots,T_m$ denote the classes of degree greater than 2. Let $t_0,\ldots,t_m$ be a set of corresponding variables. The Gromov–Witten potential of X may be written in the form

$$F(t,\lambda) = F^0_{\beta=0} + F^1_{\beta=0} + \sum_{g\geq 2} F^g_{\beta=0} + \sum_{g\geq 0}\sum_{\beta\neq 0} F^g_\beta. \tag{34.4}$$

The potential $\tilde{F}(t,\lambda)$ will denote the non-constant map contributions:

$$\tilde{F}(t,\lambda) = \sum_{g\geq 0}\sum_{\beta\neq 0} F^g_\beta.$$

34.1.1. $\beta = 0$ Contributions. The first three terms in Eq. (34.4) are the contributions of the constant maps. The genus 0 constant contribution records the classical intersection theory of X:

$$F^0_{\beta=0} = \lambda^{-2} \sum_{0\leq i_1,i_2,i_3\leq m} \frac{t_{i_1}t_{i_2}t_{i_3}}{3!}\int_X T_{i_1}\cup T_{i_2}\cup T_{i_3}.$$

The genus 1 constant contribution is obtained from a virtual class calculation:

$$F^1_{\beta=0} = \sum_{i=1}^p t_i \langle T_i\rangle_{g=1,\beta=0} = -\sum_{i=1}^p \frac{t_i}{24}\int_X T_i \cup c_2(X).$$

The Gromov–Witten invariants $\langle T_i \rangle_{g=1,\beta=0}$ can be calculated by identifying the virtual class of $\overline{\mathcal{M}}_{1,1}(X,0)$ with

$$\mathbf{e}(T_X \otimes \mathbb{E}^\vee) \cap [X \times \overline{\mathcal{M}}_{1,1}],$$

using the ideas of Sec. 26.1.2. Similarly, the genus $g \geq 2$ contributions are

$$F^g_{\beta=0} = \langle 1 \rangle_{g,\beta=0} = (-1)^g \frac{\lambda^{2g-2}}{2} \int_X (c_3(X) - c_1(X) \cup c_2(X)) \cdot \int_{\overline{\mathcal{M}}_g} \lambda^3_{g-1},$$

obtained from the virtual class equation:

$$[\overline{\mathcal{M}}_g(X,0)]^{\mathrm{vir}} = \mathbf{e}(T_X \otimes \mathbb{E}^\vee) \cap [X \times \overline{\mathcal{M}}_g].$$

The Hodge integrals that arise here have been computed:

$$\int_{\overline{\mathcal{M}}_g} \lambda^3_{g-1} = \frac{|B_{2g}|}{2g} \frac{|B_{2g-2}|}{2g-2} \frac{1}{(2g-2)!},$$

where B_{2g} and B_{2g-2} are Bernoulli numbers. The $\beta = 0$ contributions to F are therefore completely understood.

34.1.2. $\beta \neq 0$ Contributions. Since the virtual dimension of the moduli space $\overline{\mathcal{M}}_g(X,\beta)$ is

$$\int_\beta c_1(X) + 3g - 3 + 3 - 3g = \int_\beta c_1(X),$$

the classes β satisfying $\int_\beta c_1(X) < 0$ do not contribute to the potential X. Therefore, the fourth term in Eq. (34.4) may be divided into two sums:

$$\begin{aligned} \sum_{g\geq 0}\sum_{\beta\neq 0} F^g_\beta &= \sum_{g\geq 0} \sum_{\beta,\ \int_{\beta\neq 0} c_1(X)=0} F^g_\beta \\ &+ \sum_{g \geq 0} \sum_{\beta,\ \int_\beta c_1(X)>0} F^g_\beta. \end{aligned}$$

In the case $\beta \neq 0$, we will write the series F^g_β in the form

$$F^g_\beta(t,\lambda) = \lambda^{2g-2} q^\beta \sum_{n\geq 0} \frac{1}{n!} \sum_{p+1 \leq i_1,\ldots,i_n \leq m} t_{i_1} \cdots t_{i_n} \langle T_{i_1} \cdots T_{i_n} \rangle_{g,\beta}.$$

The degree 2 variables $t_1, \ldots, t_p$ have been formally suppressed in q:

$$q^\beta = q_1^{\int_\beta T_1} \cdots q_p^{\int_\beta T_p}, \quad q_i = e^{t_i}.$$

34.2. Proposal

Let X be a threefold as in Sec. 34.1. We will define new invariants

$$n^g_\beta(T_{i_1}, \ldots, T_{i_n})$$

for every genus g, curve class $\beta \neq 0$, and vector of classes $(T_{i_1}, \ldots, T_{i_n})$. Our primary interest will be in the case where the following conditions hold:

(i) $\deg(T_{i_k}) > 2$ for all k.
(ii) $n + \int_\beta c_1(X) = \sum_{k=1}^n \deg(T_{i_k})$.

The invariants will be defined to satisfy the divisor equation (which allows for the extraction of degree 2 classes T_i). Also, $n^g_\beta(T_{i_1}, \ldots, T_{i_n})$ will be defined to vanish when condition (ii) is violated.

The invariants $n^g_\beta(T_{i_1}, \ldots, T_{i_n})$ are defined by the following equation:

$$\begin{aligned} \tilde{F}(t,\lambda) &= \sum_{g\geq 0} \sum_{\beta,\ \int_{\beta\neq 0} c_1(X)=0} n^g_\beta \lambda^{2g-2} \sum_{d>0} \frac{1}{d} \left(\frac{\sin(d\lambda/2)}{\lambda/2}\right)^{2g-2} q^{d\beta} \\ &+ \sum_{g\geq 0} \sum_{\beta,\ \int_\beta c_1(X)>0} \sum_{n\geq 0} \frac{1}{n!} \sum_{p+1\leq i_1,\ldots,i_n\leq m} t_{i_1} \cdots t_{i_n} \\ &\cdot n^g_\beta(T_{i_1} \cdots T_{i_n}) \lambda^{2g-2} \left(\frac{\sin(\lambda/2)}{\lambda/2}\right)^{2g-2+\int_\beta c_1(X)} q^\beta. \end{aligned}$$

CONJECTURE 34.2.1 (Pandharipande). *The invariants $n^g_\beta(T_{i_1}, \ldots, T_{i_n})$ are integers for all threefolds X, genera $g \geq 0$, curve classes $\beta \neq 0$, and integral cohomology classes $(T_{i_1}, \ldots, T_{i_k})$.*

The motivation for Conjecture 34.2.1 is drawn from the Gopakumar–Vafa invariants in the Calabi–Yau case and calculations of degenerate contributions in the Fano case.

34.2.1. Case $\int_\beta c_1(X) = 0$. If X is Calabi–Yau, this case holds for all classes $\beta \neq 0$. Degree d multiple covers of β contribute to the Gromov–Witten invariant of $d\beta$. While the Gopakumar–Vafa formula is not exactly a multiple cover formula, the integrality constraint is believed to be "equivalent" to a suitably defined cover formula. We take the Gopakumar–Vafa formula to define n^g_β for curve classes β satisfying $\int_\beta c_1(X) = 0$ in arbitrary threefolds.

34.2.2. Case $\int_\beta c_1(X) > 0$. If X is Fano, this case holds for all classes $\beta \neq 0$. The moduli space of maps $\overline{\mathcal{M}}_g(X,\beta)$ has positive virtual dimension. The Gromov–Witten invariants $N^g_\beta(T_{i_1},\dots,T_{i_n})$ of X then depend upon a vector of cohomology classes

$$\gamma = (T_{i_1},\dots,T_{i_n}).$$

Let $Y_{i_k} \subset X$ be general topological cycles dual to the classes T_{i_k}. Integers in the Gromov–Witten theory of X can be predicted through degenerate contributions.

Let us assume we are in an ideal situation with respect to the moduli spaces of maps to X. Let $\mathcal{M}^{\mathrm{Bir}}_g(X,\beta)$ denote the moduli space of birational maps from non-singular genus g domain curves. We assume first:

(i) The spaces $\mathcal{M}^{\mathrm{Bir}}_g(X,\beta)$ are generically reduced and of the expected dimension for all $h \leq g$.

There is then an enumerative integer $n^g_\beta(\gamma)$ defined to equal the number of genus g maps of class β with non-singular domains meeting all the cycles Y_i. However, $n^g_\beta(\gamma)$ will not equal $N^g_\beta(\gamma)$. The difference arises from the following observation. For each $h < g$, there are $n^h_\beta(\gamma)$ maps with non-singular genus h domains of class β satisfying the required incidence conditions. The Gromov–Witten invariant $N^g_\beta(\gamma)$ receives a degenerate contribution from each of these lower genus solutions (via reducible genus g maps factoring through covers of the lower genus curves). As the genus h solution represents the class β, the covers must be of degree 1.

Dimension counts show maps multiple onto their image and maps with reducible images are not *expected* to contribute to $N^g_\beta(\gamma)$. This is the second ideal assumption:

(ii) Maps in $\overline{\mathcal{M}}_g(X,\beta)$ multiple onto their image or with reducible image do not satisfy incidence conditions to the cycles Y_i.

Let $C \subset X$ be a non-singular, genus h curve of class β satisfying incidence conditions to the cycles Y_i. Assume further that C is infinitesimally rigid with respect to these incidence conditions. The contribution $C_h(r,X,\beta)$ of C to the Gromov–Witten invariant $N^{h+r}_\beta(\gamma)$ is then well defined: it is an integral over the moduli space $\overline{\mathcal{M}}_{h+r}(C,[C])$. This contribution is easily seen to be independent of γ. The final ideal assumption is:

(iii) For all $h < g$, the solution maps counted by $n^g_\beta(\gamma)$ are non-singular embeddings infinitesimally rigid with respect to the incidence conditions.

The ideal relation between Gromov–Witten theory and the enumerative invariants is

$$N^g_\beta(\gamma) = \sum_{h=0}^{g} C_h(g-h,X,\beta)\, n^h_\beta(\gamma). \tag{34.5}$$

The validity of the relation Eq. (34.5) for $N^g_\beta(\gamma)$ requires assumptions (i), (ii), and (iii) for g, β, and γ.

The contributions $C_h(r,X,\beta)$ are calculated:

$$\sum_{r=0}^{\infty} C_h(r,X,\beta)\lambda^{2r} = \left(\frac{\sin(\lambda/2)}{\lambda/2}\right)^{2g-2+\int_\beta c_1(X)}. \tag{34.6}$$

Formula (34.6) together with Eq. (34.5) provides the motivation for the Fano case of Conjecture 34.2.1. While most threefolds X will fail to satisfy the ideal conditions (i)–(iii), the integrality constraints are conjectured to hold in all cases.

If Conjecture 34.2.1 is true, perhaps the recasting of Gromov–Witten theory in terms of the cohomology of the moduli of sheaves can be undertaken for all threefolds X.

34.3. Consequences for Algebraic Surfaces

Instead of reviewing the evidence for these integrality conjectures, we present here some peculiar consequences of the strongest one, Conjecture 34.2.1.

The motivation for the Fano case of Conjecture 34.2.1 came from studying threefolds X with ideal properties with respect to the moduli of maps. As the statement of the conjecture applies to all X, it is natural to examine the consequences for threefolds at the opposite extreme.

Very ill behaved spaces of maps can be found in product varieties. Let X be a threefold of the form

$$X = S \times E,$$

where S is an algebraic surface and E is an elliptic curve. Let $\beta \in H_2(X,\mathbb{Z})$ be the vertical curve class:

$$\beta = [L] \times [P],$$

where

$$[L] \in H_2(S, \mathbb{Z}), \ [P] \in H_0(E, \mathbb{Z})$$

are curve and point classes respectively. The moduli of maps to X is a product:

$$\overline{\mathcal{M}}_g(X, \beta) = \overline{\mathcal{M}}_g(S, [L]) \times E.$$

The virtual dimension of $\overline{\mathcal{M}}_g(X, \beta)$ is $\int_{[L]} c_1(S)$, while the virtual dimension of $\overline{\mathcal{M}}_g(S, [L])$ is $\int_{[L]} c_1(S) + g - 1$. Therefore $\overline{\mathcal{M}}_g(X, \beta)$ will in general have excess dimension. In fact the virtual class of the moduli of maps to X is determined by

$$[\overline{\mathcal{M}}_g(X, \beta)]^{\text{vir}} = (-1)^g \lambda_g \cap [\overline{\mathcal{M}}_g(S, [L])]^{\text{vir}} \times [E]. \tag{34.7}$$

Eq. (34.7) relates the Gromov–Witten theory of the vertical class β to Hodge integrals over the moduli space of maps to S.

Let $Y \in H^4(S, \mathbb{Z})$ and $P \in H^2(E, \mathbb{Z})$ denote the point classes in S and E respectively. Let $n = \int_{[L]} c_1(S) - 1$. By the virtual class relation (34.7), we find:

$$\langle Y \cup P, Y, \ldots, Y \rangle^X_{g,\beta} = \int_{[\overline{\mathcal{M}}_{g,n}(S,[L])]^{\text{vir}}} (-1)^g \lambda_g \cdot \prod_{i=1}^{n} \text{ev}_i^*(Y).$$

If $g = 1$, then $\lambda_1 \in H^2(\overline{\mathcal{M}}_{1,1})$ satisfies the well-known relation

$$\lambda_1 = \frac{1}{12} \Delta_0. \tag{34.8}$$

Then, by the splitting and divisor axioms of Gromov–Witten theory,

$$\langle Y \cup P, Y, \ldots, Y \rangle^X_{1,\beta} = -\frac{1}{24} ([L] \cdot [L]) \ \langle Y, \ldots, Y \rangle^S_{0,[L]}. \tag{34.9}$$

We have now expressed $g = 1$ invariants of $X = S \times E$ in terms of $g = 0$ invariants of S.

Conjecture 34.2.1 therefore implies integrality conditions for the Gromov–Witten theory of *every* algebraic surface. Let $[L] \in H_2(S, \mathbb{Z})$ be a curve class. Let $N^g(L)$ denote the genus g Gromov–Witten invariant of S:

$$N^g(L) = \langle Y, \ldots, Y \rangle^S_{g,[L]}.$$

Let $g(L)$ denote the genus of a non-singular curve in class $[L]$ determined by the adjunction formula

$$2g(L) - 2 = \int_{[L]} [L] - c_1(S).$$

Conjecture 34.2.1 and Eq. (34.9) directly yield the following integrality obtained from the invariant $n^1_\beta(Y \cup P, Y, \ldots, Y)$.

CONJECTURE 34.3.1. *Let S be an algebraic surface, and let $[L]$ be a curve class satisfying $\int_{[L]} c_1(S) > 0$. Then*

$$-\frac{1}{12} g(L) N^0(L)$$

is an integer.

Similarly, if $g = 2$, then $\lambda_2 \in H^4(\overline{\mathcal{M}}_2)$ satisfies a boundary relation in $\overline{\mathcal{M}}_2$ analogous to Eq. (34.8). Proceeding as above, the following integrality for $n^2_\beta(Y \cup P, Y, \ldots, Y)$ is predicted by Conjecture 34.2.1.

CONJECTURE 34.3.2. *Let S be an algebraic surface and $[L]$ a curve class satisfying $\int_{[L]} c_1(S) > 0$. Then*

$$\frac{1}{2880} \left(12g(L)^2 + g(L)c(L) - 24g(L)\right) N^0(L) + \frac{1}{240} \chi(S) \ N^1(L)$$
$$+ \frac{1}{240} \sum_{[A]+[B]=[L]} \binom{c(L)-1}{c(A)} ([A] \cdot [B])([B] \cdot [B]) N^1(A) N^0(B)$$

is an integer. Here $\chi(S)$ is the topological Euler characteristic of S, and $c(Z)$ denotes $\int_{[Z]} c_1(S)$.

While these conjectures appear strange, they are true in all the cases studied to date. For example, consider $S = \mathbb{P}^2$. Let $[L]$ be the class of a cubic curve. Then $g(L) = 1$ and $N^0(L) = 12$ (see Theorem 25.1.1). The integrality of Conjecture 34.3.1 is then satisfied.

34.4. Elliptic Rational Surfaces

The best evidence for Conjectures 34.3.1 and 34.3.2 has been obtained by J. Bryan in the case of the elliptic rational surface. We explain his argument here.

Let S be $\mathbb{P}^2$ blown up at nine points. Let F be the anti-canonical class and let E be the class of one of the exceptional divisors. We will consider curve classes

$$L_k = E + kF$$

for $k \geq 0$. Intersection calculations in S yield

$$\int_{[L_k]} c_1(S) = 1, \quad g(L_k) = k.$$

Therefore to prove Conjecture 34.3.1 for the pair S and L_k, we must check the integrality of

$$-\frac{1}{12}kN^0(L_k). \tag{34.10}$$

The generating series

$$F(q) = \sum_{k\geq 0} N^0(L_k)q^k$$

has been proven to be related to a modular form

$$F(q) = \prod_{m=1}^{\infty}(1-q^m)^{-12}.$$

The integrality of (34.10) can be obtained by differentiation:

$$-\sum_{k\geq 0}\frac{1}{12}kN^0(L_k)q^{k-1} = -\frac{1}{12}\frac{d}{dq}F(q) = \prod_{m=1}^{\infty}(1-q^m)^{-13}\cdot\frac{d}{dq}\prod_{m=1}^{\infty}(1-q^m).$$

Bryan has found a proof of Conjecture 34.3.2 for the classes L_k. A more subtle analysis of congruence properties of quasi-modular forms is required; Ramanujan's congruence modulo 5 plays a role.

Conjectures 34.3.1 and 34.3.2 are the first in an infinite sequence of integrality conjectures for λ_g integrals in the genus g Gromov–Witten theory of surfaces. For $g = 1$ and $g = 2$, the Hodge class can be removed to yield conjectures in pure Gromov–Witten theory.

34.5. Outlook

A fundamental understanding of the integer invariants of Gopakumar–Vafa (and their likely extension to general threefolds) should revolutionize the study of Gromov–Witten theory in dimension 3. While the Gromov–Witten invariants are very difficult to compute directly, there is reason to believe the Gopakumar–Vafa invariants may admit more effective approaches. The investigation of these ideas is a very promising future direction for the subject.

CHAPTER 35

Mirror Symmetry at Higher Genus

In this chapter we discuss the application of mirror symmetry to the calculation of higher loop correlation functions in topological string theory. Aspects of the genus 1 amplitude can be most easily exemplified in the case of the two-dimensional target torus T^2. The genus 1 closed string amplitude on T^2 is calculated in Sec. 35.2 and compared with the Ray-Singer torsion in Sec. 35.3. The topological open string annulus amplitude on T^2 is discussed in Sec. 35.4.

The B-model calculation of closed string higher-genus amplitudes on three complex-dimensional Calabi–Yau manifolds uses the holomorphic anomaly equation discussed in Ch. 31. In this chapter the holomorphic anomaly equation is further studied and solved for the example of the quintic in $\mathbb{P}^4$ and the non-compact Calabi–Yau space $\mathcal{O}(-3) \to \mathbb{P}^2$ in Sec. 35.5 and Sec. 35.6.

35.1. General Properties of the Genus 1 Topological Amplitude

We will consider first the closed topological string. At worldsheet genus 1 the simplest amplitude is the vacuum amplitude Eq. (31.14). It can be defined directly in the operator formalism as[1]

$$F_1 = \frac{1}{2}\int_{\mathcal{F}}\frac{d^2\tau}{\tau_2}\mathrm{Tr}_{\mathcal{H}}(-1)^F F_L F_R q^{H_L}\bar{q}^{H_R}, \tag{35.1}$$

where $q = \exp(2\pi i\tau)$, τ is the complex structure of the worldsheet torus and $\mathcal{F}$ its fundamental region. Without the insertion of the left and right fermion number operators $F_{L/R}$, the operator $(-1)^F$, which yields $+/-$ on states with even/odd worldsheet fermion numbers, would project straight to the Ramond–Ramond ground states of the energy operator $H_L + H_R$. The integrand reduces in this case to the Witten index, which for the σ-model just yields the Euler number of the target space.

[1]For later convenience, let us define $F_1 =: \frac{1}{2}\int_{\mathcal{F}}\frac{d^2\tau}{\tau_2^2}I_F$ to be the quantity I_F, which is to be integrated w.r.t. the $SL(2,\mathbb{Z})$ invariant measure over $\mathcal{F}$.

With the insertions of $F_{L/R}$, F_1 has an interesting and calculable dependence on the complex moduli $\sigma, \bar{\sigma}$ of the target space. It factorizes almost, but not quite, into a holomorphic and an anti-holomorphic piece. The failure of factorization is described universally by the so-called holomorphic anomaly equation. It is clear from the path-integral definition of F_1 in the B-model that it is the determinant of an elliptic operator, which varies over the complex moduli space. This situation was studied by Quillen and others and it was noted that the determinant of such operators can be captured by the curvature of a certain bundle (determinant line bundle) on moduli space. At genus 1 this so-called Quillen anomaly and the holomorphic anomaly are identical. This way of evaluating the genus 1 index by calculation of the determinant can be generalized to the open string sector as we will discuss in Sec. 35.4 for the case of the two-dimensional torus target with D-branes.

Using the methods of topological field theory, we can extend the holomorphic anomaly equation to higher genus. As discussed already in Ch. 31 the holomorphic anomaly is in the general case captured by a system of differential equations that determine F_g up to an unknown holomorphic section of a line bundle $\mathcal{L}^{2g-2}$ over the complex moduli space. The fixing of this section may be viewed as fixing the "boundary condition". This holomorphic section is typically specified by a finite set of numbers, which depend on the particular background. To fix it one can use both the space-time interpretation of topological string amplitudes discussed in Ch. 32 and the mathematical definition of F_g in various limits.

In particular there exists a holomorphic limit $F_g^{\mathrm{top}}(\sigma) = \lim_{\bar{\sigma} \to -i\infty} F_g(\sigma, \bar{\sigma})$, with σ fixed. Mirror symmetry allows us to trade the complex structure parameter of the target space σ with the complexified Kähler structure parameter t of the mirror target space. The expansion of $F_g^{\mathrm{top}}(t)$ around the large volume $\mathrm{Im}(t) \to \infty$ has a very interesting interpretation as generating function for the Gromov–Witten invariants of holomorphic maps with various degrees from a genus g worldsheet into the mirror target space, see Eq. (31.19). The same expansions count after a reorganization the integer number of BPS-states, see Eq. (35.54) and Eqs. (35.56)–(35.59). The holomorphic anomaly thus provides recursion relations, which determine infinite series of Gromov–Witten invariants once a few are known. Note that, in this holomorphic limit, $F_g(\sigma)$ is no longer a nice object (it is no longer a section of the bundle of which it was defined to be a section). However, the topological localization formula relating F_g to Gromov–Witten invariants requires taking this limit. Thus, despite the simplification of the computation of Gromov–Witten invariants in precise mathematical terms at higher genus, it is clear that the full F_g, including its non-holomorphic pieces, are the natural objects to study. At any rate, in this section we compute both F_g and its holomorphic limit for several cases.

35.2. The Topological Amplitude F_1 on the Torus

The critical dimension for the topological string, where all F_g can be non-trivial, is 3, see Sec. 24.4, Eq. (24.3) and , in Sec. 31, Eq. (31.3). However for the genus 1 amplitude discussed here and in Sec. 35.4 the corresponding obstruction vanishes in any dimension.

We take advantage of this fact and discuss first in detail the various interesting interpretations of F_1 in the simple case of the target space T^2, where everything can be calculated explicitly. However, whenever it is easy to generalize to the Calabi–Yau case, we will give the corresponding expressions to prepare for the calculation of F_1 on Calabi–Yau spaces.

35.2.1. The Generating Function for the Gromov–Witten Invariants of Holomorphic Maps from T^2 to T^2.

Normally in mirror symmetry the B-model calculation is much simpler then the A-model calculation, mainly because the latter involves generically general complicated summation over the worldsheet instantons, i.e., the holomorphic maps of the worldsheet into the target space. However in the T^2 case it is instructive and simple enough to sum over the worldsheet instantons.

We start with the A-model and study the holomorphic maps from the worldsheet torus T^2_τ to the target space torus T^2_σ. We can then construct F_1^{top} as the generating function of the Gromov–Witten invariants of these maps from basic properties of the latter.

Both tori will be defined as the complex plane divided by a two-dimensional lattice, i.e., $T_i^2 = \mathbb{C}/\Gamma_i$. The generators of the two-dimensional lattices Γ_i are given by the complex numbers (e_1^τ, e_2^τ) and (e_1^σ, e_2^σ) respectively. By the holomorphic automorphism of $\mathbb{C}$: $z \to \lambda z$ we may normalize $(1, \tau) = \left(\frac{e_1^\tau}{e_1^\tau}, \frac{e_2^\tau}{e_1^\tau}\right)$, so that τ parametrizes the complex structure of the worldsheet torus. Similarly $\sigma = \frac{e_2^\sigma}{e_1^\sigma}$ parametrizes the complex structure of the target space torus.

Clearly a map $X : \mathbb{C} \to \mathbb{C}$ is well defined as a map between the tori only if $X^* : \Gamma_\tau \hookrightarrow \Gamma_\sigma$, i.e., if

(35.2)

$$\begin{pmatrix} e_2^\tau \\ e_1^\tau \end{pmatrix} = \begin{pmatrix} r & m \\ s & n \end{pmatrix} \begin{pmatrix} e_2^\sigma \\ e_1^\sigma \end{pmatrix} \quad \text{or equivalently} \quad \tau = \frac{r e_2^\sigma + m e_1^\sigma}{s e_2^\sigma + n e_1^\sigma} = \frac{r\sigma + m}{s\sigma + n},$$

with $M = \begin{pmatrix} r & m \\ s & n \end{pmatrix} \in \mathrm{GL}(2,\mathbb{Z})$, see Fig. 1. Note that the T_τ^2 torus covers $d = |\det(M)|$ times the T_σ^2 torus, i.e., d is the degree of the map, which is given by

$$X(z) = (s\sigma + n)z. \tag{35.3}$$

It is lattice compatible as it maps the normalized lattice generators to lattice vectors, namely $X(1) = s\sigma + n$ and $X(\tau) = r\sigma + m$.

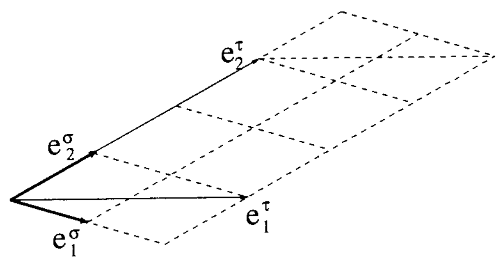


FIGURE 1. Lattice configurations that allow for a degree 6 map.

In string theory the Gromov–Witten invariants emerge from the evaluation of the variational integral[2] over all maps from the worldsheet to the target space torus weighted by the string action, whose stationary configurations are holomorphic maps $X_0 : T_\tau^2 \to T_\sigma^2$ of degree d. For the topological string the full path-integral localizes to holomorphic configurations and its measure collapses to a virtual fundamental class ϕ_d, which has to be integrated over the moduli space $\mathcal{M}_d$ of holomorphic maps. As is common in the definition of path-integrals with symmetries, one has eventually to divide by them. Here the symmetry acts as the automorphism group of $\mathcal{M}_d$.

If $d > 0$, then $\mathcal{M}_d$ is discrete so the integral is over points only, normalized to count one for each map. It is discrete because the condition

[2]The string analogue of the path-integral.

Eq. (35.2) localizes the integral over $d\tau$ in Eq. (35.1) to contributions from compatible complex structures.

The automorphism group of $\mathcal{M}_d$ is

$$\mathbb{Z}_2 \times \mathbb{Z}_d \times \mathrm{SL}(2,\mathbb{Z}). \tag{35.4}$$

The $\mathbb{Z}_2$ comes from the hyperelliptic involution of T_τ^2. Note that it reverses the orientation. Maps with $d > 1$ are all multi-covering maps. In this case we have a non-trivial $\mathbb{Z}_d$ action on $\mathcal{M}_d$ induced by the cyclic permutation of the sheets. The last factor in Eq. (35.4) is due to the holomorphic one to one maps from $T_\tau^2 \to T_\tau^2$ labeled by $\mathrm{SL}(2,\mathbb{Z})$. As is well known, only if $M \in \mathrm{SL}(2,\mathbb{Z})$ can we have a holomorphic one-to-one map between the tori, with a given orientation.

Taking into account the two orientations, which are killed[3] by the factor $\frac{1}{2}$ from the $\mathbb{Z}_2$, one gets for $d = 1$ the contribution 1 after dividing by $\mathrm{SL}(2,\mathbb{Z})$.

Multi-covering maps are given by $M \in \mathrm{GL}(2,\mathbb{Z})$ with $|\det(M)| = d > 1$. Again we have to divide by $\mathrm{SL}(2,\mathbb{Z})$, i.e., to sum only over equivalence classes $M \sim AM$ with $A \in \mathrm{SL}(2,\mathbb{Z})$. By a theorem of Hermite, the inequivalent summation is in this case over $\begin{pmatrix} r & 0 \\ m & \pm n \end{pmatrix}$ with $r, n \geq 1$ and $m = 0, \ldots, r-1$. In the sum below the two choices $\pm n$ are again killed by the $\mathbb{Z}_2$.

It remains to discuss the constant map piece $d = 0$. This is actually the only case in which an integral over the moduli space has to be performed. In the relevant limit, $\lim_{t,\bar{t} \to -i\infty} F_1$, we can evaluate Eq. (35.1) in the path-integral formulation explicitly for arbitrary Calabi–Yau manifolds M of complex dimension n. The integration over the fermion zero modes reduces the path-integral to

(35.5)

$$\frac{1}{2}(-1)^{n-1} \int_{\mathcal{F}} \frac{d\tau}{4\pi(\tau_2)^2} \int_M k \wedge c_{n-1}(TM) = \frac{(-1)^{\dim(M)-1}}{24} \int_M k \wedge c_{n-1}(TM),$$

where k is the Kähler form and $c_n(TM)$ is the nth Chern class of the tangential bundle of the target space.

[3]Note that $t \to -t$ for orientation reversing maps.

Putting everything together we can write down the generating function

$$F_1^{\text{top}} = -\frac{2\pi i t}{24} + \sum_{n,r=1}^{\infty}\sum_{m=0}^{r-1}\frac{e^{2\pi i n r t}}{nr} \tag{35.6}$$
$$= -\log \eta(t).$$

Here $\text{Im}(t)$ is the area (volume) of the target space and $\eta(t)$ is the Dedekind eta function $\eta(t) = q_t^{\frac{1}{24}}\prod_{n=1}^{\infty}(1-q_t^n)$ with $q_t = \exp(2\pi i t)$.

Instanton contributions are classical solutions X_0 of the equation of motions, which contribute to the path-integral. The exponential factor $e^{2\pi i d t}$ comes from the $e^{-S(X_0)}$ term in the path-integral and suppresses the instanton contributions for large volume. The constant map contribution has no such suppression and scales linearly with the volume $t \sim \int_{T^2_g} k$.

As we have seen before and will review in Sec. 35.2.2, it is natural to complexify t by an anti-symmetric two-form b to $t = b + iA$.

35.2.2. Direct Calculation of F_1. On T^2 we can solve the first quantized string theory and calculate its correlation functions explicitly. Given these expressions we can understand the action of the duality group, including mirror symmetry, for this simple example. In particular we can evaluate Eq. (35.1) and obtain

$$\begin{aligned} Z(\tau,\bar\tau) &= \text{Tr}\,(-1)^F F_L F_R q^{H_L}\bar q^{H_R} \\ &= \frac{\prod_{n=1}^{\infty}(1-q^n)^2(1-\bar q^n)^2}{\prod_{n=1}^{\infty}(1-q^n)^2(1-\bar q^n)^2}\sum_{(p_L,p_R)\in\Gamma} q^{\frac12 p_L^2}\bar q^{\frac12 p_R^2}. \end{aligned} \tag{35.7}$$

This can be derived using the methods discussed in Ch. 11. Eq. (35.7) is very close to computation of the Witten index, which would vanish for T^2. The insertion of $F_L F_R$ is responsible for leading to a non-vanishing result. There are four elements in the ground state $\{|0\rangle_k^F\} = \{|0\rangle, \Psi_0|0\rangle, \bar\Psi_0|0\rangle, \Psi_0\bar\Psi_0|0\rangle\} \sim \{1, dz, d\bar z, dz\wedge d\bar z\}$ on the target-space torus, cf. Eq. (13.86). The $F_L = \Psi_0^*\Psi_0$ and $F_R = \bar\Psi_0^*\bar\Psi_0$ are needed to absorb these fermionic zero mode. Note that $(-1)^F F_L F_R = 1/4$ for all the four ground states. In the operator formulation the numerator and denominator come from the fermionic and bosonic oscillator modes respectively. Γ is the four-dimensional "Narain lattice", parametrized by the geometrical background data, the metric G and anti-symmetric tensor field B. This was discussed for some special cases in Ch. 11. We define

$$I_F := \tau_2\text{Tr}(-1)^F F_L F_R q^{H_L}\bar q^{H_R} = \tau_2\sum_{(p_L,p_R)\in\Gamma} q^{\frac{p_L^2}{2}}\bar q^{\frac{p_R^2}{2}}. \tag{35.8}$$

Let us now discuss the remaining lattice sum from the bosonic zero modes, a generalization of Eq. (11.66). Worldsheet modular invariance requires Γ to be even and self-dual. This Narain lattice can be parametrized as

$$p_{L/R} = \frac{1}{\sqrt2}[n_i + m^j(B\pm G)_{ji}]e^{*\,i}, \tag{35.9}$$

where e^* spans the dual lattice of the torus $(G^{-1})^{ij} = \langle e^{*i}e^{*j}\rangle$. In the two dimensional case we have $3+1$ real background parameters, three coming from the metric, which may be parametrized by $e_2^\sigma = R_2 e^{i\phi}$, $e_1^\sigma = R_1$, and one, b parametrizing the value of the constant anti-symmetric two-dimensional tensor. It is convenient to combine these four parameters into two complex parameters

$$\sigma = \frac{G_{12}}{G_{22}} + i\frac{\sqrt{\det G}}{G_{22}} = \frac{e_2^\sigma}{e_1^\sigma}, \quad t = b + i\sqrt{\det G}, \tag{35.10}$$

with $A := \sqrt{\det G} = R_1R_2\sin\phi$. They represent the complex structure σ of the target space torus and its complexified area t, called the complexified Kähler structure. The inverse relation is

$$(G+B)_{ij} = \frac{A}{\sigma_2}\begin{pmatrix} |\sigma|^2 & \sigma_1 \\ \sigma_1 & 1\end{pmatrix} + \begin{pmatrix} 0 & b \\ -b & 0\end{pmatrix}, \quad (G^{-1})^{ij} = \frac{1}{A\sigma_2}\begin{pmatrix} 1 & -\sigma_1 \\ -\sigma_1 & |\sigma|^2\end{pmatrix}.$$

Time and space translation operators can be identified from $e^{2\pi i\tau_1 P - 2\pi\tau_2 H} = q^{\frac{p_L^2}{2}}\bar q^{\frac{p_R^2}{2}}$ (up to integers from the oscillators) as

$$\begin{aligned} H &= \tfrac12(p_L^2 + p_R^2) \\ &= \tfrac12(n_iG^{ij}n_j + 2m^iB_{ij}(G^{-1})^{jk}n_k + m^i(G - BG^{-1}B)_{ij}m^j), \\ P &= \tfrac12(p_L^2 - p_R^2) = m^in_i, \end{aligned} \tag{35.11}$$

where m^i and n_i are recognized as the winding and the momentum quantum numbers of the string. With a little calculation this can be expressed in the variables Eq. (35.10) using

$$\begin{aligned} p_L^2 &= \frac{1}{2t_2\sigma_2}|(n_2 - \sigma n_1) - t(m^1 + \sigma m^2)|^2, \\ p_R^2 &= \frac{1}{2t_2\sigma_2}|(n_2 - \sigma n_1) - \bar t(m^1 + \sigma m^2)|^2. \end{aligned} \tag{35.12}$$

In this form it is immediately apparent that T-duality on one circle, i.e., the exchange of

$$M: \quad t \leftrightarrow \sigma, \qquad n_1 \leftrightarrow m^1 \tag{35.13}$$

leaves (35.12,35.11) and hence (35.7) invariant[4]. Obviously, this exchange of complex structure and Kähler structure identifies this duality operation as mirror symmetry on the torus. Another important symmetry is the so-called axionic shift

$$T: \quad t \to t+1, \qquad \begin{array}{l} n_2 \to n_2 - m^1, \\ n_1 \to n_1 + m^2 \end{array}. \tag{35.14}$$

This symmetry can be seen directly from the path-integral, because the only b-dependent term in the action is $S_B = \frac{1}{4\pi\alpha'}\int_\Sigma d\sigma^2\ \epsilon^{\alpha\beta}B\epsilon_{\mu\nu}\partial_\alpha X^\mu \partial_\beta X^\nu$, where $b = BR_1R_2\sin\phi$ is the flux of the B-field through the torus. Hence a shift $b \to b+1$ gives an irrelevant $2\pi n_w$ phase in the exponential of the path-integral, where $n_w \in \mathbb{Z}$ is the winding number of the map X.

EXERCISE 35.2.1. *Find the generators of the full $O(2,2,\mathbb{Z})$ operation on n_i, m^i which leaves $P = (p_L^2 - p_R^2)/2 = n_1m^1 + n_2m^2 + N$ invariant up to a sign. Find its induced action on (σ, t), so that (35.7) is invariant. Give a physical interpretation of the generators.*

Solution: With $\mathrm{Im}(-\frac{1}{t}) = \frac{t_1}{|t|^2}$ we see that $S: [(n_1 \leftrightarrow m^1, n_2 \leftrightarrow -m^2), t \leftrightarrow -\frac{1}{t}]$ is a symmetry. Together with (35.14) this generates a $\mathrm{PSL}(2,\mathbb{Z})$ action on t. Conjugation by (35.13) gives the standard $\mathrm{PSL}(2,\mathbb{Z})$ action on σ, which defines the fundamental region $\mathcal{F}_\sigma = \mathbb{C}/\mathrm{PSL}(2,\mathbb{Z})$ for the space-time complex structure. In addition we have charge conjugation $C: [(\sigma, t) \to (-\bar\sigma, -\bar t)]$ and worldsheet parity $P: [(m^1 \to -m^1, m^2 \to -m^2), (t \to -\bar t)]$, which transforms $b \to -b$ and exchanges P_L with P_R.

To bring Eq. (35.8) to the expression which emerges naturally in the path-integral calculation of the amplitude we have to make a Poisson resummation, which is relegated to appendix A. As a result we can write

$$I_F = A\sum_M \exp\left[-2\pi i t\det(M) - \frac{\pi A}{\tau_2\sigma_2}\left|(1,\sigma)M\begin{pmatrix}\tau\\1\end{pmatrix}\right|^2\right]. \tag{35.15}$$

[4]The transformation of the shape of the torus is shown on the cover of this book, where b is set to zero. It is necessary and possible to show that all duality symmetries to be discussed are symmetries of all correlation functions.

Here the sum is over all integers in $M = \begin{pmatrix} m^2 & k^2 \\ m^1 & k^1 \end{pmatrix}$. For each summand the action of $\mathrm{SL}(2,\mathbb{Z})$ on $M \to MA$ can be undone by the fractional action $N(\tau) = \frac{n_{11}\tau + n_{12}}{n_{21}\tau + n_{22}}$ with $N = A^{-1}$ on τ. To obtain F_1 one has to integrate $\frac{1}{2}I_F$ against the $\mathrm{SL}(2,\mathbb{Z})$-invariant measure $\frac{d\tau}{\tau_2^2}$. Therefore one can trade the sum over elements in an $\mathrm{SL}(2,\mathbb{Z})$ orbit of M_i by an integration over the $\mathrm{SL}(2,\mathbb{Z})$ orbits $\mathcal{F}_i$ of $\mathcal{F}$ instead of $\mathcal{F}$. It turns out that there are three types of orbits to be distinguished:

1) $M_1 \equiv 0$: then $\mathcal{F}_1 = \mathcal{F}$ and the integration yields

$$I_1 = \frac{A}{2}\int_{\mathcal{F}}\frac{d^2\tau}{\tau_2^2} = \frac{A\pi}{6} = -2\pi i\frac{t-\bar t}{24}.$$

2) $\det(M_2) \neq 0$: then by the theorem of Hermite, M_2 can be represented by $M_2 = \begin{pmatrix} r & m \\ 0 & \pm n \end{pmatrix}$ with $r, n \geq 1$, $m = 0, \ldots, r-1$ and $\mathcal{F}_2$ becomes $\mathcal{F}_2 = 2H_+$, H_+ the upper half-plane. The integration yields

$$I_2 = A\sum_{\substack{0\le m<r\\ n\neq 0}} e^{-2\pi i r n}\int_{-\infty}^{\infty} d\tau_1\int_0^\infty \frac{d\tau_2}{\tau_2^2}e^{-\frac{\pi A}{\tau_2\sigma_2}|r\tau+m+p\sigma|^2} = -\log\prod_{n=1}^{\infty}|1-q_t^n|^2,$$

with $q_t = \exp(2\pi i t)$. We have encountered the first two orbits already in Sec. 35.2.1. From them we get $I_1 + I_2 = -\log(|\eta(t)|^2)$, whose holomorphic part coincides with Eq. (35.6).

3) $M_3 \neq 0$: $\det(M_3) = 0$ in this case $M_3 = \begin{pmatrix} 0 & k^2 \\ 0 & k^1 \end{pmatrix}$ and $\mathcal{F}_3 = \{\tau \in H_+ : |\tau_1| < \frac{1}{2}\}$. The integral

$$I_3 = \frac{A}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} d\tau_1\int_0^\infty \frac{d\tau_2}{\tau_2^2}\sum_{k^2,k^1\neq 0} e^{-\frac{\pi A}{\tau_2\sigma_2}|k^1+k^2\sigma|^2}$$

is divergent and has to be regulated by the factor $(1 - e^{-\frac{N}{\tau_2}})$ with $N \to \infty$, which yields $I_3 = -\log(|\eta(\sigma)|^2) - \frac{1}{2}\log(t_2\sigma_2) + C$. The constant C depends on the regularization, but $\partial_t F_1$ and $\partial_\sigma F_1$ are unambiguously defined. With $C = 0$ one gets

$$F_1 = -\log(\sqrt{\sigma_2}|\eta(\sigma)|^2) - \log(\sqrt{t_2}|\eta(t)|^2) \tag{35.16}$$

in complete accordance with Eqs. (35.6, 35.22), except that here we get the mirror symmetric, i.e., $(t \leftrightarrow \sigma)$ symmetric, expression.

We can simplify the previous calculation by taking the $\bar{t} \to -i\infty$ limit already in Eq. (35.15), which specializes, up to a t-independent (infinite) constant from case 3), to

$$\lim_{\bar{t}\to -i\infty} I_F = \frac{t-\bar{t}}{2i} + \sum_{M\in GL(2,\mathbb{Z})} e^{2\pi i t|\det(M)|}\frac{\tau_2^2}{|\det(M)|}\delta(\tau - M(\sigma)) + \mathcal{O}(e^{-2\pi i\bar{t}})$$

and the derivative $\frac{\partial}{\partial t}I_F$ can be integrated immediately w.r.t. the invariant measure $\frac{d\tau}{\tau_2^2}$ to give

$$\frac{\partial}{\partial t}F_1^{\rm top} := \frac{1}{2}\int_{\mathcal{F}}\frac{d\tau}{\tau_2^2}\frac{\partial}{\partial t}\lim_{\bar{t}\to -i\infty} I_F = -\frac{2\pi i}{24} + \frac{1}{2}\sum_{M(\sigma)\in\mathcal{F}} e^{2\pi i t|\det(M)|},$$

where the sum is over the orbit of type 2.). Hence one gets

$$\begin{aligned}\frac{\partial}{\partial t}F_1^{\rm top} &= -\frac{2\pi i}{24} + \sum_{n,r=1}^{\infty}\sum_{m=0}^{n-1} e^{2\pi i t n r}\\ &= -\frac{2\pi i}{24} + \sum_{n=1}^{\infty}\frac{nq^n}{1-q^n} = -\frac{\partial}{\partial t}\log\eta(t),\end{aligned}$$

in complete accordance with the analysis in Sec. 35.2.1 and Sec. 35.3.1.

35.3. The Ray-Singer Torsion and the Holomorphic Anomaly

Based on the equivalence between the path-integral and the operator formulation for purely imaginary τ, we can relate Eq. (35.1) via the path-integral to determinants of a Laplacian for B-model topological strings. Recall that in the B-model case, the topological string localizes to constant maps, and thus this is not very surprising. We will discuss this localization for arbitrary Calabi–Yau n-folds M. Let $\Delta_V = \Delta_1 + \Delta_2 = \bar{\partial}_V\bar{\partial}_V^\dagger + \bar{\partial}_V^\dagger\bar{\partial}_V$ be the Laplacian of a $\bar{\partial}$-operator coupled to a vector bundle V,

$$\bar{\partial}_V : \Omega^{q,p}(V) \to \Omega^{q,p+1}(V).$$

Locally such an operator looks like $\bar{\partial}_V = d\bar{z}^i(\partial_{\bar{z}^i} + a_i(z))$, where $a^i(z)$ are smooth matrix-valued functions. These operators are in one-to-one correspondence with the holomorphic structures on V and we call the moduli space of the latter $\mathcal{M}$. The spectrum $\mathcal{S}$ of Δ contains, by Hodge decomposition, the eigenvalues of Δ_1 and Δ_2. Furthermore if $\Delta_i^{(p)}$ is the Laplacian acting on p-forms, then $\mathcal{S}(\Delta_1^{(p)}) = \mathcal{S}(\Delta_2^{(p-1)})$ and $\mathcal{S}(\Delta_2^{(p)}) = \mathcal{S}(\Delta_1^{(p+1)})$. As the fermion number changes with the degree of the form, one would like to

consider $\prod_{p=0}^{n-1}(\det{}'\,\Delta_V^{(p)})^{(-1)^p}$, but this cancels due to the above symmetry[5]. So the simplest non-trivial expression to consider for Kähler manifolds is

$$I(V) = \prod_{p=0}^{n-1}(\det{}'\Delta_2^{(p)})^{(-1)^p} = \prod_{p=0}^{n-1}(\det{}'\Delta_V^{(p)})^{(-1)^p p}.$$

If V is flat this quantity, if properly normalized, is known as the holomorphic Ray–Singer torsion. One can use a heat kernel integral to regularize the determinant

$$\log I(V) = \int_\epsilon^\infty \frac{ds}{s}\mathrm{Tr}'(-1)^p p\exp(-s\Delta_V^{(p)}). \tag{35.17}$$

To relate this to the expression Eq. (35.1) we have to know the geometric realization of $F_{L/R}$. The Hilbert space of the B-model is given by

$$\mathcal{H} = \bigoplus_{p,q}\wedge^p\bar{T}_M^*\otimes\wedge^q T_M. \tag{35.18}$$

The identification of the topological states with the cohomology of $(n-q,\,p)$-forms in $H^{n-q,p}(M)$ is induced by the identification $G^+ = \frac{1}{2}(\bar{\partial} + \partial^\dagger)$, $\bar{G}^- = \frac{1}{2}(\bar{\partial} - \partial^\dagger)$, and hence $Q_B = G^+ + \bar{G}^- = \bar{\partial}$. The identification of the (p,q)-sectors of Eq. (35.18) with $H^{n-q,p}$ is by contraction of the indices in $\wedge^q T_M$ with the ones of the unique holomorphic $(n,0)$-form. $F_{L/R}$ is then given by

$$F_{L/R} = \frac{1}{2}((p-q) \pm i(k-k^\dagger)), \tag{35.19}$$

where k denotes the wedging with the Kähler form and $k^\dagger$ denotes the contraction with the Kähler form on M. Together with k_3, normalized to have eigenvalues $(p+q-n)/2$ on the $(n-q,p)$ elements in (35.18), k and $k^\dagger$ generate the well known sl(2) action on the cohomology of the Kähler manifold M[6]. Eq. (35.19) is justified by the fact that the expression has the right commutation relation[7] with the G's and the right parity. If F_LF_R is inserted in the trace $-(k-k^\dagger)^2$ can be replaced by $(p+q-n)^2$. Insertions of p^2, q^2, pn, qn and n^2 terms all give rise to various combinations of moduli

---

[5] As usual the prime denotes the product over nonzero eigenvalues only.

[6] This can be found e.g., in section 0.7 of [121]. Note the identification $k = L$, $k^\dagger = \Lambda$ and $h = 2k_3$.

[7] $[k,\partial^\dagger] = i\bar{\partial}$, $[k^\dagger,\partial] = i\bar{\partial}^\dagger$, $[k^\dagger,\bar{\partial}] = -i\partial^\dagger$ and $[k,\bar{\partial}^\dagger] = -i\partial$.

independent arithmetic genera $\chi_k = \sum_p (-)^p h_{k,p}$. So comparing Eq. (35.1) with Eq. (35.17) for purely imaginary τ and using $H = \Delta_V$ one can rewrite (35.20)

$$F_1 = \frac{1}{2}\int_{\mathcal{F}} \frac{d^2\tau}{\tau_2} \mathrm{Tr}_{\mathcal{H}}(-1)^{p+q}\, p\, q\, q^{H_R}\bar{q}^{\bar{H}_L} = \frac{1}{2}\sum_q (-1)^q q \log(I(\wedge^q T^*)).$$

For M a Riemann surface it was observed by Quillen that $I(V)$ determines a metric on the determinant line bundle over $\mathcal{M}$, whose curvature $\partial\bar{\partial}\log(I(V))$ is proportional to the Kähler form of $\mathcal{M}$. For general dimension of M this result has been extended to the formula (35.21)

$$\partial\bar{\partial}\log(I(V)) = \partial\bar{\partial}\sum_{p=0}^{n}(-1)^p \log(\det(g^{(p)})) + 2\pi i \int_M \mathrm{Td}(T)\mathrm{Ch}(V)|_{(1,1)}.$$

Here $g^{(p)}$ is the inner product in the zero mode space of $\Delta_2^{(p)}$ and T is the holomorphic tangent bundle of the Calabi–Yau, viewed as a bundle over the total space of M over $\mathcal{M}$, the universal Calabi–Yau space.

35.3.1. The Quillen Anomaly for T^2. For the case of F_1 on the two dimensional torus with $V = T^*$, the logarithm of the inner product in the space of zero modes is just the Kähler potential for the Weil–Petersson metric, the metric on the upper half-plane $g_{\sigma\bar{\sigma}} = \frac{1}{\mathrm{Im}^2(\sigma)}$. Hence the formulas Eqs. (35.20, 35.21) reduce to

$$\partial_\sigma\partial_{\bar{\sigma}}F_1 = -\frac{1}{2(\sigma-\bar{\sigma})^2}.$$

This equation is easily solved,

$$F_1 = -\frac{1}{2}\log(\sigma_2|f(\sigma)|^2) = -\log(\sqrt{\sigma_2}|f(\sigma)|),$$

where we encounter the simplest example of a holomorphic ambiguity, namely $f(\sigma)$. To fix it we note that due to space-time duality invariance F_1 must be invariant under the $\mathrm{PSL}(2,\mathbb{Z})$ action on the σ-plane. Hence f must compensate for the transformation of $\sqrt{\sigma_2}$ and thus transform as $f(-\frac{1}{\sigma}) = -i\sigma f(\sigma)$, i.e., it must be a modular form of weight 2. Furthermore, on physical grounds we know that F_1 must be finite inside the fundamental domain $\sigma \neq i\infty$. This fixes $f = \eta(\sigma)^2$ uniquely and so

$$F_1 = -\log(\sqrt{\sigma_2}|\eta(\sigma)|^2). \tag{35.22}$$

Note that if we identify σ with t as suggested by mirror symmetry we get in the $\lim_{\bar{t}\to -i\infty} F_1$ the function Eq. (35.6). In particular the $-\frac{2\pi i t}{24}$ comes from $\eta(t) \sim q_t^{\frac{1}{24}}$ in the $\mathrm{Im}(t) \to \infty$ limit.

35.3.2. Comparing the Holomorphic Anomaly and the Ray–Singer Torsion. It is interesting to contrast Eqs. (35.20, 35.21) directly with the holomorphic anomaly equation as derived in Ch. 31,

$$\partial_{\bar{j}}\partial_i F_1 = \frac{1}{2}\mathrm{Tr}\, C_i\bar{C}_{\bar{j}} - \frac{1}{24}G_{i\bar{j}}\mathrm{Tr}(-1)^F, \tag{35.23}$$

where the C_i, $\bar{C}_{\bar{j}}$ are matrix representatives of the action of the the chiral, anti-chiral fields ϕ_i, $\bar{\phi}_{\bar{j}}$ on the Ramond ground states. For more details see Ch. 17. The $\mathrm{Tr}\, C_i\bar{C}_{\bar{j}}$ reproduces the zero mode part of Eq. (35.21) as it appears in Eq. (35.20).

Let us turn to the topological term and rewrite this term in Eq. (35.20) as

$$\frac{1}{2}\mathrm{Td}(T)\sum_{q=0}^{n}(-1)^q q\mathrm{Ch}(\wedge^q T^*) = \frac{n}{4}c_n(T) + \frac{c_{n-1}(T)}{2} - \frac{1}{24}c_n(T)c_1(T) + \cdots.$$

If we integrate the latter w.r.t. $2\pi i\int_M$ and restrict to the $(1,1)$ part, we get indeed $2\pi i\chi(M)c_1(T)/24 = -\chi(M)G_{i\bar{j}}/24$, i.e. the identification of the last term in Eqs. (35.20, 35.21) and (35.23).

35.4. The Annulus Amplitude F_{ann} of the Open Topological String

35.4.1. B-model Calculation of F_{ann} on the Two-dimensional Torus. In this section we want to generalize the study of the one-loop amplitudes on a space-time torus T^2 to the open string. The relevant amplitude is the annulus amplitude[8]

$$F_{\mathrm{ann}} = -\frac{1}{2}\int_0^\infty \frac{ds}{s}\mathrm{Tr}'(-1)^F F\, \exp(-sH). \tag{35.24}$$

The picture becomes richer as mirror symmetry exchanges not only the complex and Kähler structure on the T^2 but also Dirichlet and Neumann boundary conditions for the open string. This action of mirror symmetry

[8]Klein bottle and Möbius strip amplitudes for the unoriented string are quite similar, but will not be discussed here.

on D-branes is discussed in Chapters 19, 37 and 39. As before we find interpretations for certain processes that depend only on the complex structure by using dual processes that depend only on the Kähler structure.

We describe here only the B-Model calculation of the annulus amplitude for the open string with Neumann boundary conditions on the torus with a rank N $U(N)$ bundle over the space-time torus. This corresponds to N D2 branes extending over the T^2. In this calculation no winding modes, but only shifted momentum modes, contribute. Therefore it can be reduced essentially to a point particle calculation for the quantity $F_{\rm ann} = \log\det\bar\partial_{\mathcal{A}}$. Alternatively one could calculate in this simple example $F_{\rm ann}$ also in the A-model after applying T-duality on one circle. We will leave this to the reader. However we will give in Sec. 35.4.2 the A-model interpretation of the B-model result.

Let us first discuss in some detail the effect of T-duality on the branes. Recall the couplings of the string (σ-model) on bounded worldsheets to the background parameters

$$
\begin{aligned}
S =&\frac{1}{\pi\alpha'}\int_\Sigma d^2\sigma\left[G_{\mu\nu}\partial_a X^\mu\partial_b X^\nu\eta^{ab} + B_{\mu\nu}\partial_a X^\mu\partial_b X^\nu\epsilon^{ab}\right]\\
&+\frac{1}{2\pi\alpha'}\oint_{\partial\Sigma} d\tau A_i\partial_\tau\xi^i.
\end{aligned}\tag{35.25}
$$

Here we wrote for simplicity only one $U(1)$ field A_i, $i=0,1,2$ on the D2 brane with internal coordinates ξ^i, $i=0,1,2$. The usual gauge transformation $B_{\mu\nu}\to B_{\mu\nu}+\partial_\mu\Lambda_\nu-\partial_\nu\Lambda_\mu$ must be compensated for on the boundary by $A_\mu\to A_\mu-\Lambda_\mu$ so that the invariant field strength is $\mathcal{F}_{\mu\nu}=B_{\mu\nu}-F_{\mu\nu}$. For simplicity we therefore may choose later a B field so that this quantity vanishes. Also we restrict ourselves to constant background fields. Simple boundary conditions with no sources at the boundaries can be determined. If the coordinates X^1, X^2 are on the T^2 these are[9]

$$
\begin{aligned}
\partial_{\sigma^1}X^0 &= 0,\\
\partial_{\sigma^1}X^1+\mathcal{F}\partial_{\sigma^0}X^2 &= 0,\\
\partial_{\sigma^1}X^2-\mathcal{F}\partial_{\sigma^0}X^1 &= 0,\\
\partial_{\sigma^0}X^a &= 0,\quad a=3,\dots,9
\end{aligned}\tag{35.26}
$$

A T-duality transformation as in Eq. (35.13), but here on the X^2 coordinate, exchanges Neumann into a Dirichlet boundary condition. Namely without

[9]Hereafter we will deal only with the X^1, X^2 coordinates.

imposing boundary conditions the relevant bosonic string coordinate is a solution of the string e.o.m., if $X^2(z,\bar z)=X^2_L(z)+X^2_R(\bar z)$ with $z=\sigma^1+i\sigma^2$, i.e., $\sigma^2=i\sigma^0$ is the Euclidean time and σ^1 the space coordinate. T-duality changes the sign of $X_R(\bar z)$ so that a Neumann boundary condition on the old coordinate $\partial_n X^2(z,\bar z)=0$ becomes on the new embedding coordinate $X^{2'}(z,\bar z)=X^2_L(z)-X^2_R(\bar z)$ a Dirichlet boundary condition $\partial_{\sigma^1}X^2=\partial_n X^2=-i\partial_t X^{2'}=-\partial_{\sigma^0}X^{2'}=0$. On an S^1 without Wilson lines this forces the open string endpoints to be at $X^{2'}(\pi)-X^{2'}(0)=-2\pi n R_2'$ with $n\in\mathbb{Z}$. Wilson lines in the R_2 directions modify this to

$$
X^{2'}(\pi)-X^{2'}(0)=-(2\pi n-\theta_j+\theta_i)R^{2'}.\tag{35.27}
$$

The T-dual boundary conditions to Eq. (35.26) are hence

$$
\begin{aligned}
\partial_{\sigma^1}(X^1+\mathcal{F}X^2)=&0,\\
\partial_{\sigma^0}(X^2-\mathcal{F}X^1)=&0.
\end{aligned}\tag{35.28}
$$

This correspond to a D1 brane or D-string with an angle ϕ relative to the X^2 axis, where $\cot(\phi)=\mathcal{F}$. The energy, hence the length, of the D-string should be finite, which means that ϕ should be chosen rational w.r.t. the $(\sigma',1)$ lattice, so that the string comes back to itself after winding (n,m) times around the two torus cycles. Here $\sigma'=\tau=\frac{R_1'}{R_2'}e^{i\alpha'}$ is the complex structure of the dual torus[10]. If ϕ is rational, we may as well consider the situations, where the D-string winds only around the X^1 direction.

This corresponds to a flat $U(1)$ bundle. We are interested in general in the situation of a flat $U(N)$ bundle. The moduli for such a bundle are given by the rank N flat connection $A_i=\theta_i dx_0-\phi_i dx_1$. All we have to do is to calculate $\det\bar\partial_{\mathcal{A}}$ in this background.

Consider first a single complex fermion with anti-periodic boundary conditions in time and space directions (A,A). We label the points on the torus by $x=x_0+\sigma x_1$, where σ parametrizes the complex structure of the torus. The fermion couples to one flat gauge field A_i and picks up an additional gauge phase

$$
\begin{aligned}
\Psi(x_0+2\pi,x_1)&=-e^{2\pi i\theta_i}\Psi(x_0,x_1),\\
\Psi(x_0,x_1+2\pi)&=-e^{2\pi i\phi_i}\Psi(x_0,x_1),
\end{aligned}\tag{35.29}
$$

when transported around the a,b cycle respectively.

[10]Rationality w.r.t. the lattice can be more conveniently formulated after exchanging the (X^1,X^2) directions.

It is straightforward to write down the partition function for a particle with twisted boundary conditions in the operator formulation. E.g., for the complex fermion subject to Eq. (35.29) we get, after canonical quantization, the Hamiltonian[11]

$$H = \sum_{-\infty}^{\infty}(n+\theta-\frac{1}{2}) : b^{\dagger}_{n+\theta-\frac{1}{2}} b_{n+\theta-\frac{1}{2}} : + \left(\frac{\theta^2}{2}-\frac{1}{24}\right).$$

Note the absence of $\frac{1}{2}$ for complex fermions. In the Hilbert space generated by the $b^{\dagger}_{n+\theta-1/2}$ subject to $\{b^{\dagger}_{r_1}, b_{r_2}\} = \delta_{r_1,r_2}$, $\{b^{\dagger}_{r_1}, b^{\dagger}_{r_2}\} = \{b_{r_1}, b_{r_2}\} = 0$, we consider $\mathrm{Tr}_{\mathcal{H}} g_i q^H$, where g_i is the twisting in the time direction $g_i b_{n+\theta-\frac{1}{2}} g_i^{-1} = -e^{2\pi i \phi_i}$. Let us introduce the complex parameter $u_i = \phi_i + \sigma\theta_i$. The determinant of the chiral Dirac operator $\det'_C(\bar\partial_{A_i})$ can be calculated using the equivalence of the operator and the path-integral formulation for purely imaginary time $2\pi\sigma_2$. Let us assume first that the field Ψ couples only to the field A_i. Then $\det'_C(\bar\partial_{A_i})$ depends essentially holomorphically on u_i via $z_i = e^{2\pi i u_i}$,

$$\det{}'_C(\bar\partial_{A_i}) = q^{\frac{\theta_i^2}{2}-\frac{1}{24}} \prod_{n=1}^{\infty}(1+z_i q^{n-\frac{1}{2}})(1+z_i^{-1}q^{n-\frac{1}{2}}),$$

up to the factor involving the regularization $\theta_i^2 = \left(\frac{1}{2i\sigma_2}(u_i-\bar u_i)\right)^2$, which constitutes a factorization anomaly. Here we denote $q := \exp 2\pi i\sigma$.

Now let the fermions on the space-time torus couple in the adjoint representation to the flat background gauge fields A_i in the Cartan subalgebra of $u(n)$, which generates the maximal torus of $U(N)$. The θ_i, ϕ_i denote the holonomies of the maximal torus of $U(N)$ around the a, b cycles of the target space torus.

To calculate Eq. (35.24) we take the following modifications into account in the evaluation of $\log\det'(\bar\partial_{A'})$. There are $N \times N$ Dirac fermions Ψ_{ij} with both chiralities in the adjoint representation of the gauge group. The fermions couple with the charges $e_i - e_j$ to the fields in the Cartan subalgebra. Due to topological twisting the fields are in the periodic sector (P,P), i.e., the sign in the boundary conditions in Eq. (35.29) is flipped. The expression would vanish if we were not considering the primed determinant, i.e., with

[11]Here we use the zeta function normalization: The analytic continuation of $\zeta(z,x) = \sum_{n=0}^{\infty}(x+n)^{-z}$, $\mathrm{Re}\, z > 1, x \neq 0, -1, -2, \ldots$ yields $\zeta(-n,x) = -\frac{B_{n+1}(x)}{n+1}$. With the definition of the Bernoulli polynomials ($B_2(x) = \frac{x}{2} - \frac{x^2}{2} - \frac{1}{12}$) it follows that $\sum_{n=0}^{\infty}(\theta - \frac{1}{2} + n)$ "=" $\frac{\theta^2}{2} - \frac{1}{24}$.

the zero modes removed. The regularized sum over the normal ordering contributions for one chirality is now $\frac{1}{2}\sum_{n>0} n = \frac{1}{24}$. Hence one arrives at (35.30)

$$F_{\mathrm{ann}} = -\frac{1}{2}\log\det{}'(\bar\partial_{A'}) = -\log\prod_{i,j=1}^{N}\exp\left(\frac{\pi(u_i-\bar u_i-(u_j-\bar u_j))^2}{4\sigma_2}\right) \times \left|\prod_{n=1}^{\infty} q^{\frac{1}{24}}(1-z_i z_j^{-1}q^n)\right|^2.$$

35.4.2. A-model Interpretation of the Annulus Amplitude F_{ann} on the Torus. We can interpret this function on the mirror side, i.e., trade σ against t, by making T-duality on one circle Eq. (35.13); i.e., for flat bundles and vanishing B-field the N two-branes wrapping the whole torus with Wilson line parameters ϕ_i, θ_i, $i = 1, \ldots, N$ become N one-branes parallel to the real axis but shifted by a distance θ_i and with Wilson line parameters ϕ_i along the brane as shown in Fig. 2

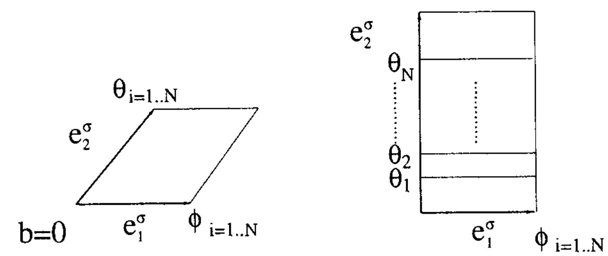


FIGURE 2. Map of the two-brane configuration by mirror symmetry into parallel one-branes

We can now consider the topological limit $\bar t \to \infty$. In this limit we get

$$F^{\mathrm{top}}_{\mathrm{ann}}(t) = \sum_{i,j} 2\pi i t[-\frac{1}{24} + \frac{(\theta_i-\theta_j)^2}{4}] + \sum_{n,m,i,j}\frac{z_i^m z_j^{-m} q^{nm}}{m}. \tag{35.31}$$

This function has an immediate interpretation in terms of open string annulus instantons. For the lowest energy configurations these are holomorphic embeddings of the annuli swept out by the open string loop (or the closed string tree-level propagation) in the T^2. In the open string interpretation one end goes m times around the torus in the direction of the one-brane i. The lowest mode of the open string corresponds to a configuration in which the string forms a line perpendicular to the brane i and ends after wrapping n times around the torus at the brane j. Its worldsheet area is therefore $nm\tau$ plus (or minus) the strip between brane i and brane j, hence $A = (\theta_i - \theta_j + nm)\tau$. From the path-integral we also get a boundary contribution of $2\pi i(\phi_i - \phi_j)$ from the Wilson lines along brane i and brane j which is. By the map to the torus, the continuous shift symmetry around the S^1 of the annulus is broken to $\mathbb{Z}_m$. Taking the exponential and dividing out by the $\mathbb{Z}_m$, we arrive at the sum in Eq. (35.31).

35.5. F_1 on Calabi–Yau in Three Complex Dimensions

In this section we will discuss the topological string amplitude F_1 on Calabi–Yau threefolds and give an example calculation.

35.5.1. Integration of the holomorphic anomaly equation. For three-dimensional Calabi–Yau spaces, Eq. (35.23) specializes to

$$\partial_i\bar{\partial}_{\bar{j}}F_1 = \frac{1}{2}C_{ijk}\bar{C}_{\bar{j}\bar{k}\bar{l}}e^{2K}G^{j\bar{k}}G^{k\bar{l}} - \left(\frac{\chi}{24}-1\right)G_{i\bar{j}}, \tag{35.32}$$

where the C_{ijk} are the three-point functions at genus 0. These three-point functions and the Kähler potential for the Weil-Petersson metric G follow from the prepotential $\mathcal{F}$, which is given in terms of periods on W, as discussed in earlier chapters.

Using $R_{i\bar{j}} = -\frac{1}{2}\partial_i\bar{\partial}_{\bar{j}}\log\det(G)$, which holds for Kähler manifolds, and the special geometry relation

$$R^k_{i\bar{j}l} = G_{i\bar{j}}\delta^k_l + G_{l\bar{j}}\delta^k_i - C_{ilm}\bar{C}_{\bar{j}\bar{p}\bar{q}}e^{2K}G^{k\bar{p}}G^{m\bar{q}}, \tag{35.33}$$

one can integrate Eq. (35.32) up to an unknown holomorphic function f to obtain

$$F_1 = \log\left(\det(G^{-1})^{\frac{1}{2}}e^{\frac{K}{2}(3+h^{2,1}-\frac{1}{12}\chi)}|f|^2\right). \tag{35.34}$$

The holomorphic ambiguity $f(z)$ can be parametrized by the vanishing or pole behavior at the discriminant loci $f = \prod_{i=1}^{k}(\Delta_k)^{r_i}\prod_{i=1}^{h^{2,1}} z_i^{x_i}$. In particular, the x_i can all be solved for from the limiting behavior (35.5) $\lim_{z_i\to 0} F^{(1)} = -\frac{1}{24}\sum_{i=1}^{h^{2,1}} t_i\int_M c_2 J_i$. The behavior of F_1 at certain types of singularities is universal — e.g., for the conifold singularity $r_{\rm con} := -\frac{1}{12}$. In this case the leading behavior of F_1 is captured by $f(z)$. In general the inverse metric and the Kähler factor can also contribute to the leading behavior.

35.5.2. The Mirror Map $t_i(z_j)$ and the Holomorphic Limit. To use the previous B-model calculation of F_1 for the prediction of Gromov–Witten invariants, we need the relation between complex structure variables on W denoted by z_i and the complexified Kähler structure variables on the mirror M, which we call t_i.

There are $2h^{2,1}(W)+2 = 2\dim(\mathcal{M})+2$ period integrals $\omega_i(\vec{z}) = \int_{\gamma_i}\Omega$, which depend on $h^{2,1}(W)$ complex structure variables z_i. There exist points (at least one) in the complex moduli space $P_\infty = \{z_i = 0\}\in\mathcal{M}$, where the $h^{2,1}(W)$ periods have logarithmic behavior $\omega_i \sim \log(z_i)+O(z)$, $i = 1,\ldots,h^{2,1}$ and a unique period is analytic[12], $\omega_0 = 1+O(z)$. The homogeneous coordinates, defined as

$$2\pi i\ t_i = \frac{\omega_i}{\omega_0} \sim \log(z_i)+O(z), \tag{35.35}$$

have the following properties, which define canonical coordinates near any point P_0.

All holomorphic derivatives

$$\partial_{t_1}\ldots\partial_{t_r}\Gamma^k_{ij}|_{P_0} = \partial_{t_1}\ldots\partial_{t_r}K|_{P_0} = 0 \tag{35.36}$$

vanish at P_0. As above, let P_0 be at $z=0$, $t=t_0$; then Eq. (35.36) implies that in the t coordinates the leading term in $\bar{\lambda}_i = (\bar{t}_i - \bar{t}_0)$ of $K = C+O(\bar{\lambda})$ and $G_{i\bar{j}} = C_{i,\bar{j}}+O(\bar{\lambda})$ is constant. When re-expressed in the coordinates z_i, the holomorphic parts[13] of K and G in the $\bar{\lambda}_i\to 0$ limit are

$$K = C - \log(\omega_0), \qquad G_{i\bar{m}} = \frac{\partial t_k}{\partial z_i}C_{k\bar{m}}. \tag{35.37}$$

Similarly in all quantities to be discussed below we will take the holomorphic $\bar{\lambda}\to 0$ limit.

The t_i at $P_\infty = \{z_i = 0\}\sim\{\mathrm{Im}(t_i)\to\infty\}$ have the additional property that they are identified by mirror symmetry with the complexified Kähler parameters of the mirror M. A simple reason for this identification comes from a symmetry consideration. We know from the path-integral that the shifts Eq. (35.14) are exact symmetries for all variables t_i in the large radius

[12]The other $h^{2,1}(W)$ periods are quadratic in the logarithms and one is cubic. These points may be characterized by their maximal unipotent monodromy.

[13]There is an important simplification in the local case. As is clear from the differential equations associated to the local case, $\omega_0 = 1$ is always a holomorphic solution, and hence for the local case the holomorphic part of the Kähler potential becomes trivial in the limit Eq. (35.37).

limit of the A-model on W. On the other hand, symmetries of the B-model, which depend on the complex structure of M, are realized as monodromies. Eq. (35.35) identifies the shifts Eq. (35.14) with the monodromies around z_i.

35.5.3. F_1 on the Quintic. Here we will calculate F_1 on the quintic hypersurface. The mirror is given by the vanishing locus of the one-dimensional perturbation family in $\mathbb{P}^4$,

$$W = \sum_{i=1}^{5} z_i^5 - 5a \prod_{i=1}^{5} z_i = 0, \tag{35.38}$$

subject to the $\mathbb{Z}_5^3$ identifications $(z_1, z_2, z_3, z_4, z_5) \sim (\alpha z_1, \alpha^{-1} z_2, z_3, z_4, z_5) \sim (\beta z_1, z_2, \beta^{-1} z_3, z_4, z_5) \sim (\gamma z_1, z_2, z_3, \gamma^{-1} z_4, z_5)$ with $\alpha^5 = \beta^5 = \gamma^5 = 1$. We note the topological data $\chi(M) = -200$, $h^{1,1} = 1$ and $\int_M c_2 J = 50$. Using them in Eq. 35.34 one obtains

$$F_1 = \log\left(\det(G^{-1})^{\frac{1}{2}} e^{\frac{31K}{3}} |f|^2\right). \tag{35.39}$$

The ambiguity f must be fixed from the behavior of F_1 at the boundaries of the moduli space. An exhaustive discussion of this moduli space and the monodromies can be found in the paper by Candelas, de la Ossa, Green and Parkes [**45**]. Here we note that the one-dimensional moduli space of the mirror quintic is compactified to $\mathbb{P}^1$. The boundaries consist of the discriminant loci. These are values of the moduli for which the manifold becomes singular. E.g., for finite a we find a solution[14] to $W = dW = 0$ for $a^5 = 1$ and $z_1 = \cdots = z_5 = 1$. The Hessian $\frac{\partial^2}{\partial z_i \partial z_j} W$ at this point is non-degenerate. Therefore we have here a nodal singularity, with local equation $\sum_{i=1}^4 \epsilon_i^2 = (a-1)$. If e.g., in a two-dimensional moduli space two nodal loci meet[15], then the Hessian has one zero eigenvalue at the meeting point and one of the local variables starts with ϵ_k^3 in the local equation. The structure of the local singularity determines the leading behavior of F_g. At $a = \infty$ Eq. (35.38) degenerates into five hyperplanes $z_i = 0$ with hierarchically more severe singularities, when $2, 3$ or 4 of the hyperplanes meet[16]. From

[14] Other solutions are identified with this solution by the $\mathbb{Z}_5^3$ symmetry.

[15] These loci in the moduli space are known as conifold divisors.

[16] This leads to the $1, \log, \log^2, \log^3$ degeneration structure of the periods at $z = \frac{1}{(5a)^5} = 0$.

the Picard-Fuchs equation in the variable $w = 1/z = (5a)^5$ and with the definition $\theta = w\frac{d}{dw}$,

$$[w\theta^4 - 5(5\theta+4)(5\theta+3)(5\theta+2)(5\theta+1)]\omega(w) = 0,$$

we see that there are four power series solutions $i = 1, \ldots, 4$ starting with $\omega_i = w^{\frac{i}{5}} + O(w^{\frac{1+i}{5}})$. By analytic continuation one notes that the symplectic basis of periods starts with $\Pi_i = c_i w^{\frac{1}{5}} + O(w^{\frac{1+1}{5}})$. We have therefore $e^K \sim e^{-\frac{1}{5}\log(w\bar{w})} \sim |z|^{\frac{2}{5}}$ and $\det(G^{-1})^{\frac{1}{2}} \sim |z|^{\frac{6}{5}}$. At $z = \infty$, F_1 should be regular as this point does not correspond to a singular manifold. This relates the parameters c, d in the ansatz $f(z) = z^c(1-5^5 z)^d$ for the ambiguity: $c = -\frac{3}{5} - \frac{31}{15} - d$. Finally in the limit $\mathrm{Im}(t) \to \infty$ F_1 behaves as $F_1 \sim \frac{50}{24}(t + \bar{t})$, which yields $d = -\frac{1}{12}$.

It turns out that not only the leading behavior of F_1 but also that of the other F_g are universal at the nodal singularity. The results for the topological invariants n_d^1 of the quintic are summarized in the table in Sec. 35.6.4.

35.5.4. F_1 on the Non-compact Calabi–Yau Space $\mathcal{O}(-3) \to \mathbb{P}^2$. The non-compact Calabi–Yau space $\mathcal{O}(-3) \to \mathbb{P}^2$ has a toric description specified by the non-complete fan spanned by $(1,0,0)$ and $(1,1,0)$, $(1,0,1)$, $(1,-1,-1)$. The Mori generator represents a relation between these points and is given by $l = (-3,1,1,1)$. The construction of the mirror geometry of the non-compact case is a generalization of Batyrev's construction. Following this construction we define Y_i, $i = 0, \ldots, 3$ to be complex coordinates subject to a $\mathbb{C}^*$ scaling action, which fulfill the relation

$$\prod_{i=0}^{3} Y_i^{l_i} = 1 \quad \text{with} \quad l = (-3,1,1,1) \tag{35.40}$$

and consider the constraint

$$p = a_0 Y_0 + a_1 Y_1 + a_2 Y_2 + a_3 Y_3 = 0. \tag{35.41}$$

To exhibit it in a more common form we can introduce new coordinates z_i that solve Eq. (35.40). They are related to Y_i by a so-called étale map $Y_i = z_i^3$ for $i = 1, 2, 3$ and $Y_0 = z_1 z_2 z_3$. The constraint Eq. (35.41) is now identified with the elliptic curve given by the cubic

$$P = \sum_{i=1}^{3} z_i^3 - 3a \prod_{i=1}^{3} z_i = 0 \tag{35.42}$$

in $\mathbb{P}^2$. Note that the map from the z_i to the Y_i is many-to-one so that one has to divide the z_i by a $\mathbb{Z}_3^2$ action to make it well defined. This orbifoldisation includes a $\mathbb{Z}_3$ action, which is an analogue of the $\mathbb{Z}_5^3$ action of the previous example. The three-dimensional mirror geometry can be defined in non-homogeneous coordinates by the equation

$$p|_{Y_0=1} = uv, \tag{35.43}$$

where $u, v \in \mathbb{C}$.

The periods on cycles with compact support can be reduced to integrals of the meromorphic form $\lambda = \oint \log(P)\frac{dz_1}{z_1}\frac{dz_2}{z_2}$ on the cubic in $\mathbb{P}^2$, where the loop integral is along a path encircling the zero $P = 0$. A direct calculation shows that $\omega_i = \int_{\Gamma_i} \lambda$ is annihilated by the differential operator

$$\mathcal{L} = \theta^3 - \tilde{z}\prod_{i=0}^{2}(\theta + \frac{i}{3}) = \tilde{\mathcal{L}}\theta, \tag{35.44}$$

where $\tilde{z} = \frac{1}{a^3}$ and $\theta = \tilde{z}\frac{d}{d\tilde{z}}$.

We note that $\theta\lambda = -3a\oint \frac{1}{P}z_3 dz_1 dz_2 =: \tilde{\omega}$ is a holomorphic differential on the elliptic curve, whose periods are annihilated by $\tilde{\mathcal{L}}$. To relate $\tilde{\omega}$ to the standard holomorphic differential $\omega = \frac{dx}{y}$, we bring Eq. (35.42) by a $GL(3)$ transformation on the coordinates z_i into the Weierstrass-form

$$y^2 = 4x^2 - g_2 x - g_3, \quad \text{with} \quad g_2 = 3a(8 + a^3), \quad g_2 = 8 + 20a^3 - a^6 \tag{35.45}$$

and perform the integration over one coordinate (y after the change of variables) in the definition of $\tilde{\omega}$. This identifies $\theta\lambda = \tilde{\omega} = -3a\frac{dx}{y} = -3a\omega$.

The differential operator in Eq. (35.44) has a solution, the so-called Meijer G-functions $G^{2,2}_{3,3}\left(-w\left|\begin{matrix}\frac{1}{3} & \frac{2}{3} & 1\\ 0 & 0 & 0\end{matrix}\right.\right)$. To be concrete we note that the following is a basis of solutions: $f_0 = 1$,

$$\begin{aligned} t = f_1 &= \frac{1}{2\pi i}\left(\log(z) + \sum_{n=1}\frac{(3n)!}{(n!)^3 n}z^n\right) \quad \text{and} \\ f_2 &= \frac{1}{(2\pi i)^2}\left(\log(z)^2 + 2\log(z)\sum_{n=1}\frac{(3n)!}{(n!)^3 n}z^n + \sum_{n=0}^{\infty}c_n z^n\right), \end{aligned} \tag{35.46}$$

with $z = -\frac{\tilde{w}}{27}$ and $c_n = 6\frac{(3n)!}{(n!)^3 n}\left(\Psi(3n+1) - \Psi(n+1) - \frac{1}{3n}\right)$. As the holomorphic solution is a constant, the mirror map Eq. (35.35) is in this case

simply given by the logarithmic solution. As observed in general the exponential $q = \exp 2\pi i t$ of the mirror map

$$q = z - 6z^2 + 63z^3 - 866z^4 + 13899z^5 - 246366z^6 + O(z^7), \tag{35.47}$$

as well as its inverse, has integral coefficients

$$z = q + 6q^2 + 9q^3 + 56q^4 - 300q^5 + 3942q^6 + O(q^7). \tag{35.48}$$

A basis of solutions that correspond to period integrals over the integral homology of the elliptic curve can be shown to be given by

$$\begin{pmatrix}\omega_2\\ \omega_0\\ \omega_1\end{pmatrix} = \begin{pmatrix}\frac{\partial\mathcal{F}}{\partial t_i}\\ 1\\ t\end{pmatrix} = \begin{pmatrix}\frac{1}{6}t^2 - \frac{1}{6}t - \frac{1}{12} + \sum_{k,p=1}^{\infty}\frac{q^{kp}kn_k^0}{p^2}\\ 1\\ t\end{pmatrix}. \tag{35.49}$$

The occurrence of the derivative of the prepotential $\mathcal{F}$ with the genus 0 instanton numbers n_k^0 can be explained from its appearance in the periods of compact Calabi–Yau manifolds.

To see this consider the compact elliptic fibration over $\mathbb{P}^2$, which can be represented as a degree 18 Calabi–Yau hypersurface X_{18} in $\mathbb{P}(1,1,1,6,9)$. It can be specialized to the geometry $\mathcal{O}(-3) \to \mathbb{P}^2$ by taking the size of the elliptic fiber to infinity. The periods in Eq. (35.49) are those periods of X_{18} that stay finite in this limit.

Eq. (35.37) and the fact that $\omega_0 = 1$ implies some simplification due to a trivial e^K factor in Eq. (35.34). More important simplifications occur in the propagators due to this fact, see Sec. 35.6.1. Note that all dependence of the F_g^{top} on the Euler number χ disappears due to this simplification.

The singularities of the geometry occur at $z = 0$, the degeneration into hyperplanes, and $1 - 3^3 z = 0$, a conifold locus. We therefore parametrize the holomorphic ambiguity by $f = z^c(1 - 3^3 z)^d$ and determine the coefficients c and d again by requiring regularity at $1/z = 0$ and $F_1 \sim \frac{c_2 J}{24}(t + \bar{t})$ in the limit $\mathrm{Im}(t) \to \infty$ of F_1. The intersections $c_2 J$ for the non-compact cases were defined by an adjunction formula. For the case at hand we have $c_2 J = -2$, which fixes, together with the regularity at $1/z = 0$, $c = 2$, $d = -\frac{1}{12}$. The genus 0 and the genus 1 instanton numbers can be found in the tables in Sec. 35.6.5.

35.6. Integration of the Higher Genus Holomorphic Anomaly Equations

In this section we will connect to the discussion of the higher-genus anomaly equations in Sec. 31 and describe their explicit integration. We will apply the result to the quintic and the non-compact $\mathcal{O}(-3) \to \mathbb{P}^2$ geometry.

The higher-genus correlators, i.e., partition functions with fields $\Phi_{i_1}, \dots, \Phi_{i_n}$ inserted, also satisfy a holomorphic anomaly equation given by[17]

$$\begin{aligned} \bar\partial_{\bar i} F^{(g)}_{i_1,\dots,i_n} &= \frac{1}{2}\bar C^{jk}_{\bar i} F^{(g-1)}_{jki_1,\dots,i_n} \\ &+ \frac{1}{2}\bar C^{jk}_{\bar i} \sum_{r=0}\sum_{s=0} \frac{1}{s!(n-s)!} \sum_{\sigma\in S_n} F^{(r)}_{ji_{\sigma(1)}\cdots i_{\sigma(s)}} F^{(g-r)}_{ki_{\sigma(s+1)}\cdots i_{\sigma(n)}} \\ &- (2g-2+n-1)\sum_{s=1}^{n} g_{\bar i, i_s} F^{(g)}_{i_1\dots i_{s-1}i_{s+1}\dots i_n}, \end{aligned} \tag{35.50}$$

where $\bar C^{ij}_{\bar i} := \bar C_{\bar i\bar j\bar k} e^{2K} g^{j\bar j} g^{k\bar k}$ and the subscripts of F denote the field insertions. The right-hand side corresponds to the boundary contributions of the moduli space of marked Riemann surfaces, $\overline{\mathcal{M}}_{g,n}$, as depicted in Fig. 3. More precisely the first term comes from pinching a handle a), the next from splitting the surface into two components by growing a long tube b), and the last c) arises when two marked points approach each other.

35.6.1. The Propagators. The solution of Eq. (35.50) is provided by the calculation of potentials for the anti-holomorphic quantities $\bar C^{ij}_{\bar i}$. In the first step, one calculates $S^{ij} \in \mathcal{L}^{-2} \otimes \mathrm{Sym}^2 T\mathcal{M}$ such that $\bar C^{ij}_{\bar i} = \bar\partial_{\bar i} S^{ij}$. This follows from (35.33) by noting that in Kähler geometry $R^k_{i\bar j l} = -\bar\partial_{\bar j}\Gamma^k_{il}$, and hence

$$\bar\partial_{\bar k}[S^{ij} C_{jkl}] = \bar\partial_{\bar k}[\delta^i_l \partial_k K + \delta^i_k \partial_l K + \Gamma^i_{kl}]. \tag{35.51}$$

The derivatives $\partial_{\bar k}$ on both sides can be removed at the price of introducing a meromorphic object f^i_{kl}, which also has to compensate for the non-covariant transformation properties of quantities on the right-hand side.

Therefore it is natural to split f^i_{kl} into quantities with simple transformation properties: $f^i_{kl} = \delta^i_k \partial_l \log f + \delta^i_l \partial_k \log f - v_{l,a}\partial_k v^{i,a} + \tilde f^i_{kl}$, where $\tilde f^i_{kl}$ now transforms covariantly, $f \in \mathcal{L}$, and $v^{i,a}$ transform as tangent vectors.

[17]To distinguish the genus index from the indices of the marked points we write $F^{(g)}$ in this section.

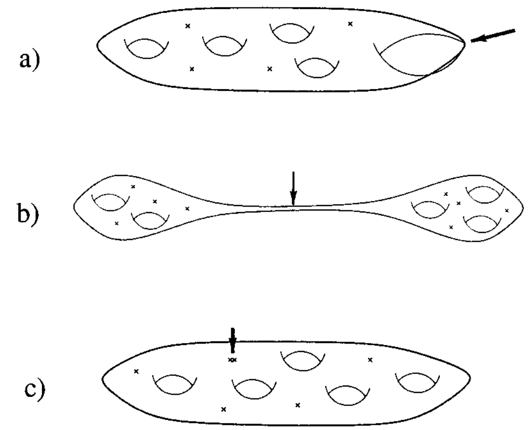


FIGURE 3. Degenerating Riemann surfaces contributing to the holomorphic anomaly

The choices of the f, $v^{i,a}$, $\tilde f^i_{kl}$ are by no means independent, however. If we specialize to the one-modulus case, as in our main examples the quintic and the non-compact $\mathbb{P}^2$, we can in addition set $\tilde f^z_{zz} = 0$ and get in the holomorphic limit the simplified expression

$$\begin{aligned} S^{zz} &= \tfrac{1}{F^{(0)}_{zzz}}\left[2\partial_z \log(e^K|f|^2) - (G_{z\bar z}v)^{-1}\partial_z(vG_{z\bar z})\right] \\ &= -\tfrac{1}{F^{(0)}_{zzz}}\partial_z \log\left(v\tfrac{\partial t}{\partial z}\right) \qquad \text{as } \overline{\lambda} \to 0, \end{aligned} \tag{35.52}$$

where $v \in T\mathcal{M}$ to render S^{zz} covariant[18]. Further potentials needed to solve for the $F^{(g)}$ in the global cases are $S \in \mathcal{L}^{-2}$, with $C_{\bar j\bar k\bar l} = e^{-2K} D_{\bar i} D_{\bar j}\bar\partial_{\bar k}\bar S$, and $S_{\bar i} := \bar\partial_{\bar i} S$, $\bar\partial_{\bar i} S^j = G_{i\bar i} S^{ij}$. $K^{ij} = -S^{ij}$, $K^{i\phi} = -S^i$ and $K^{\phi\phi} = -2S$ can be interpreted as propagators in the topological gravity theory, where ϕ is the dilaton, the first descendant of the puncture operator. S^i is derived from

$$\bar\partial_{\bar z} S^z = \frac{1}{F^{(0)}_{zzz}}\bar\partial_{\bar z}\left[2\partial_z \log(e^K|f|^2)^2 - v^{-1}\partial_z(v\partial_z K)\right].$$

In the local case there is a choice of the different holomorphic ambiguities so that $K^{i\phi}$ and $K^{\phi\phi}$ vanish.

The derivation of $F^{(g)}$ proceeds recursively. One first considers the holomorphic anomaly equation of $F^{(g)}$ and, using $\bar C^{ij}_{\bar i} = \bar\partial_{\bar i} S^{ij}$, one can write the

[18]There is a considerable simplification in the second line for the local case due to the fact that K is constant in the holomorphic limit and f can chosen to be constant as well.

right-hand side, e.g., for $g=2$, as

$$\frac{1}{2}\bar{\partial}_{\bar{i}}\left[S^{jk}(F^{(1)}_{jk}+F^{(1)}_{j}F^{(1)}_{k})\right]-\frac{1}{2}S^{jk}\bar{\partial}_{\bar{i}}\left[F^{(1)}_{jk}+F^{(1)}_{j}F^{(1)}_{k}\right].$$

Using the definition of the Riemann tensor as commutator, as adapted to special geometry, i.e., $[\bar{\partial}_{\bar{i}},D_j]^l_{\ k}=-G_{\bar{i}j}\delta^l_k-G_{\bar{i}k}\delta^l_j+C_{jkm}\bar{C}^{ml}_{\bar{i}}$, one lets the $\bar{\partial}_{\bar{i}}$ derivative act on $F^{(g-1)}$ and repeats the procedure until an expression $\bar{\partial}_{\bar{i}}F^{(g)}=\bar{\partial}_{\bar{i}}[\cdots]$ is derived, where the $[\cdots]$ contains the propagators and lower-genus correlation functions and the holomorphic ambiguity f_g. In the genus 2 case this yields for the local one parameter models the expression[19]

$$F^{(2)}_{\rm top} = -\frac{1}{8}S_2^2F^{(0)}_{,4}+\frac{1}{2}S_2F^{(1)}_{,2}+\frac{5}{24}S_2^3(F^{(0)}_{,3})^2 -\frac{1}{2}S_2^2F^{(1)}_{,1}F^{(0)}_{,3}+\frac{1}{2}S_2(F^{(1)}_{,1})^2+f_2. \tag{35.53}$$

The pictorial representation of this equation can be seen below.

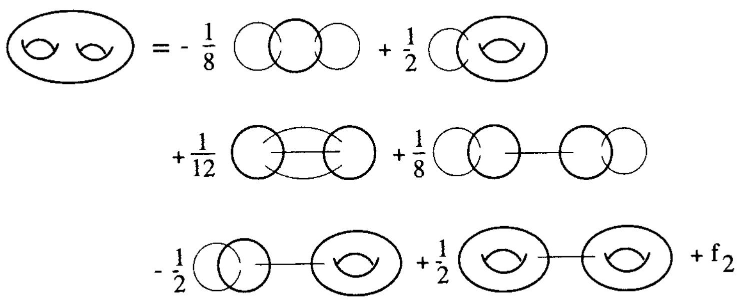


FIGURE 4. Pictorial reduction of the genus 2 vacuum amplitude to boundary contributions involving lower genus correlators for the local case.

Note that relative to the compact case the six correlators involving the descendant of the dilaton disappear in the non-compact case.

35.6.2. Extracting the n^g_d Invariants. Once the holomorphic ambiguity is fixed we can take the holomorphic limit Eq. (35.37) and use the mirror map Eq. (35.35) to calculate $F^{(g)}_{\rm top}(t)$. This is to be compared with the general expansion of the amplitudes in terms of the GV invariants n^g_d

[19]More general expressions can be found in the cited literature, see Ch. 40.

discussed in Ch. 32 and Ch. 33:

(35.54)
$$F(\lambda)=\sum_{g=0}^{\infty}F_g\lambda^{2g-2}=\sum_{\{d_i\},r\geq 0,k>0}n^g_{d_i}\frac{1}{k}\left(2\sin\frac{k\lambda}{2}\right)^{2r-2}\exp\left[-2\pi k\sum_i d_it_i\right].$$

Here d_i is the degree of the corresponding holomorphic curve w.r.t. a basis of two-forms k_i, $i=1,\ldots,h^{1,1}(M)$ spanning the Kähler cone. The invariants n^g_d defined by Eq. (35.54) are expected to be integers. This comes naturally from the physical interpretation of these numbers as five-dimensional BPS states associated to the M-theory two-brane wrapping the corresponding holomorphic curve in the Calabi–Yau threefold Ch. 32 and also suggests a new mathematical definition of the Gromov–Witten invariants. This was further explored in Ch. 33.

An important piece of information for fixing the holomorphic ambiguity is the large volume behavior of $F^{(g)}_{\rm top}$. In this limit only the degree 0 map survives. As for F_1 (cf. Eq. (35.5)) its contribution can be calculated for all genera directly in the A-model. The moduli space of the constant maps splits into the manifold itself and the moduli space of the worldsheet Riemann surface $\mathcal{M}_g$. The invariant is given by $\frac{1}{2}e(M)\int_{\mathcal{M}_g}c^3_{g-1}(H)$ where $e(X)=\int_M c_3$ denotes the Euler characteristic of M and H denotes the Hodge bundle (coming from the space of holomorphic one-forms on the Riemann surface) over the moduli space. Performing the second integration one gets

$$\langle 1\rangle^M_{g,0}=(-1)^g\frac{\chi}{2}\int_{\mathcal{M}_g}\lambda^3_{g-1}=(-1)^g\frac{e(M)}{2}\frac{|B_{2g}B_{2g-1}|}{2g(2g-2)(2g-2)!}. \tag{35.55}$$

Let us extract explicitly the contributions for genus 0, 1, 2, and 3 from Eq. (35.54),

$$F_0=-\frac{K^0t^3}{3!}+\frac{n}{2}t^2+\frac{t}{24}\int_X c_2J-i\frac{\chi}{2(2\pi)^3)}\zeta(3)+\sum_{d=1}^{\infty}n^0_d\,{\rm Li}_3(q^d), \tag{35.56}$$

where ${\rm Li}_n(q)=\sum_{k=1}^n\frac{q^k}{k^n}$. Here only the last term is a prediction of Eq. (35.54). The classical terms are added for completeness. Furthermore one gets

$$F_1=\frac{t\int c_2J}{24}+\sum_{d=1}^{\infty}\left(\frac{1}{12}n^0_d+n^1_d\right){\rm Li}_1(q^d), \tag{35.57}$$

$$F_2 = \frac{\chi}{5760} + \sum_{d=1}^{\infty} \left(\frac{1}{240} n_d^0 + n_d^2 \right) \mathrm{Li}_{-1}(q^d), \tag{35.58}$$

$$F_3 = -\frac{\chi}{1451520} + \sum_{d=1}^{\infty} \left(\frac{1}{6048} n_d^0 - \frac{1}{12} n_d^2 + n_d^3 \right) \mathrm{Li}_{-3}(q^d). \tag{35.59}$$

etc.

35.6.3. Fixing the Holomorphic Ambiguity for the Quintic. In the literature on the n_d^g the holomorphic ambiguities were fixed up to genus 4. F_g is a section of $\mathcal{L}^{2g-2}$. It is split into a part that comes from the integration of the anomaly equation F_g^a and the holomorphic ambiguity f_g. As explained in Sec. 35.6.1 this is not an invariant split. It is affected by the holomorphic gauge choice in the definition of the propagators. However we can fix this gauge so that F_g^a is regular at $z = \infty$ and $\Delta_{con} = 0$ and has the correct large volume behaviors as in Eqs. (35.5, 35.55). As $z = \infty$ is a non-singular point we make the ansatz

$$f_g(\psi) = \sum_{k=0}^{2g-2} \frac{A_k^g}{\Delta_{con}^k}. \tag{35.60}$$

To determine the A_k is sufficient for fixing the holomorphic ambiguity. This was done for the quintic up to genus four using the vanishing of low-degree holomorphic curves and the fact that the A_{2g-2} can be fixed using the target space physics.

35.6.4. Results and Enumerative Predictions for the Quintic. The n^g were calculated for the quintic up to genus 4. In fact all n_d^g were found to be integers. Many of the higher-genus enumerative results have been checked explicitly from the algebraic geometry approach to curve counting, in particular using the moduli space of the two-branes wrapping the holomorphic curves, see Ch. 33.

d	n_d^0	d	n_d^1
1	2875	1	0
2	609250	2	0
3	317206375	3	609250
4	242467530000	4	3721431625
5	229305888887625	5	12129909700200
6	248249742118022000	6	31147299733286500
7	295091050570845659250	7	71578406022880761750
8	375632160937476603550000	8	154990541752961568418125

d	n_d^2	d	n_d^3
1	0	1	0
2	0	2	0
3	0	3	0
4	534750	4	8625
5	75478987900	5	−15663750
6	871708139638250	6	3156446162875
7	5185462556617269625	7	111468926053022750
8	22516841063105917766750	8	130346459840858345500 0

d	n_d^4
1	0
2	0
3	0
4	0
5	15520
6	−7845381850
7	111468926053022750
8	25509502355913526750

35.6.5. Results and Enumerative Predictions for the Non-compact $\mathcal{O}(-3) \to \mathbb{P}^2$ Geometry. The ansatz Eq. (35.60) parametrizes the ambiguity for the $\mathcal{O}(-3) \to \mathbb{P}^2$ geometry. In the Klemm and Zaslow paper [**160**] it was fixed up to genus 4 using explicit A-model calculations for the Gromov–Witten invariants. For genus 5, additional information about the universality of A_8 was employed. The corresponding n_d^g are listed below.

d	n_d^0	d	n_d^1	d	n_d^2	d	n_d^3
1	$3^\diamond$	1	0	1	0	1	0
2	$-6^\diamond$	2	0	2	0	2	0
3	$27^\diamond$	3	$-10^\diamond$	3	0	3	0
4	$-192^\diamond$	4	$231^\diamond$	4	$-102^\diamond$	4	$15^\diamond$
5	$1695^\diamond$	5	$-4452^\diamond$	5	$5430^\diamond$	5	$-3672^\diamond$
6	-17064	6	80958	6	-194022	6	$290853^\diamond$
7	188454	7	-1438086	7	5784837	7	-15363990
8	-2228160	8	25301064	8	-155322234	8	649358826

d	n_d^4	d	n_d^5
1	0	1	0
2	0	2	0
3	0	3	0
4	0	4	0
5	$1386^\diamond$	5	$-270^\diamond$
6	$-290400^\diamond$	6	$196857^\diamond$
7	29056614	7	-40492272
8	-2003386626	8	4741754985

The diamond on the numbers above indicate that they have been checked using the direct BPS count from the cohomology of the Jacobian fibration as described in Ch. 33. For completeness we list below some higher-genus predictions, which were derived using only the latter approach.

d	n_d^0	d	n_d^1	d	n_d^2	d	n_d^3	d	n_d^4
1	0	1	0	1	0	1	0	1	0
2	0	2	0	2	0	2	0	2	0
3	0	3	0	3	0	3	0	3	0
4	0	4	0	4	0	4	0	4	0
5	21	5	0	5	0	5	0	5	0
6	-90390	6	27538	6	-5310	6	585	6	-28
7	42297741	7	-33388020	7	19956296	7	-9001908	7	2035271

35.7. Appendix A: Poisson Resummation

Let k, b be in the dual lattice to n; then by completeness $\sum_{n\in\mathbb{Z}} f(n) = \sum_{k\in\mathbb{Z}} \hat{f}(k)$ with $\hat{f}(k) = \int_{-\infty}^{\infty} f(n)\exp(2\pi ikn)dn$. Here $f(n) = \exp(-\pi nMn + 2\pi ibn)$ and $\hat{f}(k) = \frac{1}{\sqrt{\det(M)}}\exp[-\pi(k-b)M^{-1}(k-b)^\dagger]$. This leads for the zero mode part of the closed string partition function on the torus (35.8) to the following Poisson resummation:

$$
\begin{aligned}
I_F &= \tau_2 \sum_{n_i, m^i} \exp\left[2\pi i\tau_1 P - 2\pi\tau_2 H\right] \\
&= \tau_2 \sum_{n_i, m^i} \exp\left[-\pi n G^{-1}\tau_2 n + 2\pi i(\tau_1 m + i\tau_2 mBG^{-1})n\right. \\
&\qquad \left. -\pi\tau_2 m(G - BG^{-1}B)m\right] \\
&= A \sum_{k^j, m^i} \exp\left[-\frac{\pi}{\tau_2}(kGk - 2\tau_1 kGm + 2i\tau_2 kBm + (\tau_1^2 + \tau_2^2)mGm)\right] \\
&= A \sum_{k^j, m^i} \exp\left[-2\pi it(m^2k^1 - k^2m^1) - \frac{\pi A}{\tau_2\sigma_2}|k^2 + m^2\tau + \sigma(k^1 + m^1\tau)|^2\right] \\
&= A \sum_{k^j, m^i} \exp\frac{\pi i}{2\tau_2\sigma_2}\left[t|(k^2 + k^1\sigma) + \bar{\tau}(m^2 + m^1\sigma)|^2\right. \\
&\qquad \left. + \bar{t}|(k^2 + k^1\sigma) + \tau(m^2 + m^1\sigma)|^2\right].
\end{aligned}
$$

CHAPTER 36

Some Applications of Mirror Symmetry

In this chapter we discuss some applications of mirror symmetry and topological string amplitudes. The first application is the idea to represent quantum field theories by strings propagating on a nearly singular background. This is known as geometric engineering of quantum field theories. Mirror symmetry in this context becomes a powerful tool in studying aspects of gauge theory dynamics, as we will discuss below. Another application involves studying the large N limit of gauge theories. As we will discuss, certain $U(N)$ gauge theories turn out to be equivalent to topological strings in a rather subtle way.

36.1. Geometric Engineering of Gauge Theories

Singularities of the target space of string theory have an interpretation as giving rise to gauge theories in target space. In this context, mirror symmetry allows one to compute certain quantities of interest in the gauge theory in question. This idea, and more generally the idea of identifying string propagation in geometric singularities with certain gauge theories, is known as "geometric engineering" of gauge theories.

Geometric engineering is a broad topic, and here we will only give a very brief introduction to it. To keep the discussion focussed we will consider one specific class of examples.

36.1.1. ADE Singularities and Gauge Groups. Let us start with the connection between ADE singularities of K3 and ADE gauge groups. As discussed in Sec. 6.6, the ADE singularities of K3 (or more precisely, resolutions or deformations of $\mathbb{C}^2/\Gamma$ where Γ is a discrete subgroup of $SU(2)$) give rise to (vanishing) two-cycles whose intersection matrix is captured by a corresponding (affine) Dynkin diagram. For example, consider a deformation of an A_1 singularity, which can be represented as the hypersurface in $\mathbb{C}^3$

defined by

$$x^2 + y^2 = \prod_{i=1}^{2} (z - a_i). \tag{36.1}$$

Take the a_i to be real. The real parts of x, y can be viewed as the circle fiber in an S^1 fibration over the z-plane that degenerates at two points $z = a_i$. This space has one compact two-cycle, which can be seen by thinking of the sphere as an S^1 fibration over an interval, with the fiber vanishing at the ends (more generally, the A_n singularity has n compact two-cycles, and the intersection matrix of these cycles is given by the Cartan matrix for $A_n = SU(n+1)$).

Now consider this A_1 geometry as a non-compact target space for (type IIA) superstring theory. When type IIA string theory is compactified on this A_1 space, the low-energy six-dimensional theory has a gauge field giving a gauge group of $U(1)$ (more generally, the number of gauge fields is the number of normalizable elements of the second cohomology H^2. By the amount of supersymmetry preserved by this compactification, this gauge field is actually a member of an $\mathcal{N} = 2$ vector multiplet A_0. There are also two other vector multiplets A_+ and A_- in the low-energy theory from D2-branes and anti-D2-branes wrapping the $\mathbb{CP}^1$ and these vector multiplets carry plus/minus charge under the $U(1)$ gauge group associated to A_0. In the limit where the area of the $\mathbb{CP}^1 \to 0$ at fixed string coupling, the vector multiplets A_+, A_- become massless and combine with A_0 to generate an $SU(2)$ gauge group. In this way, Type IIA on an A_1 singularity gives rise to the corresponding $SU(2)$ gauge group in six dimensions. There is a similar story for the A_n, D_n, E_6, E_7, E_8 singularities /groups. For a physicist, this is the meaning of the magical relation between ADE singularities and ADE gauge groups.

This now suggests new ways to manufacture non-abelian gauge theories from abelian gauge theories in the presence of geometric singularities. This approach can be used to translate gauge theory questions to geometric questions: all we have to do is to understand precisely what geometry gives rise to the gauge theory of interest and what the dictionary is between the data encoded in the geometry and that of the gauge theory. There are subtleties to this story — for example we do not know the singularity corresponding to an arbitrary gauge theory with arbitrary matter content, and conversely there exist singularities that give rise to new unrecognizable field theories.

We shall avoid these subtleties in this short discussion and focus on one example: the case of (pure) $\mathcal{N} = 2$ supersymmetric $SU(2)$ gauge theory in four dimensions, studied by Seiberg and Witten and also familiar in the context of Donaldson theory. How do we engineer this gauge theory? We know that Type IIA compactified on a Calabi–Yau threefold gives rise to an $\mathcal{N} = 2$ theory in four dimensions. Since we have seen that an $SU(2)$ gauge theory in six dimensions comes from a shrinking $\mathbb{CP}^1$, we see that we must have a Calabi–Yau threefold that contains an A_1 cycle $\mathbb{CP}^1$ fibered over some space. The simplest example (a trivial fibration) is a $\mathbb{CP}^1 \times \mathbb{CP}^1$ in a Calabi–Yau threefold. Let us call the (class of the) fiber F and the base B. When F shrinks, we get an $SU(2)$ theory in four dimensions (of course there are gravitons, etc. in the four-dimensional effective theory, but these can be decoupled by scaling the Planck mass to infinity). The effective action, in genus 0, is of the form

$$\int \tau(a) F^+ \wedge F^+ \tag{36.2}$$

where the gauge coupling $\tau = \partial^2 F_0 / \partial a^2$ is the second derivative of the genus 0 prepotential F_0, as was noted before. The coupling constant of the gauge theory is given by $\tau = 1/g^2 = \mathrm{vol}(B)$ and a is the Kähler parameter of the fiber F, as can be seen from the compactification ansatz. We have here two Kähler moduli, one for the base and the other for the fiber, and both correspond to fields in four dimensions (as discussed earlier, it is natural to think of the Calabi–Yau moduli as varying over the four-dimensional space, thus being interpreted as four-dimensional fields). From a gauge theory point of view, $1/g^2$ is viewed as a parameter. We can take the limit where the volume of the base B goes to infinity, so that the dynamics of the field $1/g^2$ get frozen out and we can view it as a parameter of the field theory. We then end up with a being the scalar in a vector multiplet associated to $U(1)$ in $SU(2)$, and $1/g^2$ the coupling constant parameter that depends on the size of B.

Seiberg and Witten found by some consistency arguments that, from the perspective of the four-dimensional gauge theory, the gauge coupling $\tau(a)$ is the sum of a perturbative part (coming from a one-loop gauge theory computation) and a non-perturbative part (summarizing the effects of gauge theory instantons). The instanton contribution is very difficult to compute and was determined only indirectly by using these various consistency arguments.

In the current setting, since we know that τ is related to the genus 0 prepotential of the topological string, it can be computed using mirror symmetry! This in turn can be viewed as contributions from holomorphic curves wrapping two-cycles in the Calabi–Yau. We find that the perturbative part of the gauge coupling is reproduced by holomorphic curves in the class $[F]$. The instanton contribution is reproduced by holomorphic curves wrapping the class $[B]$ multiple times. Note that this is consistent with the fact that the instanton number n contribution is expected to scale like $e^{-n/g^2} \sim e^{-n\mathrm{vol}(B)}$. So we have an alternate way of computing the contribution of gauge theory instantons by reformulating them in terms of topological strings and using mirror symmetry to compute them. This approach demonstrates directly the connection between topological string computations and Donaldson invariants.

This approach, which gives a flavor of the uses of mirror symmetry, can be generalized to other gauge groups, with or without extra matter fields. It is usually applied in the context of non-compact Calabi–Yau's, which are the relevant case for field theories (only the geometry near the singularity of the fiber is needed). Recall that these are also the easiest cases for establishing the predictions of mirror symmetry.

36.2. Topological Strings And Large N Chern–Simons Theory

Consider a field theory in some dimension that includes as its field variable some $N \times N$ matrix-valued fields where the action is invariant (at least) under global conjugations by unitary $N \times N$ matrices acting on the fields. An important example of this class of theories includes $SU(N)$ gauge theories (possibly coupled to matter). An interesting large N limit of theories of this kind can be considered. Let us try to motivate this limit from the simplest example, namely a zero-dimensional quantum field theory of an $N \times N$ matrix X with action $S = (1/g^2)\mathrm{tr}(X^2 + X^3)$.

This theory has an obvious $U(N)$ invariance, which is given by the conjugation of the matrix X. In this way, X can be viewed as belonging to the representation of $U(N)$ given by the tensor product of the fundamental representation and its conjugate. The idea of trying to take a large N limit is, to extract the N-dependence of the correlation functions of $U(N)$-invariant

correlation functions from the Feynman diagrams. Viewing X as the adjoint representation of $U(N)$, i.e., as the tensor product of the fundamental and anti-fundamental representations, we can adopt an oriented double-line representation for edges of Feynman graphs with each line carrying an index that transforms in the fundamental or anti-fundamental representation, depending on the orientation. The edges of the Feynman diagram, i.e., the propagators, would thus be represented as depicted in Fig. 1.

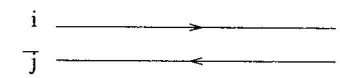


FIGURE 1. The edges of the Feynman diagram for $U(N)$ gauge theories

The cubic interaction will be represented as shown in Fig. 2.

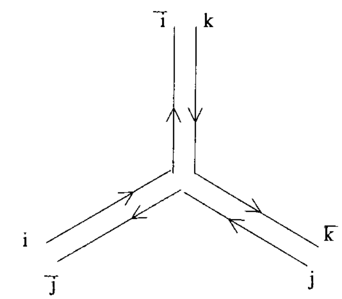


FIGURE 2. The cubic interaction vertex

A Feynman graph is depicted in Fig. 3. These are also sometimes referred to as "ribbon graphs".

Recalling the rules for the weight of the Feynman graphs, each propagator gets weighted with a factor of g^2, each vertex with a factor of $1/g^2$ and each boundary circle involves a trace over an index in the fundamental (or anti-fundamental) represention and thus gets weighted with a factor of N. We see that the g- and N- dependence of a graph with E edges (propagators), V vertices (cubic interaction vertex) and F faces (or holes) is captured in a universal way by

$$(g^2)^E(1/g^2)^V N^F = (Ng^2)^{E-V} N^{F-E+V}$$

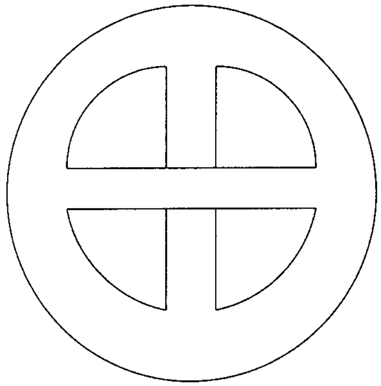


FIGURE 3. Feynman or "ribbon" graphs suitable for large N expansion

with a prefactor that is independent of g and N. Consider filling each of the holes of the Feynman graphs. Filling in the boundaries will give a closed, orientable Riemann surface. For example the graph depicted in Fig. 3 gives rise to a torus when the holes are filled.

EXERCISE 36.2.1. *Verify that by filling the holes one obtains an orientable Riemann surface and for the case depicted in Fig. 3 it has genus* 1.

Using the fact that

$$F - E + V = \chi = 2 - 2r,$$

where χ is the Euler characteristic of the Riemann surface obtained by filling in the faces and r is the genus, and defining $t = g^2 N$, we can rewrite the t- and N- dependence of the amplitude as

$$N^{2-2r} t^{2r-2+F}.$$

From this weight factor it is natural to consider a particular regime of the parameters where we take $N \gg 1$ with t fixed (i.e., with $g^2 \sim 1/N \ll 1$). In this limit the graphs corresponding to low-genus r dominate in the computation. The leading contribution comes from graphs that have the topology of S^2. It is not too hard to see that the gauge theory diagrams that give rise to genus 0 surfaces are planar diagrams (i.e., diagrams that can be drawn in the plane without self-crossings). Note that the computation will still involve a non-trivial function of t for each genus because the numbers of faces

F can be essentially arbitrary for each genus. So for example the partition function organizes itself in the form

$$F(t, N) = \sum_r N^{2-2r} f_r(t).$$

We can also view this expansion as a small $\lambda = g^2$ expansion and write it as

$$F(t, \lambda) = \sum_r (\lambda/t)^{2r-2} f_r(t) = \sum_r \lambda^{2r-2} F_r(t)$$

where $F_r(t) = f_r(t)/t^{2r-2}$. This final form looks very much like what one expects to find for the partition function of a string theory, where λ is identified with the string coupling constant and t is some kind of modulus of the target space. Note that even though we discussed the scalings in the context of zero-dimensional QFTs the organization of Feynman diagrams and their g, N dependence is the same for any dimension (e.g., it does not change if X is an adjoint-valued function on a manifold).

't Hooft's conjecture states that for the large N limit of field theories there is such a dual description. Note that the validity of this conjecture is a priori not very clear. There are holes in the Riemann surface that we are filling in to get a compact Riemann surface without boundaries. One might think, therefore, that it is more natural to consider Riemann surfaces with punctures, instead of filling in the holes. But the conjecture is that somehow in a dual description the holes have disappeared!

The first concrete example of 't Hooft's conjecture came from the work of Kontsevich, who showed that a matrix model like the one we considered above gave rise to graphs that could be viewed as triangulating the moduli space of Riemann surfaces. In this case, $F(t, \lambda)$ computes intersection forms involving Mumford classes on the moduli space of Riemann surfaces, allowing the large N matrix model to be interpreted as two-dimensional, pure topological gravity.

In recent years, similar, but much more sophisticated, examples have emerged. For instance, the large N limit of $\mathcal{N} = 4$ supersymmetric $SU(N)$ Yang-Mills theory on S^4 was shown to be equivalent to Type IIB closed string theory on $\mathrm{AdS}_5 \times S^5$ (this is what is referred to as the AdS-CFT correspondence).[1] This duality arises by considering N coincident D3-branes of Type IIB, whose dynamics is described by $SU(N)$ Yang-Mills theory

[1] Here AdS_k denotes "anti-de Sitter" k-space, i.e., the k-dimensional Minkowski-signature metric solving Einstein's equations with a negative cosmological constant.

on the worldvolume S^4. In the large N limit, this geometry changes by surgery. The D-branes disappear, inducing an $\mathrm{AdS}_5 \times S^5$ geometry. This is a realization of 't Hooft's conjecture relating large N Yang-Mills to a closed string theory in extra dimensions. Even though verifying this conjecture is beyond our present day ability, there has been a large body of evidence supporting it.

Returning to the main theme of the present book, we have considered topological strings on Calabi–Yau manifolds. We can ask the reverse question: Could this string theory be dual to some large N gauge system?

In order to probe this question it turns out that we have to consider D-branes in the context of topological theories. Recall that in the context of topological A-models, consistency conditions dictate that the D-branes lie on Lagrangian submanifolds of the target. Thus in the presence of D-branes we consider holomorphic maps $\Sigma \to M$ such that the boundary $\partial\Sigma$ maps to a Lagrangian submanifold of M. An example of this situation is given by the Calabi–Yau threefold known as the conifold, given by T^*S^3 which we have discussed before. This space has a Lagrangian S^3 (the zero section) on which we can wrap, say, N D3-branes (or what a physicist would call N Euclidean D2-branes) . The corresponding computation is almost trivial: there are apparently no holomorphic maps with boundary on the S^3, and the computation localizes to maps that are degenerate (zero-area) ribbons which explicitly realize the Feynman diagrams of a gauge theory known as the $SU(N)$ Chern–Simons theory on the S^3. The meaning of the foregoing is that the topological A-model on T^*S^3 with N D3-branes on the S^3 has the interpretation in target space physics as including a sector involving Chern–Simons theory on S^3.

The action for the Chern–Simons gauge theory is given by

$$S_{CS} = \frac{ik}{8\pi}\int_{M^3} \mathrm{Tr}\left(\frac{1}{2}A \wedge dA + \frac{1}{3}A \wedge A \wedge A\right). \tag{36.3}$$

This action makes sense on any three-manifold, independent of a choice for the metric. It is a topological field theory. For consistency, k has to be an integer, as global gauge transformations shift $S \to S + 2\pi i n k$, so that the path-integral including the weight e^{-S} is gauge invariant only if k is an integer. Note that the consistency of the quantum field theory requires only that $\exp(-S)$ be single-valued, and not necessarily S itself, as that is what appears in the path-integral. Note also that the Chern–Simons theory

is topological, in the sense that it does not depend on the metric on the manifold and for example (at least formally) the partition function depends only on the topology of the (real) three-manifold. In the context of D-branes wrapping, a Lagrangian submanifold $2\pi i/k$ gets identified classically with the string coupling constant λ, which receives a quantum correction shifting the identification to $\lambda = 2\pi i/(k+N)$.

More generally, the statement about the equivalence of the topological string theory in the presence of D-branes to Chern–Simons theory is true with S^3 replaced by an arbitrary Lagrangian three-manifold L inside a Calabi–Yau threefold. The target space field theory will include a Chern–Simons theory on L. However, in general, if there are non-trivial two-cycles, there are also corrections to the Chern–Simons action coming from finite-action holomorphic maps from Riemann surfaces with boundaries to the target geometry, where the boundary is mapped to L. In particular, for any such map one obtains a correction to the partition function of the Chern–Simons theory of the form

$$e^{-\int_\Sigma f^*(k)}\mathrm{tr}(P\exp\int_{\gamma_1} A)\cdots\mathrm{tr}(P\exp\int_{\gamma_m} A),$$

where $f^*(k)$ is the pull-back of the Kähler form, γ_i represent one-cycles on L, and $P\exp(\cdots)$ denotes the holonomy of the gauge connection along the corresponding cycle.

Even though we do not need it for the purposes of this section, let us briefly comment on the B-model version of the Chern–Simons theory. This will be useful in applications of mirror symmetry to D-branes. In particular, we would be able to translate the question of counting holomorphic maps from Riemann surfaces with boundaries to some simpler computation on the B-model side. The topological B-model is related to the holomorphic version of Chern–Simons theory. For example, consider N D6-branes wrapping a Calabi–Yau threefold. Then in the target theory we obtain an action

$$S = \frac{ik}{16\pi}\mathrm{tr}\int(\overline{A}\partial\overline{A} + \frac{2}{3}\overline{A}^3)\wedge\Omega,$$

where Ω is the holomorphic three-form and $\overline{A}$ denotes a $(0,1)$-form representing the connection of a holomorphic rank N bundle.

EXERCISE 36.2.2. *Recall that in the B-model one can also have other even-dimensional D-branes, i.e., D0-, D2- and D4-branes. Write the corresponding action in the target space by "reducing" the above holomorphic Chern–Simons. By this reduction one means replacing the components of the connection normal to the D-brane by a section of the normal bundle on the D-brane.*

Let us return to our discussion of the A-model D-branes and the Chern–Simons theory on them. The correlation functions of Chern–Simons theory can be exactly computed by relating it to two-dimensional CFT. In particular the Hilbert space of the $SU(N)$ theory with coupling constant k can be related to the conformal blocks of $SU(N)$ affine Kac–Moody algebra at level k. As a consequence, the partition function of $SU(N)$ Chern–Simons theory on S^3 at level k can be computed exactly.

Now returning to 't Hooft's conjecture, we want to ask, what is the closed string theory that is the large N limit of the Chern–Simons theory on S^3? Since the Chern–Simons theory has an alternate description as topological string theory on T^*S^3 in the presence of N D2-branes on S^3, the general idea that the large N dual involves a geometric transition suggests that in the large N limit the D3-branes (and the S^3 on which they live) disappear, and the geometry develops a compact S^2. In other words, we go to a limit where the conifold develops a singularity and then the singularity is resolved by blowing up an S^2 (see Fig. 4).

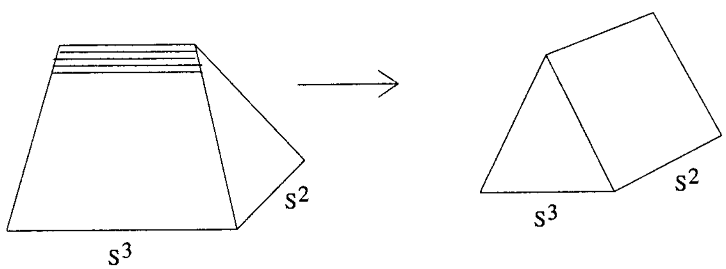


FIGURE 4. Conifold singularity resolved by blowing up S^2

Here the D-branes would have disappeared, as there is no non-trivial compact three-cycle.[2]

[2] Notice that we have A-model topological strings on both sides of the duality; this is *not* mirror symmetry.

It turns out that this guess is in fact correct. More precisely, the $SU(N)$ Chern–Simons theory on S^3 is equivalent to the closed string topological A-model on the total space of $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ over $\mathbb{CP}^1$. The size of the $\mathbb{CP}^1$ (the complexified Kähler class) is given by $t = \lambda N = 2\pi i N/(k+N)$, where $\lambda = 2\pi i/(k+N)$ and the free energies of the Chern–Simons theory and the topological A-model are related by $F(N,k) = F(t,\lambda)$. Also note that the "area" of the $\mathbb{CP}^1$ is imaginary: $t = 2\pi i N/(k+N)$. This should be familiar by now — since t is the complexified Kähler class, so while the area proper is zero, there is a nonzero B-field turned on along the S^2. This conjecture can be explicitly checked because both sides are computable exactly. As mentioned before, the partition function of Chern–Simons theory is computable and the partition function of topological strings on $\mathcal{O}(-1) \oplus \mathcal{O}(-1)$ over $\mathbb{CP}^1$ has already been discussed.[3]

So far we have talked about the partition function on both sides. It is natural to extend this to the correlation functions on both sides. On the Chern–Simons side there are Wilson-loop observables in the theory, related to knot invariants. These correspond to insertions in the path-integral of the trace of the holonomy of the connection along the knot (or more generally along links). In other words, the Wilson-loop observables of Chern-Simons theory are obtained by taking a knot γ in S^3 and defining $U_\gamma = P\exp(\int_\gamma A)$, the path-ordered exponential of the holonomy of the gauge field around the loop. The observable is then $\operatorname{tr} U_\gamma$, the trace being taken in the fundamental representation. We can then consider arbitrary correlation functions of these observables, such as

$$\langle \operatorname{tr} U_{\gamma_1} \cdots \operatorname{tr} U_{\gamma_m} \rangle.$$

Again it turns out that there are well-developed methods to compute these correlation functions on the Chern–Simons side and they give rise to knot (or link) invariants. (The results are invariants, as the integral is over all gauge connections; so the correlations only depend on the link and the representations of the group involved.) We would like to know what these are computing on the dual resolved conifold in the context of the topological A-model.

We know that the Chern–Simons theory on S^3 is the same as the topological A-model on T^*S^3 in the presence of D-branes on the S^3. We can

[3] This duality is expected to generalize to all compact three-manifolds with positive curvature — lens spaces, etc.

associate, in a canonical way, to each knot γ on S^3 a Lagrangian cycle $S_\gamma \subset T^*S^3$ that intersects S^3 in the knot γ.

EXERCISE 36.2.3. *Show how this can be done by using the conormal bundle of γ, i.e., the subset of T^*S^3 over γ that is zero on $T\gamma$.*

The topology of this cycle is $\mathbb{R}^2 \times S^1$, and since it is Lagrangian, it satisfies the A-model boundary conditions, and so we can allow D-branes (say, M of them) to lie along S_γ. In total, we have N D-branes on S^3 (with an $SU(N)$ Chern–Simons theory describing their dynamics) and M D-branes on S_γ (and correspondingly an $SU(M)$ Chern–Simons theory on their worldvolume), and the two sets of D-branes intersect along the knot γ. Now there is also a new sector of open strings that is allowed, stretching from one set of D-branes to the other, and transforming in the bi-fundamental representation of $U(M) \times U(N)$. Eliminating these extra fields is equivalent to inserting in the gauge theory path-integrals

$$Z(U,V) = \exp\Big(\sum_{n=1}^{\infty} \frac{1}{n} \mathrm{tr}\, U^n \mathrm{tr}\, V^{-n}\Big) \tag{36.4}$$

where $U = P\exp(\int_\gamma A)$, $V = P\exp(\int_\gamma \tilde{A})$, where A and $\tilde{A}$ are the Chern–Simons gauge fields on S^3 and S_γ respectively.

EXERCISE 36.2.4. *This arises by a computatation in a one-dimensional QFT identified with two real bosons propagating on the knot γ. One obtains for the partition function of this theory $-\log(\det(d + A + \tilde{A})$. By doing the explicit computation of this determinant, demonstrate the above formula.*

We wish to view the D-branes on the non-compact Lagrangian submanifold S_γ as spectators (which means we wish to treat them as classical fields). In other words we want to integrate out the degrees of freedom living on the D-branes on the S^3 to arrive at an effective theory for the D-branes on S_γ. We define $F(t,V)$ by

$$\begin{aligned} \exp(-F(t,V)) =& \frac{1}{\int [DA] \exp(-S_{CS}(A;S^3))} \\ & \times \int [DA] \exp\left[-S_{CS}(A;S^3) + \sum_{n=1}^{\infty} \frac{1}{n} \mathrm{tr}\, U^n \mathrm{tr}\, V^{-n} \right] \\ =& \langle Z(U,V) \rangle_{S^3}. \end{aligned} \tag{36.5}$$

which is the generating functional for all observables in the Chern–Simons theory on S^3 associated to the knot γ. The effective theory on the branes on S_γ is given by $S = S_{CS}(\tilde{A}; S_\gamma) + F(t,V)$.

How is this reflected on the dual side with blown up $\mathbb{CP}^1$? Since we have added extra non-compact D-branes to the original geometry, we would expect to see some D-branes left over in the geometry on the other side of the transition, after taking the large N limit. The question of which Lagrangian submanifold one obtains on the other side is not known for the general case. For some special cases (such as torus knots) it is known. For example, consider the trivial knot (or the "unknot") in S^3. In this case, S_γ can be characterized as the fixed-point set of an involution on T^*S^3. On continuing to the other side, we end up with a Lagrangian cycle with topology $\mathbb{R}^2 \times S^1$ where S^1 is the equator of the $\mathbb{CP}^1$ and the $\mathbb{R}^2$ encompasses two of the four non-compact directions.

EXERCISE 36.2.5. *Demonstrate this by finding an anti-holomorphic involution on the total space of $\mathcal{O}_{\mathbb{CP}^1}(-1) \oplus \mathcal{O}_{\mathbb{CP}^1}(-1)$.*

As far as the effective theory on L is concerned, this should be the Chern–Simons theory, with corrections coming from the fact that now there are holomorphic maps of Riemann surfaces into the space with boundary on L. For instance, in genus 0, these are just disks that cover the northern or southern hemisphere of the $\mathbb{CP}^1$ with boundary on the equator. The contribution of these holomorphic maps from the topological string side has the general structure

$$F(t,V) = \sum_{g=0}^{\infty} \sum_{h=1}^{\infty} \sum_{n_1,\dots,n_h}^{\infty} \lambda^{2g-2+h} F_{g;n_1,\dots,n_h}(t) \mathrm{tr}\, V^{n_1} \cdots \mathrm{tr}\, V^{n_h} \tag{36.6}$$

where $F_{g;n_1,\dots,n_h}$ is the topological string amplitude on a genus g surface with h boundaries. The factors $\mathrm{tr}\, V^n$ correspond to the worldsheet boundary wrapping the equator of the $\mathbb{CP}^1$ n times. In general it is difficult to compute these numbers (from the mathematical side the technique for doing so is in its nascent stages). However, for some simple cases, such as the unknot, one can use a simple physical reasoning to compute it and compare it to the Chern–Simons prediction.

The Chern–Simons computation yields (after suitable analytic continuation, so that we can compare to results on the other side)

$$F(t,V) = -i\sum_{n=1}^{\infty}\frac{\operatorname{tr}V^n + \operatorname{tr}V^{-n}}{2n\sin(n\lambda/2)}e^{-nt/2}. \tag{36.7}$$

We now wish to compare this with the predictions of the topological strings on the side involving a blown-up $\mathbb{CP}^1$. To do this one uses a target-space (M-theory) interpretation of the topological string, similar to what we developed in the context of Gopakumar–Vafa invariants. In this description,we have D-branes ending on the equator. From a Schwinger-type computation, one finds the general structure

$$F(t,V) = i\sum_{n=1}^{\infty}\sum_{\mathcal{R},Q,s}\frac{N_{\mathcal{R},Q,s}}{2n\sin(n\lambda/2)}e^{n(is\lambda - tQ)}\mathrm{Tr}_{\mathcal{R}}V^n, \tag{36.8}$$

where Q is the charge of the D2-brane, $\mathcal{R}$ is the representation of a "dual gauge group", s is the spin of the particle and $N_{\mathcal{R},Q,s}$ is the net number of charged states with these quantum numbers. The integral structure that arises in this formula is unexpected from the viewpoint of the Chern–Simons theory or the theory of knot invariants. In the case at hand (the unknot) we have only two non-trivial D2-branes corresponding to the D2-branes wrapping the northern hemisphere and southern hemisphere. In particular $N_{\mathcal{R},Q,s}$ are vanishing, except when $Q = 1/2, s = 0$ and $\mathcal{R}$ corresponds to the fundamental and the conjugate representation (corresponding to the northern and southern hemispheres respectively). This answer agrees with that coming from the Chern–Simons side. The computation has also been done for a more complicated torus knot on the Chern–Simons side and reexpressed in the form suggested by the topological string and (lo and behold) integral invariants arose as predicted.[4]

Clearly, direct techniques from topological string theory need to be developed to compute holomorphic maps with boundaries to the target geometry. The above results based on physical reasoning may provide a clue, as the expected answers are known for many cases using the large N duality of Chern–Simons with topological string theory. These techniques would yield a promising new insight into defining knot invariants.

[4]In fact the above integral invariants can be further refined.

CHAPTER 37

Aspects of Mirror Symmetry and D-branes

We consider how mirror symmetry acts on D-branes. The correspondence of these states in mirror theories leads to many predictions. In this context mirror symmetry relates, for example, classical integrals to "boundary" Gromov–Witten invariants counting holomorphic curves from discs. These D-brane states are described in mathematical terms, where the statement of mirror symmetry becomes equivalent to certain mathematical conjectures. This correspondence is interpreted mathematically by Kontsevich to imply a striking equivalence of categories. The D-brane correspondence under mirror symmetry also leads to the Strominger–Yau–Zaslow conjecture about the structure of Calabi–Yau manifolds with mirror symmetry (special Lagrangian fibrations) and how to find the mirror manifold. The mirror symmetry action on a D-brane is described in detail for the elliptic curve. We describe a mathematical technique for constructing a functor establishing the equivalence of brane categories on mirror manifolds.

37.1. Introduction

In this chapter we will explore some aspects of mirror symmetry that arise from looking at the transformation of D-branes under mirror symmetry. *N.B.:* As a self-contained treatment of all the assertions and results we present would require substantial length, we rely instead on a descriptive exposition. The interested reader can find references Ch. 40.

We have studied the classical version of mirror symmetry as predicting Gromov–Witten invariants, as well as its proof and generalization to non-Calabi–Yau manifolds via the gauged linear sigma model. We have also briefly talked about how mirror symmetry acts on D-branes in the simple case of T^2 in Ch. 19. Here we will see that incorporating D-branes into mirror symmetry leads to the SYZ conjecture about how to find the mirror manifold geometrically, in cases where the mirror of a Calabi–Yau is again a Calabi–Yau.[1] Some evidence supports the validity of this conjecture. Another point of view shows that D-brane correlation functions should reveal new enumerative invariants involving maps from Riemann surfaces with boundaries to Calabi–Yau manifolds, with the boundaries mapping to Lagrangian submanifolds. Finally, Kontsevich's approach formalizes the notion of D-branes as objects in a category. Mirror symmetry, then, is interpreted as an equivalence of these categories. We will describe this equivalence in detail for the case of an elliptic curve or torus (one-dimensional Calabi–Yau), extending our previous discussion of it in Ch. 19.

In the wake of the general physics proof of classical mirror symmetry covering Fano as well as Calabi–Yau manifolds, one can also find the mirror of D-branes in the Fano context (see the discussion in Ch. 39). We see that we are only nearing a complete mathematical understanding of how mirror symmetry acts on D-branes. In particular a precise mathematical understanding of D-brane categories is still lacking in the non-Calabi–Yau case, or Calabi–Yau manifolds whose mirror is not a Calabi–Yau.

37.2. D-branes and Mirror Symmetry

37.2.1. String Theory. The supersymmetric quantum field theories used in classical mirror symmetry are in fact super*conformal* field theories, invariant under local rescalings of the worldsheet metric.[2] This symmetry allows us to integrate over all two-dimensional metrics merely by integrating over the finite-dimensional space of conformal classes of metrics. This is what makes string theory possible (see the introduction to Ch. 31, e.g.). Four-dimensional states are created in two-dimensional language, and physical scattering amplitudes involve computing two-dimensional correlators and integrating over all possible two-dimensional geometries of Feynman graphs. Thus string theory is a prescription for how to perform the Feynman perturbation series. It is, by its nature, a "perturbative" formulation of physics — inadequate, in some respects.

Of course, the four-dimensional field theory which is an approximation to string theory at low energies must contain gravity (that's the point). It

[1]As we have noted before this is not always the case. For example rigid CYs have Landau-Ginzburg mirrors but no geometric mirror.)

[2]That they are classically conformal is immediate from the form of the bosonic action $\int \langle d\phi, d\phi\rangle \sqrt{g} d^2x$. Indeed, a local rescaling sends $g \to \Lambda(x)g$, $\sqrt{g} \to \Lambda\sqrt{g}$, and the Λ cancels the $\Lambda^{-1}g^{-1}$ used in the rescaled inner product $\langle d\phi, d\phi\rangle$.

also can contain gauge fields and matter. In such theories, it is natural to ask about non-perturbative states, such as magnetic monopoles. These are static configurations of gauge fields that are not connected to the trivial configuration. A localization of energy density allows us to think of these configurations as particles. Though they have finite mass, it would take an infinite-energy perturbation to create them from the vacuum. As another example, consider a $(1+1)$-dimensional field theory with a potential $V = (\phi^2 - 1)^2$. The vacua correspond to $\phi = \pm 1$, but there may also be finite-energy configurations with boundary conditions $\phi(-\infty) = -1$, $\phi(+\infty) = +1$, for example. Different connected components of finite-energy boundary conditions at spatial infinity determine different non-perturbative ("solitonic") sectors. (Solitons in Landau–Ginzburg theories were discussed in Ch. 18.) States with such boundary conditions are by nature non-perturbative, and their interactions with perturbative states demand a path-integral (not Hamiltonian) approach to quantum physics.

37.2.2. D-branes. What are the non-perturbative states of string theory? Which states correspond to non-perturbative states of the effective four-dimensional field theory from strings? Without a non-perturbative formulation of string theory, it might seem hopeless to try to answer these questions. However, Polchinski's discovery of D-branes changed our understanding of string theory dramatically. His main result is that the vacuum states of a conformal field theory on a Riemann surface with a boundary (and appropriate boundary conditions) can be thought of as non-perturbative states of the full closed string theory.[3]

Among the space of states in a non-perturbative sector are the minimal-energy states. In a supersymmetric theory, the central charge, often expressed as a total derivative, can be determined by boundary conditions at spatial infinity — so the non-perturbative sectors are evidenced by a non-vanishing central charge, Z. The central charge also provides a mass bound for states in that sector: $m \geq |Z|/2$ (true for Z as a matrix operator). This can be seen as follows. We write the $\mathcal{N} = 2$ supersymmetry algebra in the zero-momentum frame (rest frame, valid for massive states):

$$\{Q^i_\alpha, \overline{Q}^j_{\overline{\beta}}\} = 2\gamma^\mu_{\alpha\overline{\beta}} P_\mu \delta^{ij} = 2\delta_{\alpha\overline{\beta}}\delta^{ij} M,$$

[3]The discussion of D-branes in Ch. 19 and Ch. 39 focuses mainly on their worldsheet description and interpretation, while Ch. 32 contains the target space interpretation.

with

$$\{Q^i_\alpha, Q^j_\beta\} = Z^{ij}\epsilon_{\alpha\beta}$$

being the part of the algebra involving the central charge.[4] When $Z = 0$ the algebra looks like $2\mathcal{N} = 4$ pairs of creation and annihilation operators, so we expect 2^4-dimensional representations. However, when $Z \neq 0$, we can define

$$a_\alpha = \frac{1}{\sqrt{2}}\left(Q^1_\alpha + \epsilon_{\alpha\beta}(Q^2_\beta)^\dagger\right),$$
$$b_\alpha = \frac{1}{\sqrt{2}}\left(Q^1_\alpha - \epsilon_{\alpha\beta}(Q^2_\beta)^\dagger\right).$$

One finds $\{a_\alpha, a^\dagger_\beta\} = \delta_{\alpha\beta}(2M + Z)$ and $\{b_\alpha, b^\dagger_\beta\} = \delta_{\alpha\beta}(2M - Z)$, all others vanishing. Again, we get oscillators, but now positive definiteness requires

$$M \geq |Z|/2,$$

with the bound saturated by those states annihilated by either a's or b's, i.e., killed by half the supercharges. Such states are called "BPS states." Note that the dimension of BPS multiplets is smaller: 2^2.

So the BPS states preserve a fraction of the supersymmetry charges. In string theory, each space-time supersymmetry is a reflection of an $\mathcal{N} = 2$ supersymmetry on the worldsheet (the relation is via a spectral flow relating the NS and R sectors, which translates to a correspondence among space-time bosons and fermions). So space-time BPS states can be described by two-dimensional states invariant under a single $\mathcal{N} = 2$ left-right combination of supercharges.

Why should there be a closed string interpretation of an open string boundary condition? Let us recall the argument of Sec. 19.3. The reason is that on a two-dimensional, Euclideanized torus, we can quantize along constant slices of either of the two coordinates. Therefore, an open string path-integral in an A-B sector (A and B labeling boundary conditions) can be described by the diagram on the left in Fig. 1, where "time" runs vertically. However, choosing the other coordinate as "time," this can be calculated equivalently as the diagram on the right, and now the boundary conditions

[4]The notation here is as follows. Each Q labels a spinor supercharge, with subscript α a chiral index of the spinor representation, and superscript i labeling the supercharge, $i = 1, 2$. μ is a space-time index, γ^μ are Dirac matrices, P^μ is the momentum, M the mass ($P^0 = M$ in the rest frame), and $Z^{ij} = \begin{pmatrix} 0 & Z \\ -Z & 0 \end{pmatrix}$ the central charge.

A and B are represented as closed string states. This is known as the *closed string channel.*

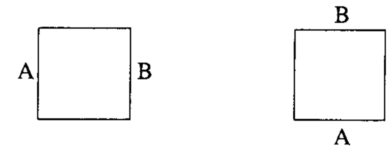


FIGURE 1. Quantization along two different coordinates allows us to express boundary conditions (A or B) as states in the closed string theory.

We should note that non-perturbative states were expected based on space-time solutions and other effective field theory reasoning, but the description as D-branes allows a *quantitative* analysis of these states via two-dimensional techniques. Thus D-branes have profoundly advanced our understanding of string theory. Perhaps the most striking discovery resulting from D-brane analysis is the exact calculation of black hole entropy — a purely quantum gravitational problem — for certain supersymmetric black holes.

Our focus will be on the characterization of D-brane boundary conditions.

37.3. D-branes in IIA and IIB String Theory

If we have two equivalent two-dimensional QFTs, this in particular means that they are equivalent on arbitrary Riemann surfaces, including those with boundaries. Recall, as in our discussion of D-branes in Ch. 19, that when there is a boundary on the worldsheet we should choose suitable boundary conditions. Thus if we have two equivalent theories, for each choice of boundary condition for one theory, we should get an equivalent choice for the other. Thus when we have two equivalent mirror theories, there is also a map for the corresponding D-branes, which is part of the data for the isomorphism of the two theories. One can also view this from the target space viewpoint. If we have two equivalent target theories, as would be the case for IIA, IIB superstrings compactified on mirror threefolds, all non-perturbative states of one should map to the other. In particular, D-branes should get mapped to D-branes.[5]

Recall that even though the operation of mirror symmetry at the level of the worldsheet is a matter of sign convention between left and right movers, at the level of realization of these theories as sigma models it has dramatic implications. The same is true for the map between the D-branes. Therefore, the conformal field theory description of corresponding D-brane states should be different.

37.3.1. A-Cycles and B-Cycles. Consider a bosonic action

$$S = S_{\text{closed}} + S_{\text{bdry}} = \int_D (\langle d\phi, d\phi\rangle dV + \phi^* B) + \oint_{\partial D} \phi^* A,$$

where $\phi : D \to M$, with $\phi : \partial D \to C$, is a map from a disk D to M sending the boundary to a submanifold C, and A is a $U(1)$ gauge connection on the boundary. We consider a supersymmetric version of this action. In the presence of a boundary condition, however, some of the worldsheet supersymmetry, if not all, will be broken. Likewise, the boundary may destroy conformality, even if the theory with D a closed surface would be conformal. If we start with a superficial $\mathcal{N} = (2,2)$ superconformal action and ask what boundary conditions preserve conformality, and an $\mathcal{N} = 2$ supersymmetric left-right combination of $(2,2)$ (which is the maximal allowed), the resulting boundary vacuum can be thought of as a nice D-brane. This is one way of deriving the D-brane conditions we will state in this chapter. In Ch. 39 this is discussed in great detail.

Another method is to take a space-time viewpoint. One writes down the D-brane worldvolume action and asks which space-time supersymmetry generators leave this action invariant. This is the approach we adopt in this chapter. Note, however, that the connection to geometry that we shall discuss will only be valid at "large radius." (Techniques of conformal field theory are valid at all radii.) Let us now describe the conditions on the submanifold and the gauge field describing the D-branes.

Boundary conditions are set by specifying a submanifold C of the Calabi–Yau, on which the fields must take values at the boundary of the Riemann

[5]Mirror symmetry does not affect the string coupling constant, so a state that is non-perturbative in one theory must be non-perturbative in the dual theory.

surface. In addition, we include a term in the action equal to the holonomy of a $U(1)$ gauge field along the loop. (Gauge invariance will demand a mixing of the field strength with the background B-field.) We call the boundary data a "supersymmetric cycle" if supersymmetry and conformal symmetry are preserved in the boundary field theory. There are two kinds of supersymmetric cycles (C, L) on a Calabi–Yau threefold M, where C is a (possibly singular and with multiplicity) submanifold of M and L is a complex line bundle over C together with a $U(1)$ connection D_A. Let us denote the Kähler form (resp. holomorphic volume form) on the Calabi–Yau threefold by ω (resp. Ω).

The type-A supersymmetric cycle is when C is a special Lagrangian submanifold of M with a flat $U(1)$ connection. That is, if $i : C \hookrightarrow M$, then

$$i^*\omega = 0 \;\; (\text{"Lagrangian"}) \text{ and } \qquad i^*\left[\operatorname{Im} e^{-i\theta}\Omega\right] = 0 \;\; (\text{"special"}), \tag{37.1}$$

$$F_A = 0, \tag{37.2}$$

where F_A is the curvature of D_A. We can write the special condition as

$$\operatorname{Im}\Omega = \tan\theta \operatorname{Re}\Omega. \tag{37.3}$$

Here and above θ is a constant. In the presence of a background B-field (an element of $H^2(M, \mathbb{R}/\mathbb{Z})$), F_A should be replaced by $F_A - B$, where B is understood to be pulled back to the submanifold.

The type-B cycle is when C is a complex submanifold of M of dimension n and the curvature two-form F_A of D_A satisfies the conditions

$$F_A^{0,2} = 0, \tag{37.4}$$

$$\operatorname{Im} e^{-i\theta}(\omega + F_A)^n = 0. \tag{37.5}$$

The first equation says that the (0,1) component of the connection determines a holomorphic structure on L. The second equation is called the MMMS (or deformed Hermitian-Yang–Mills) equation and it is equivalent to the equation

$$\operatorname{Im}(\omega + F_A)^n = \tan\theta \operatorname{Re}(\omega + F_A)^n. \tag{37.6}$$

For example, when C is the whole Calabi–Yau manifold M of dimension three, when we expand the second equation we get $F \wedge \omega^2/2 - F^3/6 = \tan\theta\left[\omega^3/6 - \left(F^2/2\right)\wedge\omega\right]$. The angle θ is called the "phase" of the supersymmetric cycle.

These equations involve semi-classical reasoning. Their validity is taken to hold at or near large radius.

Mirror symmetry at the non-perturbative level states that all D-branes (and all questions involving D-branes) correspond to D-branes on the mirror manifold. We now explore the consequences of this assertion.

37.4. Mirror Symmetry as Generalized T-Duality

The title of this section is no surprise, given the physical proof of mirror symmetry by the gauged linear sigma model which T-dualizes the phase of complex fields. However, another look at mirror symmetry from the point of view of D-branes leads to new mathematical conjectures about how mirror symmetry works in cases where the mirror theories are both realized as geometric Calabi–Yau manifolds. This generalizes the discussion we had in the context of D-branes on T^2 in Ch. 19.

Some very simple observations reveal intriguing predictions about mirror symmetry. Consider a mirror pair M and $\widetilde{M}$.[6] If we look at B-cycles on M there are two immediate distinguished candidates: the zero-cycle and the six-cycle. As we will assume M is simply connected, we can forget about the data involving flat line bundles (there is only the trivial one). The six-cycle is M itself and has no moduli as a holomorphic submanifold: its moduli space $\mathcal{M}_D$ as a D-brane is a single point. The zero-cycle is a point in M and the moduli space of choices $p \in M$ is all of $M : \mathcal{M}_D = M$.

On, $\widetilde{M}$, there must be corresponding objects S (to the six -cycle) and T (to the zero-cycle). They must have the same moduli spaces as their partners. (The reason is that we can build this correspondence at large radius, where classical geometry is a good approximation to the actual quantum moduli space.) The situation is summarized as follows.

Let us focus on $T \subset \widetilde{M}$ in Table 1. We are led to the Strominger-Yau-Zaslow conjecture: *$\widetilde{M}$ has a distinguished submanifold whose D-brane moduli space produces the mirror manifold, M.* This gives an intrinsic characterization of the mirror manifold.

We can learn more. The moduli space $\mathcal{M}_D(T) = M$ has complex dimension 3. As we shall see, this will imply that T has first Betti number equal to 3. We learn that T has the Betti numbers of a three-torus, and we will assume T is a torus. (Similarly, we find $b_1(S) = 0$.) Further, $\mathcal{M}_D(T)$ naturally

[6] We assume that M and $\widetilde{M}$ are both Calabi–Yau manifolds here.

M D-brane	Moduli Space, $\mathcal{M}_D$	$\widetilde{M}$ D-brane
0	M	T
6	*	S

TABLE 1. D-branes

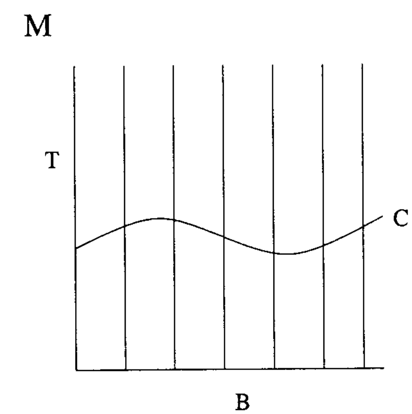


FIGURE 2. An illustration of M as a fibration over a base, B by special Lagrangian tori (T). A special Lagrangian submanifold C which is a section of this fibration is also shown.

fibers over $\mathcal{M}_{\mathrm{sLag}}(T)$, the moduli space of T as a special Lagrangian submanifold. The map $\mathcal{M}_D(C) \to \mathcal{M}_{\mathrm{sLag}}(C)$ for any cycle, C, is obtained by simply forgetting all bundle data. The fiber T' of this map is the data of all flat bundles on C. The space T' consists of a torus of dimension $b_1(C)$, i.e., a circle of possible $U(1)$ holonomies for each loop in $C : T' = \mathrm{Hom}(\pi_1 C, S^1)$. Therefore, we learn that M *admits a fibration with torus fibers:*

$$\begin{array}{c} T' \to \mathcal{M}_D(T) = M \\ \downarrow \pi \\ \mathcal{M}_{\mathrm{sLag}}(T). \end{array}$$

(In the diagram, T' represents a generic fiber $\pi^{-1}(p)$.)

In fact, more can be shown through an analysis of the metric on D-brane moduli space. Let us review the reasoning that leads to a sharper conjecture. From the space-time point of view, a D-brane represents an extended object — a location of energy density. The low-lying states of the boundary conformal field theory correspond to motions or fluctuations of this object about its minimal configuration. Thus, massless states of the boundary conformal field theory correspond to moduli of the classical space of minimal configurations — tangent vectors of $\mathcal{M}_D$. If we allow the D-brane to move in space-time, it sweeps out a $(3+1)$-dimensional worldvolume, which can be described by an action $S = \int d^3x dt\, \mathcal{L}$, where $\mathcal{L}$ is the Lagrangian density. The bosonic part of $\mathcal{L}$ is $\sqrt{-g}$, where g is the (determinant of the) metric on $C \times \mathbb{R}$ induced from the map $C \times \mathbb{R} \to M \times \mathbb{R}$. Note that for static configurations, S is minimized by area-minimizing maps. If we include the gauge field, g must be replaced by $g - \mathcal{F}$, where $\mathcal{F} = F - B$ is the (B-field-corrected) field strength. This is known as the Dirac–Born–Infeld action.

We consider then a three-dimensional family of supersymmetric three-tori in $\widetilde{M}$. If we are near a situation where the three-tori are flat, then neighboring three-tori do not intersect (this will be clear from Sec. 37.4.1) and provide a local fibering of the Calabi–Yau. Therefore, we can use coordinates $y^a \sim y^a + 1$ and x^m, $a, m = 1, \ldots, 3$ on $\widetilde{M}$ and write the metric $\widetilde{g}$ in these coordinates:

$$\widetilde{g} = \widetilde{g}_{ab} dy^a dy^b + \widetilde{g}_{am} dy^a dx^m + \widetilde{g}_{ma} dx^m dy^a + \widetilde{g}_{mn} dx^m dx^n.$$

The y^a are torus coordinates, and the x^a are the three coordinates on $\mathcal{M}_{\rm sLag}(T)$. Now $\mathcal{M}_D(T)$ is parametrized by x^m and u^a, where u^a describe the flat connection $\sum_a u^a dy^a$ on $T = \pi^{-1}(x)$. One then makes a low-energy approximation by assuming that the time-dependence of the D-brane configuration is given by a path in $\mathcal{M}_{\rm sLag}(T)$ with time derivative $\dot{x}^a$, and $\dot{y}^a$ are determined from $\dot{x}^a$ by requiring the motion of the D-brane to be induced by normal (perpendicular) vectors: i.e.,

$$\dot{y}^a = \widetilde{g}^{ab} \widetilde{g}_{bm} \dot{x}^m.$$

Here $\widetilde{g}^{ab}$ is the inverse metric on the torus (not on $\widetilde{M}$): $\sum_{b=1}^{3} \widetilde{g}^{ab} \widetilde{g}_{bc} = \delta^a{}_c$. Finally, we allow a gauge field time-dependence $\dot{u}^a$ as well.

Expanding the action S to second order in time derivatives yields a metric g_D on D-brane moduli space $\mathcal{M}_D(T)$, which we will soon equate with

M. This metric has no u-dependence. Though such isometries (u translations) are impossible for a simply-connected Calabi–Yau, we note that the true D-brane moduli space metric will be corrected by instantons arising from holomorphic maps from disks with boundary lying along the D-brane, and this introduces u-dependence. Ignoring such corrections, we have u-independence; the u parametrize flat tori and we can perform the usual $R \leftrightarrow 1/R$ duality ("T-duality") on each circle. If we T-dualize g_D in this way, we get a new metric and find that *we recover* $\widetilde{g}$*!* The mirror metric is therefore related by T-duality, and the mirror manifold is the mirror fibration. This argument will be physical in the large radius limit of $\widetilde{M}$ away from singular fibers, where one obtains $M = \mathcal{M}_D$ at the large complex structure limit (where the metric has $U(1)$ isometries away from singular fibers). Near the large complex structure limit, the instanton sum should converge. We can thus identify the mirror manifold with the manifold obtained by dualizing the torus fibration. Further, one finds that the torus fibers of $M = \mathcal{M}_D$ are themselves special Lagrangian. Reversing the reasoning with this torus fibration, we arrive at the conjecture that *every Calabi–Yau manifold with a mirror admits a fibration by special Lagrangian tori near its large complex structure limit, and the mirror manifold is obtained by dualizing this fibration.* The mirror is thus the moduli space of these toroidal D-branes.

EXAMPLE 37.4.1. *The simplest example is the torus $(S^1)^n \times (S^1)^n$, which is an $(S^1)^n$ fibration over $(S^1)^n$. The dual fibration is also a torus and the two are related by ordinary T-duality. For simply-connected compact Calabi–Yau manifolds, the fibrations are singular over loci in the bases, and the construction of the dual fibrations is rather subtle. One can also consider orbifolds of these theories to obtain mirror pairs of Calabi–Yau with $SU(N)$ holonomy. Namely, suppose T^n has a discrete isometry group G, as a subgroup of $SO(n)$. We can orbifold this theory by simultaneous G action on the base and the fiber. The choice of the action on the base is the same for mirror pairs, but the choice of the action on the fiber may not be the same. A simple example is $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ acting on T^3, for which the choice of the fiber action is the same for the mirror pairs (up to a subtlety having to do with "discrete torsion").*

37.4.1. Comments on Special Lagrangian Moduli Space. Here we collect, without proof, some results on the geometric structure of special Lagrangian moduli space.

Let $L \hookrightarrow M$ be a smooth special Lagrangian submanifold, with $i : L \to M$ the inclusion map. So $i^*\omega = 0$ and $i^*\mathrm{Im}\,\Omega = 0$, for some choice of phase of Ω. Now consider a family $f(t)$ of immersions of L in M such that $f(0) = i$. Then, taking derivatives at zero, it is easy to see that

$$\frac{d}{dt} f^*\omega|_{t=0} = d\theta_{\dot{f}},$$

where $\theta_{\dot{f}}$ is a one-form on L defined by the normal vector field $\dot{f} = f_*\frac{d}{dt}$ representing the first-order deformation. Explicitly, $\theta_{\dot{f}}(V) = \omega(\dot{f}, f(0)_*V)$. By preserving the Lagrangian condition to first order, we learn that the tangent space (at L) to the moduli space of deformations of L as a Lagrangian submanifold is equal to the closed one-forms on L. Similarly, it can be shown that preserving the special condition to first order requires $d^\dagger\theta_{\dot{f}} = 0$, where the adjoint is taken with respect to the induced metric on L. McLean also showed that these first-order deformations are unobstructed, i.e., the moduli space of deformations of L as a special Lagrangian submanifold, $\mathcal{M}_{\mathrm{sL}}$, is smooth at L. We learn

$$T\mathcal{M}_{\mathrm{sL}}|_L = \mathcal{H}^1(L),$$

where $\mathcal{H}^1(L)$ denotes the harmonic one-forms.

In the case of D-branes, we will want to add the data of a $U(1)$ bundle, up to gauge transformation. The choices, for a fixed curvature (solving the D-brane equation), are equal to $H^1(L;\mathbb{R})/H^1(L;\mathbb{Z})$, Therefore, the moduli space of D-branes $\mathcal{M}_D$ fibers over $\mathcal{M}_{sL}$ with toroidal fibers. Note that, using the metric on L, we can equate $H^1(L;\mathbb{R})$ with $\mathcal{H}^1(L)$.

In fact, the cohomology space $H^1(L)$ serves as a local model for the moduli space $\mathcal{M}_{sL}$. To see this, consider the set $U \subset \mathcal{M}_{sL} \times M$ consisting of pairs (L', p), where $p \in L'$ (a kind of universal deformation). Note that U maps to M, so we can consider ω pulled back to U (which we will also denote ω). Now consider a coordinate y for $\mathcal{M}_{sL}$, and a vector $v = \frac{\partial}{\partial y}$. Choose any lift of v (also denoted v) in TU. Then $i_v\omega$ (interior multiplication) defines a one-form on U, which over the point in $\mathcal{M}_{sL}$ represented by L is a one-form on L. In the notation above, $i_v\omega$ is just θ_v.

Now let $i_v\omega$ be the element of $\mathcal{H}^1(L,\mathbb{R})$ corresponding to v. Then if $\gamma_1, \dots, \gamma_s$ is an integral basis for $H_1(L,\mathbb{Z})$ $(s \equiv b_1(L))$, we can define the

one-forms α_i on $\mathcal{M}_{sL}$ at L by

$$\alpha_i(v) = \int_{\gamma_i} i_v\omega.$$

By letting the γ_i vary over $L' \in \mathcal{M}_{sL}$, we get one-forms $\alpha_1, \dots, \alpha_s$ on $\mathcal{M}_{sL}$, defined up to monodromies (of H_1 around loops in $\mathcal{M}_{sL}$).

The same procedure can be done using $\mathrm{Im}\,\Omega$ instead of ω. We get an $(n-1)$-form $i_v\mathrm{Im}\,\Omega$, and choosing a basis $\Gamma_1, \dots, \Gamma_s$ for $H_{n-1}(L,\mathbb{Z})$ (here $n = \dim_{\mathbb{C}} M$), we can define one-forms β_i on $\mathcal{M}_{sL}$ by $\beta_i(v) = \int_{\Gamma_i} i_v\mathrm{Im}\,\Omega$. A consequence of the special Lagrangian condition is that

$$*i_v\omega = -i_v\mathrm{Im}\,\Omega, \tag{37.7}$$

where $*$ is taken with respect to the induced metric on L.

It can be shown that the α_i define a *frame of closed one-forms* on $\mathcal{M}_{sL}$, and the same is true of the β_i. This means that we can define coordinates y_i such that $\alpha_i = dy_i$, and coordinates $\check{y}_i$ such that $\beta_i = d\check{y}_i$. The only choices we made were an integral basis for H_1 (or H_{n-1}), and integration constants. Different choices would be related by $y_i' = \sum_j a_{ij}y_j + b_i$, with $(a_{ij}) \in GL(n,\mathbb{Z})$ and $b_i \in \mathbb{R}$ (similarly for $\check{y}_i$), i.e., by an affine transformation. This means that we have found an *affine structure* on $\mathcal{M}_{sL}$ — two such structures, actually — meaning that this manifold can be coordinatized such that the transition functions are affine transformations.

There is a natural metric g on $\mathcal{M}_{sL}$ defined by

$$g(v,w)|_L = -\int_L i_v\omega \wedge i_w\mathrm{Im}\,\Omega$$

(using Eq. (37.7) this becomes the natural Weil–Petersson metric $\int_L (i_v\omega, i_w\omega)dV_L$). This metric has a potential, i.e., in coordinates $y_1, \dots, y_s$ we can write $g_{ij} = \frac{\partial^2 K}{\partial y_i \partial y_j}$. Similarly, for the coordinates $\check{y}_j$, we can find a potential $\check{K}$. In fact, $\check{y}_i = \frac{\partial K}{\partial y_i}$ (and *vice versa*). The two are related by Legendre transform: $\check{K} = \sum_{i=1}^s \check{y}_i y_i - K$.

These properties will be used (implicitly, mostly) in Sec. 37.9. Finally, we note that since the moduli space of D-branes fibers over $\mathcal{M}_{sL}$ with fiber $H^1(L,\mathbb{R})/H^1(L,\mathbb{Z})$ over a point L, we can form $\mathcal{M}_D$ by quotienting the total space of $T^*\mathcal{M}_{sL}$ or $T\mathcal{M}_{sL}$ by lattices defined, say, by the α_i or β_i respectively. One has a natural symplectic structure, one a natural complex structure. Mirror symmetry between two manifolds is expected to interchange these structures. In particular dual Calabi–Yau manifolds have been described as compactifications of dual torus fibrations arising in this way.

37.5. Mirror Symmetry with Bundles

One lesson from physics is that new symmetries can appear at points where states become massless. In a quantum field theory with a charged complex scalar field ϕ, a potential $(|\phi|^2 - a^2)^2$ demands a vacuum value ϕ_0 with $|\phi_0|^2 = a^2$. This breaks the gauge symmetry of the vacuum to the little group preserving ϕ_0. New fields ϕ' are determined by expanding around ϕ_0: $\phi' \equiv \phi - \phi_0$. The multiplet ϕ now decomposes: massless fields correspond to directions from ϕ_0 along the space $|\phi|^2 = a^2$, while massive fields correspond to directions normal to this space. Some of the formerly massless gauge fields become massive, due to the $A\phi_0 A\phi_0$ term in the expansion of the kinetic energy $|D_A(\phi' + \phi_0)|^2$, where D_A is the covariant derivative. This symmetry breaking is called "the Higgs mechanism." When $a = 0$ the full gauge symmetry is restored.

Exactly the same phenomenon occurs when D-branes coincide. Open strings stretched between D-branes are massive, but become massless when the distance between D-branes becomes zero, so they sit on top of each other. Therefore, if D-branes A and B coincide, then the two $U(1)$ gauge fields can combine with the A-B and B-A open string sectors to form massless fields in the fundamental representation of $U(2)$. In fact, the Lie algebra decomposes into $u(1)$ and $su(2)$, which we can treat separately. We get enhanced gauge symmetry upon coincidence.

Consider, then, N D6-branes wrapping the whole Calabi–Yau manifold, M. We get an $SU(N)$ gauge bundle. If the bundle is trivial, it must correspond on the mirror $\widetilde{M}$ to a submanifold in the homology class N times that of the torus, T. A non-trivial bundle E can be thought of as including lower-dimensional branes on top of ("bound to") the six-brane. In general, then, we have a map from the Chern character class $\mathrm{ch}(E)$ of the bundle — or the Poincaré dual cycles in $H_{\mathrm{even}}(M)$ — to the homology class of the corresponding Lagrangian submanifold, which lies in $H_{\mathrm{odd}}(\widetilde{M})$.[7]

[7]Finding the full map of integral homologies may not be so easy in general.

What can be learned from such an identification? Before answering this, let us first recall that classical mirror symmetry involves equating three-point functions of states A_i of the topological theory — or massless states ("marginal operators") in the superconformal field theory. In the B-model (on $\widetilde{M}$, say), $A_i \in H^1(T\widetilde{M})$ and

$$\langle A_i A_j A_k \rangle = \int (A_i \wedge A_j \wedge A_k)^\sharp \wedge \Omega = \partial_i \partial_j \partial_k \mathcal{F},$$

where the $\sharp$ indicates that we equate $\Lambda^3 T$ with the trivial bundle (using Ω), so $A_i \wedge A_j \wedge A_k$ is in $H^{0,3}(\widetilde{M})$. Let us now return to the present D-brane setting, and look on the B-cycle (in M) side.[8] The structure parallels the closed string case, only now we have a boundary superconformal field theory preserving an $\mathcal{N} = 2$ (a combination of left and right) and we can form the topological theory whose states live in $H^1(End(E))$, infinitesimal deformations of the holomorphic structure of E. The three-point function is described by the cubic term in a holomorphic Chern–Simons theory:

$$\langle A_i A_j A_k \rangle = \int \mathrm{Tr}\,(A_i \wedge A_j \wedge A_k) \wedge \Omega.$$

We now turn to the mirror models, first the closed then open cases. In the closed string mirror A-model (on M, say) the A_i are in $H^{1,1}(M)$ and the calculation $\langle A_i A_j A_k \rangle$ involves an instanton sum over holomorphic curves meeting cycles that are Poincaré dual to the A_i. In the D-brane open string calculation on the A-brane $C \subset \widetilde{M}$, the A_i are in $H^1(C, \mathbb{C})$ and $\langle A_i A_j A_k \rangle$ is a sum over holomorphic maps ϕ from a disk D with boundary ∂D mapping to C, such that three points on the boundary meet cycles Poincaré dual to the A_i. Again, these instantons are weighted by $\exp \int_D \phi^* \omega$, where ω is the complexified Kähler form, and also by the holonomy $\exp \oint_{\partial D} \phi^* \beta$, where β is the $U(1)$ connection. We also must sum over different cycles in C which are wrapped by the boundary ∂D.

To write down the three-point function for A-cycles, let us set some notation. Let γ_a be a basis for $H_1(C)$, $a = 1, \ldots, b_1(C)$. Let D_a be a minimal-area disk in $\widetilde{M}$ and $r_a = \int_{D_a} \omega$ its area. Let $u_a = \oint_{\gamma_a} \beta$ be the

[8]The choice of A/B cycles in what follows may be confusing. Note that in type IIA string theory, branes are even-dimensional, with odd-dimensional world volumes, and can wrap all of spacetime as well as an odd-dimensional submanifold. In type IIB, branes can wrap spacetime plus an even dimensional submanifold.

holonomy. We combine these two numbers in a complex coordinate

$$w_a = u_a + ir_a; \qquad \widetilde{q}_a = \exp(2\pi i w_a).$$

Now since two disks with the same boundary differ by a closed two-cycle, any disk with boundary γ_a has area $r_a + \sum_i d_i t_i$, with $i = 1, \ldots, b_2(\widetilde{M})$, with $t_i = \int_{C_i} \omega$ (the t_i will be complexified in the usual way to the complex coordinate $q_i = \exp 2\pi i t_i$; the B-field contribution to r_a is absorbed by the u_a) where C_i are a basis for $H_2(\widetilde{M})$. Now consider $\langle A_a A_b A_c \rangle$. Let $U_a \subset C$ be a two-cycle Poincaré dual to the one-form A_a, etc. Let

$$K_{\vec{d},\vec{m}}(a,b,c)$$

be the "number" of holomorphic maps ϕ from a disk D into $\widetilde{M}$ such that

- $\vec{m}$ describes the image of the boundary:
$$\phi(\partial D) = \sum_{a=1}^{b_1(C)} m_a \gamma_a;$$
- $\vec{d}$ describes the image disk, i.e., the class of $D - \sum_a m_a D_a$ (a closed two-cycle) is given by $\sum_{i=1}^{b_2(\widetilde{M})} d_i C_i$; and
- three cyclically-ordered points $(0, 1, \infty)$ on the boundary of the disk are mapped into U_a, U_b, U_c respectively.

By "number" here we mean the top Chern class of some obstruction bundle, as is familiar to readers from Theorem 26.1.2 and Sec. 26.2. Note that the theory of Gromov–Witten invariants for Riemann surfaces with boundary is still a rather nascent area of study. Algebraic geometry may not be as strong a calculational tool as it is for the usual Gromov–Witten invariants.

Now assuming we can map from bundles to special Lagrangian cycles, so that we have a correspondence $E \leftrightarrow (C, A)$ (A is the flat connection), and assuming we can map infinitesimal deformations on both sides (we denote both by A_i), then we have the proposed equality

$$\int_M \mathrm{Tr}\,(A_i \wedge A_j \wedge A_k) \wedge \Omega = \sum_{\vec{d},\vec{m}} K_{\vec{d},\vec{m}}(i,j,k) \cdot m_i m_j m_k \prod_{a=1}^{b_1(C)} \widetilde{q}_a^{\,m_a} \prod_{m=1}^{b_2(\widetilde{M})} q_m^{d_m}.$$

In fact, similarly to the closed Riemann surface case, this equation can be written as the third derivative of a generating function of open string Gromov–Witten invariants. Such formulas have been checked mathematically in several non-trivial examples, where both sides of the equation could be computed (and a mirror map could be found). Counting disk instantons is a problem similar to the closed string instantons, though rigorous proofs have not yet been established.

This is just the first piece of the prediction of mirror symmetry applied to D-branes. One can also consider other topologies for the Riemann surface with arbitrary holes and handles, just as one does in the context of closed topological strings. For example for the annulus the mirror map leads to a reformulation of counting of holomorphic maps from annuli ending on A-branes, in terms of Ray–Singer torsion on the mirror B-brane.

37.6. Mathematical Characterization of D-branes

37.6.1. A-cycles. We have discussed A-cycles as being special Lagrangian submanifolds with flat bundles on them. This is the viewpoint we will continue to use. However, we will mention that Hitchin has offered a description of D-branes as "gerbes." For Calabi–Yau threefolds, gerbes can be defined by codimension 3 real submanifolds, analogously to the relation between divisors and line bundles. Gerbes contain the line bundle information on the special Lagrangian as well. Though this characterization of D-branes may ultimately be useful, we will stick to our more pedestrian point of view here.

37.6.2. B-cycles. The story for B-cycles is more interesting. How should we think of a holomorphic submanifold with a bundle over it? If the submanifold is the full space, we have seen that we can think of a B-cycle as a vector bundle. However, if we think of a holomorphic sheaf, its support must be a holomorphic submanifold (or union of them). Perhaps sheaves are the proper language for speaking about B-cycles. Coherent sheaves are sheaves of sections that are locally quotients of a finitely-generated free group of sections by a finitely-generated group of relations.

EXAMPLE 37.6.1. *For example, the structure sheaf $\mathcal{O}_p$ of a point $p = (a, b) \in \mathbb{C}^2$ can be written in a sequence $\mathcal{O} \oplus \mathcal{O} \to \mathcal{O} \to \mathcal{O}_p \to 0$, where the first map sends (f, g) to $(x - a)f + (y - b)g$. Note that this map has a kernel, but the local presentation in terms of $\mathcal{O}$'s is all that is needed to show coherence.*

It seems reasonable to assume that D-branes can be described locally by a set of equations, giving the associated sheaves the coherent property.

(But who knows?) In any case, the category of coherent sheaves seems large enough to house most D-branes, and is probably enough to gain some intuition. However, some puzzles arise. One is that for K3, there is a symmetry between sheaves called a Fourier–Mukai transform, which is believed to be a symmetry of the string theory. However, this symmetry acts on a larger class of objects than sheaves. It acts on cohomological *complexes of sheaves*, up to an equivalence we now describe. Two complexes $\mathcal{E}^\bullet$ and $\mathcal{F}^\bullet$ are deemed equivalent if there is a map between them (in either direction) inducing the same cohomology sheaves, i.e., a "quasi-isomorphism" (also if there are two different quasi-isomorphisms from a third complex of sheaves, or even a chain of such linkages). This equivalence has been constructed by "inverting" quasi-isomorphisms, and the category of bounded complexes up to this equivalence is called the *derived category of coherent sheaves*. It will be discussed in greater detail in Sec. 38.3.

EXAMPLE 37.6.2. *Completing the sequence from Example 37.6.1, we have*

$$0 \to \mathcal{O} \to \mathcal{O} \oplus \mathcal{O} \to \mathcal{O} \to \mathcal{O}_p \to 0$$

(the first map sends f to $-(y-b)f \oplus (x-a)f$), from which we can see that $E^\bullet \equiv \mathcal{O} \to \mathcal{O} \oplus \mathcal{O} \to \mathcal{O}$ is quasi-isomorphic to $\mathcal{O}_p$, hence isomorphic in the derived category.

As of this writing, perhaps the closest understanding we have of the physics of quasi-isomorphism is as follows. As mentioned, there are strings connecting pairs of D-branes that become massless when the D-branes coincide. What about a D-brane/anti-D-brane system? One might expect that such a configuration is unstable, with the stable configuration described by the vacuum, after mutual annihilation. An indication of an instability would be the presence of a tachyonic state in the string theory. Indeed, "tachyonic" means negative mass-squared, and the mass-squared is the second derivative of the potential — so tachyons correspond to unstable critical points of a field potential. It turns out that such tachyons exist, and again correspond to strings stretching between the brane and anti-brane.[9] For the D-brane/anti-D-brane pair, the gauge fields on the branes and these tachyons comprise

[9]Sen's tachyon condensation conjecture is that the tachyonic field rolls to a true (non-tachyonic) minimum of the field potential, at which point the D-brane and anti-D-brane have annihilated each other. This conjecture has now been formulated as a universal problem in open string field theory, and is well-evidenced in numerical computations.

the data of a "superconnection"[10]

$$\begin{pmatrix} A & T \\ \overline{T} & B \end{pmatrix},$$

where A and B are connections on the branes (bundles) E and E' and the tachyons have bi-representation indices (corresponding to the two ends of the open string) making T a section of $\mathrm{Hom}\,(E, E')$ and $\overline{T}$ a section of $\mathrm{Hom}\,(E', E)$, respectively. Thus we have a two-term complex of bundles with connections $E \to E'$, with the tachyon providing the map, as well as a reverse map. This is not quite the derived category, as we can't interpret integer shifts of grading (though perhaps it can serve as a "rolled-up" version of one), but perhaps offers a clue as to what quasi-isomorphism means. That is, the unstable tachyonic configuration may be equivalent to a complex $0 \to E \to E \to 0$, which, if the map in the middle is an isomorphism, is equivalent (quasi-isomorphic) to the zero object (vacuum), indicating annihilation. For recent developments in the understanding of the relationship between tachyons and the derived category, see work of Douglas et al [**75, 76**].

37.7. Kontsevich's Conjecture

Observations involving topological open string models and the need to enlarge our understanding of mirror symmetry to include D-branes (and not just the conformal ones) leads to the question of what is the proper mathematical formulation of the mirror symmetry in the context of D-branes. Kontsevich has proposed a remarkable definition of what mirror symmetry should be in the context of D-branes: an equivalence of two categories. On one side we have the derived category of coherent sheaves, and on the other side we have special Lagrangian submanifolds with flat (or otherwise fixed-curvature) $U(1)$ connections. As we mentioned previously, *all* questions involving D-branes and their *interactions* must be equivalent. This amounts to an equivalence betwen the *categories* of D-branes.

Let us recall that categories in mathematics are comprised of a class of objects; for every pair of objects $\mathcal{E}_1$, $\mathcal{E}_2$, a set of morphisms, $\mathrm{Hom}\,(\mathcal{E}_1, \mathcal{E}_2)$; and

[10]In fact it seems that the natural notion in this context may be a connection in the supergroup $U(N|M)$, and the topological B-model becomes equivalent to holomorphic Chern-Simons for this supergroup.

for pairs of morphisms $\phi \in \mathrm{Hom}(\mathcal{E}_1, \mathcal{E}_2)$ and $\psi \in \mathrm{Hom}(\mathcal{E}_2, \mathcal{E}_3)$ a composition $\psi \circ \phi \in \mathrm{Hom}(\mathcal{E}_1, \mathcal{E}_3)$.

EXAMPLE 37.7.1. *Here are some examples of categories. Objects: sets, topological spaces, sheaves, groups, vector spaces, complexes. Respective morphisms: functions, continuous functions, maps of sheaves, group homomorphisms, linear maps, morphisms of complexes. Compositions: compositions are compositions of maps in all examples. Note that, in general, the morphisms are not necessarily maps and merely have the structure of a set. In such cases, the composition needs to be defined carefully. We will encounter such an example in the case of A-cycles.*

Kontsevich's conjecture (really a definition) states that M and $\widetilde{M}$ are mirror pairs if and only if

$$\mathcal{D}^b(M) \cong \mathcal{F}^0(\widetilde{M}).$$

On the left is the bounded derived category of coherent sheaves (B-cycles), and on the right is essentially the category of A-cycles, whose compositions depend on holomorphic maps from disks. Mathematically, it is a category derived from Fukaya's A^∞ category. We will discuss the derived category $\mathcal{D}^b(M)$ in more detail in Ch. 38. Below we will discuss the Fukaya category in more detail.

From the physical point of view, interpreting the objects on the left-hand side as the B-branes and the right-hand side as the object for A-branes, and noting that the operations on the objects in each category relate to topological amplitudes at the level of the disc (as will be discussed below for the Fukaya category) makes one direction of the above conjecture obvious. The reverse direction is also plausible physically because one expects from the target space point of view that the properties of BPS D-branes characterize the full sigma model on Calabi–Yau. Note however that, from the physics proof of mirror symmetry given in Ch. 20, we do not expect the mirror symmetry to involve just a pair of manifolds. In particular, more generally we obtain an LG theory mirror to a sigma model on a manifold. The above conjecture should be suitably generalized to also include this more general case.

	physics	math	math
	$M/\widetilde{M}$	M	$\widetilde{M}$
Obj	B/A-cycle D-branes	cpxes of sheaves $\mathcal{E}^\bullet$	sLags + bundles (C, E)
Hom	massless flds	$\mathrm{Hom}(\mathcal{E}_1^\bullet, \mathcal{E}_2^\bullet)$	$\bigoplus_{p \in C_1 \cap C_2} \mathrm{Hom}(E_1, E_2)\vert_p$
Comp	3-pt functs	comp of morphisms	comp by hol maps

TABLE 2. D-brane categories

37.7.1. Fukaya's A^∞ Category. Let us describe Fukaya's A^∞ category.[11] Fukaya's definition involves not special Lagrangian, but merely Lagrangian submanifolds, up to something called Hamiltonian deformation. This gives something very similar to the category of sLags, as we shall see in Ch. 38. The idea is to consider the open topological string theory on the disk, with several branes.

An A^∞ category is one with multi-compositions m_k (not just pairwise) involving k morphisms, $k = 1, \ldots, \infty$, which obey various compatibility equations. A^∞ categories are to ordinary categories with associative compositions as A^∞ algebras are to associative algebras. The compositions are "associative up to homotopy," a concept we now describe.

Objects. The objects of $\mathcal{F}(\widetilde{M})$ are special Lagrangian submanifolds of $\widetilde{M}$ — i.e., minimal Lagrangian submanifolds — endowed with flat bundles E with monodromies having eigenvalues of unit modulus[12], and one additional structure we will discuss momentarily. Thus, an object $\mathcal{U}$ is a pair:

$$\mathcal{U} = (L, E).$$

[11] Readers may wish to read the next section in parallel with this one, where a detailed example of Fukaya's category and the Kontsevich equivalence is worked out for the case of the elliptic curve.

[12] Kontsevich considered only unitary local systems, or flat $U(n)$ bundles. The Jordan blocks will be related to non-stable vector bundles over the torus.

The additional structure we need is the following. A Lagrangian submanifold L of real dimension n in a complex n-fold, $\widetilde{M}$, defines not only a map from L to $\widetilde{M}$ but also the Gauss map from L to V, where V fibers over $\widetilde{M}$ with fiber at x equal to the space of Lagrangian planes at $T_x\widetilde{M}$. The space of Lagrangian planes has fundamental group equal to $\mathbb{Z}$, and we take as objects special Lagrangian submanifolds together with lifts of the Gauss map into the fiber bundle over $\widetilde{M}$ with fiber equal to the universal cover of the space of Lagrangian planes.

Hom's. The morphisms $\mathrm{Hom}(\mathcal{U}_i, \mathcal{U}_j)$ are defined as

$$\mathrm{Hom}(\mathcal{U}_i, \mathcal{U}_j) = \bigoplus_{p \in L_i \cap L_j} \mathrm{Hom}(E_i|_p, E_j|_p),$$

where the second "Hom" in the above represents homomorphisms of vector spaces underlying the local systems at the points of intersection. There is a $\mathbb{Z}$-grading on the Homs. If p is a point in $L_i \cap L_j$, then it has a Maslov index $\mu(p) \in \mathbb{Z}$.

Generally speaking, a category has composable morphisms that satisfy associativity conditions. This is not generally true for the category $\mathcal{F}(\widetilde{M})$. However, we have instead on $\mathcal{F}(\widetilde{M})$ an additional interesting structure making $\mathcal{F}(\widetilde{M})$ an A^∞ category. Associativity will hold cohomologically, in a way. The equivalence of categories that we will prove in Sec. 37.8 will involve a true category $\mathcal{F}^0(\widetilde{M})$, which we will construct from $\mathcal{F}(\widetilde{M})$ after describing the A^∞ structure.

37.7.2. A^∞ Structure. The category $\mathcal{F}(\widetilde{M})$ has an A^∞ structure, by which we mean the composable morphisms satisfy conditions analogous to those of an A^∞ algebra. An A^∞ algebra is a generalization of a differential, graded algebra. Namely, it is a $\mathbb{Z}$-graded vector space, with a degree 1 map, m_1, which squares to zero ($(m_1)^2 = 0$). There are higher maps, $m_k : A^{\otimes k} \to A$, as well.

DEFINITION 37.7.2. *An A^∞ category, $\mathcal{F}$ consists of*

- *A class of objects* $\mathrm{Ob}(\mathcal{F})$;
- *For any two objects, X, Y, a $\mathbb{Z}$-graded abelian group of morphisms* $\mathrm{Hom}(X, Y)$;

- *Composition maps*

$$m_k : \mathrm{Hom}(X_1, X_2) \otimes \mathrm{Hom}(X_2, X_3) \otimes \cdots \otimes \mathrm{Hom}(X_k, X_{k+1}) \to \mathrm{Hom}(X_1, X_{k+1}),$$

$k \geq 1$, of degree $2-k$, satisfying the condition

$$\sum_{r=1}^{n} \sum_{s=1}^{n-r+1} (-1)^{\varepsilon} m_{n-r+1}\big(a_1 \otimes \cdots \otimes a_{s-1} \otimes m_r(a_s \otimes \cdots \otimes a_{s+r-1}) \otimes a_{s+r} \otimes \cdots \otimes a_n\big) = 0$$

for all $n \geq 1$, where $\varepsilon = (r+1)s + r(n + \sum_{j=1}^{s-1} \deg(a_j))$.

An A^∞ category with one object is called an A^∞ algebra. The first condition ($n = 1$) says that m_1 is a degree 1 operator satisfying $(m_1)^2 = 0$, so it is a co-boundary operator which we can denote d. The second condition says that m_2 is a degree 0 map satisfying $d(m_2(a_1 \otimes a_2)) = m_2(da_1 \otimes a_2) + (-1)^{\deg(a_1)} m_2(a_1 \otimes da_2)$, so m_2 is a morphism of complexes and induces a product on cohomologies. The third condition says that m_2 is associative at the level of cohomologies.

The A^∞ structure on Fukaya's category is given by summing over holomorphic maps (up to projective equivalence) from the disc D, which take the components of the boundary $S^1 = \partial D$ to the special Lagrangian objects. An element u_j of $\mathrm{Hom}(\mathcal{U}_j, \mathcal{U}_{j+1})$ is represented by a pair

$$u_j = t_j \cdot a_j,$$

where $a_j \in L_j \cap L_{j+1}$, and t_j is a matrix in $\mathrm{Hom}(E_j|_{a_j}, E_{j+1}|_{a_j})$.

$$m_k(u_1 \otimes \cdots \otimes u_k) = \sum_{a_{k+1} \in L_1 \cap L_{k+1}} C(u_1, \ldots, u_k, a_{k+1}) \cdot a_{k+1},$$

where (notation explained below)

$$C(u_1, \ldots, u_k, a_{k+1}) = \sum_{\phi} \pm \mathrm{e}^{2\pi i \int \phi^* \omega} \cdot P\mathrm{e}^{\oint \phi^* \beta} \tag{37.8}$$

is a matrix in $\mathrm{Hom}(E_1|_{a_{k+1}}, E_{k+1}|_{a_{k+1}})$. Here we sum over (anti-)holomorphic maps $\phi : D \to \widetilde{M}$, up to projective equivalence, with the following conditions along the boundary (see Fig. 3): there are $k+1$ points $p_j = \mathrm{e}^{-2\pi i \alpha_j}$ such that $\phi(p_j) = a_j$ and $\phi(\mathrm{e}^{-2\pi i \alpha}) \in \mathcal{L}_j$ for $\alpha \in (\alpha_{j-1}, \alpha_j)$. In the above, $\omega = b + ik$ is the complexified Kähler form, the sign is determined by an orientation in the space of holomorphic maps (it will always be positive for us), and

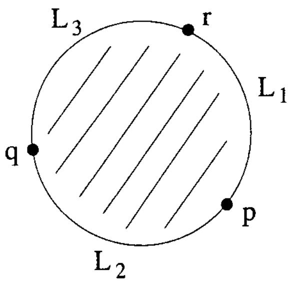


FIGURE 3. Holomorphic maps contributing to $m_2((p,a),(q,b))$, where $(p,a) \in \mathrm{Hom}(\mathcal{U}_1,\mathcal{U}_2)$ — i.e., $p \in L_1 \cap L_2$ and $a \in \mathrm{Hom}(E_1,E_2)|_p$, — and $(q,b) \in \mathrm{Hom}(\mathcal{U}_2,\mathcal{U}_3)$. The composition is given by a sum over maps ϕ satisfying the boundary conditions pictured, and is $m_2((p,a),(q,b)) = \sum_{r\in L_1\cap L_3} \pm e^{2\pi i \int \phi^*\omega} \cdot Pe^{2\pi i \oint \phi^*\beta}$, where β is the connection on the boundary formed by the connections on L_1, L_2, L_3, linked by p and q, as described in the text.

P represents a path-ordered integration, where β is the connection of the flat bundle along the local system on the boundary. Note that in the case of all trivial local systems ($\beta \equiv 0$), the weighting is just the exponentiated complexified area of the map. The path-ordered integral is defined by

$$Pe^{\oint \phi^*\beta} = Pe^{\int_{\alpha_k}^{\alpha_{k+1}} \beta_k d\alpha} \cdot t_k \cdot Pe^{\int_{\alpha_{k-1}}^{\alpha^k} \beta_{k-1} d\alpha} \cdot t_{k-1} \cdot \cdots \cdot t_1 \cdot Pe^{\int_{\alpha_{k+1}}^{\alpha_1} \beta_1 d\alpha}$$

(this formula is easily understood by reading right to left). Following the integration along the boundary, we get a homomorphism from $\mathcal{E}_1$ to $\mathcal{E}_{k+1}$.

Fukaya has shown the A^∞ structure of this category. We will skip this demonstration and point the reader to the references in Ch. 40. Now to define a true category which can be checked against the derived category, we simply take H^0 of all the morphisms (recall that they have the structure of complexes). The composition, previously homotopy associative, becomes associative. We call the resulting category $\mathcal{F}^0(\widetilde{M})$.

37.8. The Elliptic Curve

In this section we outline the proof of Kontsevich's version of mirror symmetry for the elliptic curve, or torus. This generalizes our discussion of

how mirror symmetry acts on D-branes in elliptic curves given in Ch. 19. The elliptic curve is a one-dimensional Calabi–Yau which is self-mirror — the complex parameter is exchanged with the complexified Kähler parameter by the mirror map.

The minimal submanifolds are just minimal lines, or geodesics (the Lagrangian property is trivially true for one-dimensional submanifolds). To define a closed submanifold in $\mathbb{R}^2/(\mathbb{Z}\oplus\mathbb{Z})$ the slope of the line must be rational, so can be given by a pair of integers (p,q). There is another real datum needed to define the line, which is the point of intersection with the line $x = 0$ (or $y = 0$ if $p = 0$). In the easiest case, the rank of the unitary system is 1, so that we can specify a flat line bundle on the circle by simply specifying the monodromy around the circle, i.e., a complex phase $\exp(2\pi i\beta)$, $\beta \in \mathbb{R}/\mathbb{Z}$. For a general local system of rank r we can take (p,q) to have greatest common divisor equal to r.

For our objects, we thus require more than the slope, which can be thought of as a complex phase with rational tangency, and therefore as

$$\exp i\pi\alpha.$$

We need a choice of α itself to make the Lagrangian a graded Lagrangian object. Clearly, the $\mathbb{Z}$-degeneracy represents the deck transformations of the universal cover of the space of slopes. Shifts by integers correspond to shifts by grading of the bounded complexes in the derived category. There is no natural choice of zero in this copy of $\mathbb{Z}$.

For our example, let us consider $\mathrm{Hom}(\mathcal{U}_i,\mathcal{U}_j)$, where the unitary systems have rank 1, and where the lines $\mathcal{L}_i$ and $\mathcal{L}_j$ go through the origin. Then $\tan\alpha_i = q/p$ and $\tan\alpha_j = s/r$, with (p,q) and (r,s) both relatively prime pairs. For simplicity, one can think of the lines as the infinite set of parallel lines on the universal cover of the torus, $\mathbb{R}\oplus\mathbb{R}$. It is then easy to see that there are

$$|ps - qr|$$

non-equivalent points of intersection. Since $\mathrm{Hom}(\mathbb{C},\mathbb{C})$ is one-dimensional, the monodromy is specified by a single complex phase T_i at each point of intersection. For rank n local systems, T_i would be represented by an $n\times n$ matrix.

The $\mathbb{Z}$-grading on $\mathrm{Hom}(\mathcal{U}_i,\mathcal{U}_j)$ is constant for all points of intersection in our example (they are all related by translation). If α_i, α_j are the real numbers representing the logarithms of the slopes, as above, then for $p \in L_i \cap L_j$ the grading is given by

$$\mu(p) = -[\alpha_j - \alpha_i],$$

where the brackets represent the greatest integer. Note that $-[x]-[-x] = 1$, which the Maslov index must obey for a onefold. The Maslov index is non-symmetric. For $p \in L_i \cap L_j$ in an n-fold, $\mu(p)_{ij} + \mu(p)_{ji} = n$. The asymmetry is reassuring, as we know that $\mathrm{Hom}(E_i, E_j)$ is not symmetric in the case of bundles. It is the extra data of the lift of the Lagrangian plane which allowed us to define the Maslov index in this way.

In our case of the elliptic curve, since $m_1 = d = 0$ the cohomology complex is the same as the original complex, and so we simply take the degree 0 piece of Hom. Since m_2 has degree 0, it survives this restriction and is associative. The higher m's are projected to zero in this category, so our equivalence will be defined by constructing a dictionary of objects and checking compatibility with m_2.

37.8.1. Categorical Equivalence. To construct an equivalence of categories, one would like to have a bijective functor relating objects and Hom's, compatible with compositions. In this section, we will describe this functor for the elliptic curve. As we will see, the construction is rather *ad hoc*. It would be preferable to have a more general construction of such a functor. In the subsequent section, we shall take up this point of view.

Here we wish to show the mirror relation $E_\tau \leftrightarrow E^\tau$, where E_τ represents an elliptic curve with modular parameter τ: $E_\tau = \mathbb{C}/(\mathbb{Z} \oplus \tau\mathbb{Z})$ and E^τ is the torus $\mathbb{R}^2/(\mathbb{Z} \oplus \mathbb{Z})$ with complexified symplectic form $\tau dx \wedge dy$. In fact, we will exhibit only the functorial map of categories and the equivalence of compositions for a simple example. The interested reader can find reference to a more detailed proof at the end of the chapter.

The derived category simplifies on the elliptic curve. In fact, any complex of coherent sheaves is quasi-isomorphic to the trivial complex formed from its cohomology. This means that the complex is essentially a "direct sum" of single-term complexes (which we can think of as sheaves), up to shifts in grading. We will ignore the shifts in grading, then, and focus on sheaves. In addition, sheaves on elliptic curves can be decomposed into a direct sum of honest bundles plus sheaves supported on points.

At its most basic, the dictionary of objects maps slope to slope. The slope of a bundle V is defined to be d/r, where $d = \int_E c_1(V)$ is the "degree" and r is the rank. We want to map this object to a line with slope d/r. As a check, let us calculate the dimension of $\mathrm{Hom}(V, V') = H^0(V^* \otimes V')$. The Grothendieck–Riemann–Roch formula tells us that if H^1 is zero, this equals

$$\int_E \mathrm{ch}(V^* \otimes V') = \int_E (r - dx)(r' + d'x) = rd' - dr',$$

where x is Poincaré dual to a point ($\int_E x = 1$). Now two lines of slope d/r and d'/r' can be described by vectors between lattice points with coordinates (d, r) and (d', r'). These make a parallelogram with area $|dr' - rd'|$, which counts the number of intersection points in the quotiented torus. In fact, the signs work out that the intersection points are in H^0 precisely when the bundles have global Hom's (recall that the lines are "graded," so intersection points contribute to either H^0 or H^1 depending on which slope is greater).

For the moment we will ignore putting monodromies on the Lagrangians or moving them away from lattice points. This corresponds to the fact that we will ignore the possibility of tensoring our bundles by line bundles of degree 0. The two real parameters describing these choices on each side can be explicitly mapped to the parameters on the other side.

Let us check whether this assignment of objects is compatible with the composition of Hom's. We try a simple example. Let $\mathcal{L}_0$ be the line of slope 0 (through the origin). Let $\mathcal{L}_1$ be the line of slope 1, and $\mathcal{L}_2$ be the line of slope 2. Note $\mathcal{L}_0 \cap \mathcal{L}_1 = p_{01}$, where p_{01} is the origin (we retain the intersection label). Now $\mathcal{L}_1 \cap \mathcal{L}_2 = p_{12}$, where p_{12} is also the origin. Finally, $\mathcal{L}_0 \cap \mathcal{L}_2 = \{p_{02}, q_{02}\}$, where p_{02} is the origin and q_{02} is the point $(1/2, 0) \in \mathbb{R}^2/\mathbb{Z}^2$. Composition in the Fukaya category takes the form

$$p_{01} \otimes p_{12} = C_{ppp} \cdot p_{02} + C_{ppq} \cdot q_{02}.$$

Let us now compute the coefficients C_{ppp} and C_{ppq}. We will need to sum over triangles in the plane bounded by lines of slope 0, 1, 2. These triangles represent the image of holomorphic maps from the disk. More precisely, we need to then quotient from the plane (universal cover) to the torus. The corners of the triangle have to sit at lattice points for C_{ppp}, while the intersection of the slope 0 and slope 2 lines must sit at $(n+1/2, m)$ for C_{ppq}. Without loss of generality we can take $\mathcal{L}_0$ and $\mathcal{L}_1$ to intersect at the origin, $(0,0)$. Then the triangle is completely determined by the intersection of $\mathcal{L}_0$

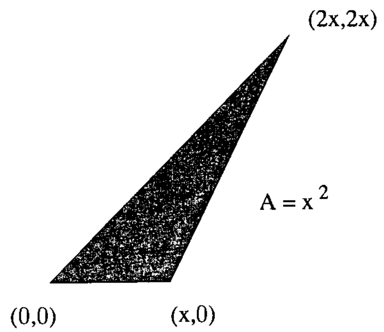


FIGURE 4. Triangles in the sum of C_{ppp} ($x \in \mathbb{Z}$) and C_{ppq} ($x \in \mathbb{Z} + \frac{1}{2}$).

and $\mathcal{L}_2$ at $(n,0)$ for C_{ppp} or $(n+1/2,0)$ for C_{ppq} (see Fig. 4). The "area" of these triangles is n^2 (resp. $(n+1/2)^2$), which we must multiply by τ according to the symplectic form $\tau dx \wedge dy$. Then

$$C_{ppp} = \sum_{n=-\infty}^{\infty} e^{i\pi 2\tau n^2}, \qquad C_{ppq} = \sum_{n=-\infty}^{\infty} e^{i\pi 2\tau (n+1/2)^2}.$$

According to our dictionary, the corresponding bundles are $L_0 \equiv \mathcal{O}$, the trivial (degree 0) line bundle; $L_1 \equiv \mathcal{O}(1)$, the degree 1 Θ-line bundle, i.e., the line bundle on E_τ whose unique section is the theta function, $\Theta[0,0]$;[13] and $L_2 \equiv \mathcal{O}(2)$.

We wish to compute the map

$$\mathrm{Hom}^0(\mathcal{O}, L) \otimes \mathrm{Hom}^0(L, L^2) \to \mathrm{Hom}^0(\mathcal{O}, L^2).$$

First note $\mathrm{Hom}^0(\mathcal{O}, L) = H^0(\mathcal{O}^* \otimes L) = H^0(L)$ which is one-dimensional, generated by $\Theta[0,0](\tau, z)$. This section also generates $\mathrm{Hom}^0(L, L^2) = H^0(L)$. Therefore, the correspondence $p \leftrightarrow \Theta[0,0]$ makes sense for p_{01} and p_{12} (and p_{02} as we shall see). Finally $\mathrm{Hom}^0(\mathcal{O}, L^2) = H^0(L^2)$ is two-dimensional, generated by $\Theta[0,0](2\tau, 2z)$ and $\Theta[1/2,0](2\tau, 2z)$. (Note the change in argument.) To complete the correspondence we should map p_{02} to $\Theta[0,0]$ and

[13] To fix notation, we let L be the line bundle on E_τ defined by $\mathbb{C}^* \times \mathbb{C}/((u,v) \sim (uq, \varphi v)$, where $u = \exp(2\pi i z)$ is a coordinate on $\mathbb{C}^*$, $q = \exp(2\pi i \tau)$, and $\varphi = q^{-1/2}u^{-1}$. Define $\Theta[a,b](\tau,z) = \sum_{n\in\mathbb{Z}} \exp(i\pi\tau(n+a)^2 + 2i\pi(n+a)(z+b))$. Then $\Theta \equiv \Theta[0,0]$ is a section of L.

$q_{02} \leftrightarrow \Theta[1/2,0]$. Composition of Θ functions is by multiplication, and the classical theta function identity reveals the product decomposition

$$\begin{aligned}\Theta[0,0](\tau,z)\cdot\Theta[0,0](\tau,z) &= \Theta[0,0](2\tau,0)\cdot\Theta[0,0](2\tau,2z)\\ &\quad+ \Theta[1/2,0](2\tau,0)\cdot\Theta[1/2,0](2\tau,2z).\end{aligned}$$

One checks explicitly from the form of the theta function given in the footnote ([13]) that

$$\Theta[0,0](2\tau,0) = C_{ppp}; \qquad \Theta[1/2,0](2\tau,0) = C_{ppq}.$$

Fukaya composition and composition in the derived category are isomorphic! The computation on the derived category side can be viewed as the suitable reduction of the cubic term of holomorphic Chern-Simons $\int \Omega\wedge \mathrm{Tr}\, A[A,A]$ for the CY threefold, to the case of the CY onefold, leading to $\int \Omega\wedge\mathrm{Tr}\,\Phi_1[A,\Phi_2]$ where $\Phi_{1,2}$ are $U(N)$ adjoint-valued sections of the trivial bundle on the elliptic curve and the A is a $U(N)$ connection on the elliptic curve. For the case at hand, where we have considered three branes, $N=3$.

Next one wants to move the lines away from lattice points and put in $U(1)$ holonomies. We consider changing the line $\mathcal{L}_2$, for example. If we move it off the lattice an amount α, the sum defining C_{ppp} will be over $\exp[i\pi 2\tau(n+\alpha)^2]$. If we add a flat connection βdx, this will appear as a phase $2\pi i(n+\alpha)(\beta)$. One sees that α and β parametrize shifts of the θ functions describing C_{ppp} and C_{ppq}, a reflection of the fact that the corresponding line bundles are shifted by tensoring with degree 0 line bundles. The reader can find more details in the references cited at the end of this chapter.

The extension to arbitrary bundles is achieved by expressing certain higher-rank bundles as push-forwards of line bundles under finite covers (isogenies) of the elliptic curve. The push-forward of $\mathcal{O}$ under an r-fold cover is $\mathcal{O}^{\oplus r}$, for example. The corresponding lines have higher-rank local systems. For sheaves supported at points, our slope-slope correspondence tells us that they are equivalent to vertical lines. One checks compositions, etc. This turns out to be enough to show the equivalence for the whole derived category, thus proving Kontsevich's conjecture in this simple but illustrative case.

One could also generalize the check of the topological string computation to other topologies and in particular to the case of the annulus. For D-branes on elliptic curves this is done in Ch. 35.

37.9. A Geometric Functor

In this section we show that it may be possible to define a natural functor relating special Lagrangian objects to holomorphic sheaves. For simplicity, we will restrict ourselves to some simplifying assumptions. Following the reasoning of Sec. 37.4, we assume that our Calabi–Yau manifold has a special Lagrangian torus fibration, and that the mirror manifold is constructed by dualizing the fibration. Further, we take the tori to be flat. We also take the base of the fibration to be non-compact, studying only the local (on the base) geometric properties of the functor. With these assumptions, we will define a natural way of going from A-cycles to B-cycles via a real family version of Fourier–Mukai transform.

Our goal will be to use the structure of dual fibrations to relate special Lagrangians plus flat $U(1)$ bundles to bundles with connections satisfying the deformed Hermitian-Yang–Mills equations. We limit ourselves to Lagrangians that are sections of the fibration, as in Fig. 2. The dual object is therefore a six-brane, so we will ignore any strictly sheaf-theoretic issues and focus on honest bundles. Such a map of objects would be an important first step in the general construction of a bijective functor, proving Kontsevich's conjecture.

The transform we construct is simple-minded. A special Lagrangian section intersects each fiber at a single point. That point represents a line bundle over the dual fiber. The collection of line bundles pieces together to form a line bundle over the dual family of fibers — i.e., the mirror.

We have our Calabi–Yau M fibered over a base B by flat special Lagrangian tori T, with x^i coordinates for the base and y^i coordinates for the fiber, $i = 1,\ldots,3$. The Kähler potential ϕ for M is then taken to be independent of the y^i, i.e. $\phi = \phi(x)$. The complex coordinate is $z^j = x^j + iy^j$.

As studied by Calabi, the Ricci tensor vanishes and $\Omega = dz^1 \wedge \ldots \wedge dz^m$ is a covariant constant if and only if ϕ satisfies a real Monge–Ampère equation

$$\det \frac{\partial^2 \phi}{\partial x^i \partial x^j} = const.$$

The Ricci-flat Kähler metric and form are

$$g = \sum_{i,j} \frac{\partial^2 \phi}{\partial x^i \partial x^j}\left(dx^i dx^j + dy^i dy^j\right),$$

and

$$\omega = \frac{i}{2}\sum_{i,j} \frac{\partial^2 \phi}{\partial x^i \partial x^j} dz^i \wedge d\bar{z}^j,$$

(henceforth we sum over repeated indices). Notice that $\Omega \wedge \bar{\Omega}$ is a constant multiple of ω^m, and it is a direct consequence of the real Monge-Ampère equation.

Also note from the form of the metric g that M is locally isometric to the tangent bundle of B with its metric induced from the metric $\sum_{i,j} \frac{\partial^2 \phi}{\partial x^i \partial x^j} dx^i dx^j$ on B. If we use the metric on B to identify its tangent bundle with its cotangent bundle, then the above symplectic form ω is just the canonical symplectic form $dp \wedge dq$ on the cotangent bundle.

We can view the universal cover of M either as TB with the standard complex structure, or as T^*B with the standard symplectic structure. A solution of the real Monge-Ampère equation is used to determine the symplectic structure in the former case and to determine the metric structure, and therefore the complex structure, in the latter case.

We will construct the transform for a special Lagrangian exhibited as a section of the fibration, i.e., a graph over the base.

Recall that a section of T^*B is Lagrangian with respect to the standard symplectic form if and only if it is a closed one-form, and hence locally exact. Therefore (or by calculation), a graph $y(x)$ in M is Lagrangian with respect to ω if and only if $\frac{\partial}{\partial x^j}(y^l \phi_{lk}) = \frac{\partial}{\partial x^k}(\phi_{lj} y^l)$, where $\phi_{ij} = \frac{\partial^2 \phi}{\partial x^i \partial x^j}$, from which we get

$$y^j = \phi^{jk} \frac{\partial f}{\partial x^k}$$

for some function f (locally), where ϕ^{jk} is the inverse matrix of ϕ_{jk}.

Now $dz^j = dx^j + i dy^j$ and on C we have $dy^j = \phi^{jl}\left(\frac{\partial^2 f}{\partial x^l \partial x^k} - \phi^{pq}\phi_{lkp}\frac{\partial f}{\partial x^q}\right)dx^k$. Therefore $dz^j = \left(\delta_{jk} + i\phi^{jl}\left(\frac{\partial^2 f}{\partial x^l \partial x^k} - \phi^{pq}\phi_{lkp}\frac{\partial f}{\partial x^q}\right)\right)dx^k$ over C. Notice that if we write $g = \phi_{jk} dx^j dx^k$ as the Riemannian metric on the base, then the Christoffel symbol for the Levi–Civita connection is $\Gamma^q_{lk} = \phi^{pq}\phi_{lkp}$. Therefore $\mathrm{Hess}\,(f) = \left(\frac{\partial^2 f}{\partial x^l \partial x^k} - \phi^{pq}\phi_{lkp}\frac{\partial f}{\partial x^q}\right) dx^l dx^k$. Hence

$$\begin{aligned} dz^1 \wedge \cdots \wedge dz^m|_C &= \det\left(I + i g^{-1}\mathrm{Hess}\,(f)\right) dx^1 \wedge \cdots \wedge dx^m \\ &= \det(g)^{-1}\det\left(g + i\mathrm{Hess}\,(f)\right) dx^1 \wedge \cdots \wedge dx^m, \end{aligned}$$

so the special Lagrangian condition (with phase) $\operatorname{Im}(dz^1 \wedge \dots \wedge dz^m)|_C = \tan\theta \cdot \operatorname{Re}(dz^1 \wedge \dots \wedge dz^m)|_C$ becomes

$$\operatorname{Im}\det(g + i\operatorname{Hess}(f)) = (\tan\theta)\operatorname{Re}\det(g + i\operatorname{Hess}(f)).$$

From these data, we want to construct a $U(1)$ connection over the mirror manifold W that satisfies the deformed Hermitian-Yang-Mills equation. The dual manifold W is constructed by replacing each torus fiber T in M by the dual torus $\widetilde{T} = \operatorname{Hom}(T, S^1)$. If we write the dual coordinates to $y^1, \dots, y^m$ as $\widetilde{y}_1, \dots, \widetilde{y}_m$, then the dual Riemannian metric on W is obtained by taking the dual metric on each dual torus fiber $\widetilde{T}$:

$$\widetilde{g} = \sum_{i,j} \left(\phi_{ij} dx^i dx^j + \phi^{ij} d\widetilde{y}_i d\widetilde{y}_j\right).$$

We need to understand the complex structure and the symplectic structure on W. First we rewrite $\widetilde{g}$ as

$$\widetilde{g} = \sum_{i,j} \phi^{ij}\left(\left(\Sigma_k \phi_{ik} dx^k\right)\left(\Sigma_l \phi_{jl} dx^l\right) + d\widetilde{y}_i d\widetilde{y}_j\right).$$

Notice that $d\left(\sum_k \phi_{jk} dx^k\right) = 0$ because ϕ_{jkl} is symmetric with respect to interchanging the indices. Therefore there exist functions $\widetilde{x}_j = \widetilde{x}_j(x)$ such that $d\widetilde{x}_j = \sum_k \phi_{jk} dx^k$ locally — then $\frac{\partial \widetilde{x}_j}{\partial x^k} = \phi_{jk}$ — and we obtain

$$\widetilde{g} = \sum_{i,j} \phi^{ij}\left(d\widetilde{x}_i d\widetilde{x}_j + d\widetilde{y}_i d\widetilde{y}_j\right).$$

So we can use $\widetilde{z}_j = \widetilde{x}_j + i\widetilde{y}_j$'s as complex coordinates on W. It is easy to check that the corresponding symplectic form is given by

$$\widetilde{\omega} = \frac{i}{2} \sum_{i,j} \phi^{ij} d\widetilde{z}_i \wedge d\overline{\widetilde{z}}_j.$$

Moreover the covariant constant holomorphic m-form on W is given by

$$\widetilde{\Omega} = d\widetilde{z}_1 \wedge \dots \wedge d\widetilde{z}_m.$$

Again, as a direct consequence of ϕ being a solution of the real Monge-Ampère equation, $\widetilde{\Omega} \wedge \overline{\widetilde{\Omega}}$ is a constant multiple of $\widetilde{\omega}^m$.

REMARK 37.9.1. *The mirror manifold W can be interpreted as the moduli space of special Lagrangian tori together with flat $U(1)$ connections. It can be checked directly that the L^2 metric, i.e., the Weil–Petersson metric, on this moduli space W coincides with our $\widetilde{g}$ above.*

In general, the relevant metric on the moduli space W is given by a two-point function computed via a path integral, which includes instanton contributions from holomorphic disks bounding the special Lagrangian torus fibers. However, for our local Calabi–Yau M such holomorphic disks do not exist. This is because M is homotopic to any one of its fibers; but any such holomorphic disk would define a non-trivial relative homology class. Therefore the metric $\widetilde{g}$ coincides with the physical metric on the moduli space W.

REMARK 37.9.2. *We note the symmetry between g (resp. ω) and $\widetilde{g}$ (resp. $\widetilde{\omega}$). For one can write ϕ^{ij} as the second derivative of some function $\widetilde{\phi}$ with respect to the $\widetilde{x}_j$'s. Simply write $x^j = x^j(\widetilde{x})$, then $\frac{\partial x^j}{\partial \widetilde{x}_k} = \phi^{jk} = \frac{\partial x^k}{\partial \widetilde{x}_j}$ and therefore $x^j = \frac{\partial \Phi}{\partial \widetilde{x}_j}$ for some function, Φ, and it is easy to check that $\widetilde{\phi} = \Phi$.*

On each torus fiber, we have canonical isomorphisms $T = \operatorname{Hom}(\widetilde{T}, S^1) = \operatorname{Hom}(\pi_1(\widetilde{T}), S^1)$, therefore a point $y = (y^1, \dots, y^m)$ in T defines a flat connection D_y on its dual $\widetilde{T}$. This is the real Fourier–Mukai transform. Explicitly, we have

$$\begin{aligned} g_y: \ & \widetilde{T} \to i(\mathbb{R}/\mathbb{Z}) = S^1, \\ & \widetilde{y} \mapsto i \textstyle\sum_{j=1}^m y^j \widetilde{y}_j, \end{aligned}$$

and $D_y = d + A = d + i dg_y = d + i \sum y^j d\widetilde{y}_j$.

In fact we get a torus family of one-forms, since y (hence A) has x- (or $\widetilde{x}$-) dependence. Namely, we obtain a $U(1)$ connection on W,

$$D_A = d + i \sum_j y^j d\widetilde{y}_j.$$

Its curvature two-form is given by,

$$F_A = dA = \sum_{k,j} i \frac{\partial y^j}{\partial \widetilde{x}_k} d\widetilde{x}_k \wedge d\widetilde{y}_j.$$

In particular

$$F_A^{2,0} = \frac{1}{2} \sum_{j,k} \left(\frac{\partial y^k}{\partial \widetilde{x}_j} - \frac{\partial y^j}{\partial \widetilde{x}_k}\right) d\widetilde{z}_j \wedge d\widetilde{z}_k.$$

Therefore, that D_A is integrable, i.e., $F_A^{0,2} = 0$, is equivalent to the existence of $f = f(\widetilde{x})$ such that $y^j = \frac{\partial f}{\partial \widetilde{x}_j} = \phi^{jk} \frac{\partial f}{\partial x^k}$ because of $d\widetilde{x}_j = \sum_k \phi_{jk} dx^k$. Namely, the cycle $C \subset M$ must be Lagrangian. Now

$$\frac{\partial y^j}{\partial \widetilde{x}_k} = \frac{\partial^2 f}{\partial \widetilde{x}_j \partial \widetilde{x}_k}.$$

In terms of the x variable, this is precisely the Hessian of f, as discussed above. Therefore the cycle $C \subset M$ being special is equivalent to

$$\mathrm{Im}\,(\widetilde{\omega} + F_A)^m = (\tan\theta)\, Re\,(\widetilde{\omega} + F_A)^m .$$

For a general type-A supersymmetric cycle in M, we have a special Lagrangian C in M together with a flat $U(1)$ connection on it. Since as before, C is expressed as a section of $\pi : M \to B$ and is given by $y^j = \phi^{jk}\frac{\partial f}{\partial x^k}$, a flat $U(1)$ connection on C can be written in the form $d + ide = d + i\sum \frac{\partial e}{\partial x^k}dx^k$ for some function $e = e(x)$. Recall that the transformation of C alone is the connection $d + i\sum y^j d\widetilde{y}_j$ over W. When the flat connection on C is also taken into account, then the total transformation becomes

$$\begin{aligned} D_A &= d + i\sum y^j d\widetilde{y}_j + ide \\ &= d + i\sum \phi^{jk}\frac{\partial f}{\partial x^k}d\widetilde{y}_j + i\sum \frac{\partial e}{\partial \widetilde{x}^j}d\widetilde{x}_j. \end{aligned}$$

Here we have composed the function $e(x)$ with the coordinate transformation $x = x(\widetilde{x})$. Notice that the added term $\sum \frac{\partial e}{\partial \widetilde{x}^j}d\widetilde{x}_j$ is exact and therefore the curvature form of this new connection is the same as the old one. In particular D_A satisfies

$$\begin{aligned} F_A^{0,2} &= 0, \\ \mathrm{Im}\, e^{i\theta}(\omega + F)^m &= 0, \end{aligned}$$

so is a supersymmetric cycle of type B in W. By the same reasoning, we can couple with C a flat connection on it of *any* rank and we would still obtain a non-abelian connection D_A on W satisfying the above equations.

In conclusion, the transformation of an A-cycle section is a B-cycle bundle. Other cases may be considered as well, but defining this transform in the general case — including singular and non-flat fibers, non-section and multi-section A-cycles, and non-transversal intersection of fibers — is a formidable problem.

37.10. The Correspondence Principle

In 1923 Niels Bohr asserted that the predictions of quantum mechanics must agree with results from classical dynamics in regimes where classical physics is appropriate. Indeed, a kickball is not very quantum mechanical. Likewise, relativity must reduce to Newtonian mechanics when $v/c \ll 1$.

This is known as the "correspondence principle." If we are to believe Kontsevich's definition of what mirror symmetry is, it must incorporate all the stunning predictions of the classical picture of mirror symmetry: Gromov–Witten invariants, Picard–Fuchs differential equations, and all that. It also should recover other developments such as Lagrangian fibrations, insofar as they are valid. In this section, following Kontsevich, we give a rather heuristic description of how such a correspondence might be achieved. We also discuss some other desirable features of Kontsevich's viewpoint.

37.10.1. Motivation. As already noted, the conjecture is closely related to how mirror symmetry works in the context of D-branes.

Another motivation of the conjecture comes from trying to understand the moduli space of topological quantum field theories. As we know, the observables correspond to cohomology classes $H^{p,q}(M)$ in the A-model and to $H^p(\Lambda^q T\widetilde{M})$ in the B-model. As these states can be used to deform the action, they can be considered as tangent vectors to the full moduli space of topological quantum field theories (if we include gravity at higher genus, there are many more operators — the "gravitational descendants"). However, classical mirror symmetry calculations involve only the Kähler cone (with directions $H^{1,1}(M)$) in the A-model and the moduli space of Calabi–Yau's (with directions $H^1(T\widetilde{M}) \cong H^{2,1}(\widetilde{M})$) in the B-model. Perhaps motivated by the fact that Fukaya had found an A_∞ structure in the open string, Kontsevich considers deformations of the algebraic structure of the Calabi–Yau which are not necessarily associative. Let us recall ordinary deformations of algebras. Consider an algebra, A, with its composition $m_2 : A \otimes A \to A$. Infinitesimal deformations of the algebraic structure m_2 can be found by putting $m_2 \to m_2 + \epsilon_2$, and linearizing the associativity constraint. Deformations modulo basis redefinitions are classified by the second Hochschild cohomology $HH^2(A)$, which has an equivalent description as $\mathrm{Ext}^2(A, A)$, where A is taken as an A-bimodule (acting on the left and right). To extend this to the algebraic structure on the sheaf of holomorphic functions, consider $\mathcal{O}_\Delta$, which is the structure sheaf of the diagonal in $M \times M$. Then $\mathcal{O}_\Delta$ is an $\mathcal{O}_M$ bimodule, and behaves much like A.[14] Kontsevich then shows that the deformations of the structure sheaf as

[14] $\mathcal{O}_\Delta$ acts as the identity functor on sheaves, $\mathcal{S}$. If we pull back $\mathcal{S}$ to $M \times M$, tensor with $\mathcal{O}_\Delta$, and push forward to M, we find $\pi_{2*}(\pi_1^*\mathcal{S} \otimes \mathcal{O}_\Delta) \cong \mathcal{S}$.

a sheaf of A_∞ algebras, which are classified by $\mathrm{Ext}^*(\mathcal{O}_\Delta, \mathcal{O}_\Delta)$, are equal to $\bigoplus_{p+q=n} H^p(\Lambda^q TM)$.

37.10.2. Recovering Classical Mirror Symmetry. Recall that classical mirror symmetry involves equating three-point functions of marginal operators, i.e., $H^1(TM)$ or $H^{1,1}(\widetilde{M})$ on the B/A side, respectively. As we saw above, we can think of $H^1(TM)$ as $\mathrm{Ext}^2(\mathcal{O}_\Delta, \mathcal{O}_\Delta)$, i.e. as an element of $\mathrm{Hom}^2(M, M)$. Therefore, we should think of $M \subset M \times M$ as a holomorphic submanifold. On the mirror side, the corresponding object is $\widetilde{M} \subset \widetilde{M} \times \overline{\widetilde{M}}$, and is a special Lagrangian submanifold. Note that we have to reverse the complex structure of the second $\widetilde{M}$. In the Fukaya category, then, we have to make sense of $\mathrm{Hom}^2(\widetilde{M}, \widetilde{M})$. Although the intersection is not finite or transverse, we can make sense of it by deforming the second $\widetilde{M}$, say, by a Morse function, f. If we do so, the Fukaya Hom complex becomes the ordinary Morse complex on $\widetilde{M}$ (cf. Sec. 3.4) and its cohomology can be identified with $H^2(\widetilde{M}) = H^{1,1}(\widetilde{M})$, as required for mirror symmetry. In couplings we will take the limit $f \to 0$.

$M \times M$	$\widetilde{M} \times \overline{\widetilde{M}}$
$\mathcal{O}_\Delta$	$\widetilde{\Delta}$
$\mathrm{Ext}^2(\mathcal{O}_\Delta, \mathcal{O}_\Delta)$	$\mathrm{Hom}^2(\widetilde{\Delta}, \widetilde{\Delta})$
$H^1(TM)$	Pick Morse function f Hom complex $\leftrightarrow$ Morse complex $\Rightarrow \mathrm{Hom}^2(\widetilde{M}, \widetilde{M}) = H^2(\widetilde{M})$
Yukawa pairing	disk instanton sum $\phi : D \to \widetilde{M} \times \overline{\widetilde{M}}$ $=$ ordinary instanton $\phi : \mathbb{P}^1 \to \widetilde{M}$

TABLE 3. Calculations on the diagonal should recover ordinary mirror symmetry.

Now the three-point coupling in the B-model is the classical Yukawa pairing. For the A-cycle $\widetilde{M}$ the triple pairing involves disk instantons. Now note that a disk instanton in $\widetilde{M} \times \overline{\widetilde{M}}$ is a pair of maps $(\phi_1, \overline{\phi_2})$ where $\phi_i : D \to \widetilde{M}$, such that $\phi_1 = \phi_2$ on the boundary. Gluing ϕ_1 and $\overline{\phi_2}$ along the boundary, we can construct an ordinary holomorphic map from the sphere. Since the Hom's are defined with reference to a perturbing function (and though we ultimately take the limit $f \to 0$) There are conditions at each of three points along the boundary of the disk – that they get mapped to critical points of f (or sets of critical points) representing the cohomology classes of the operators inserted there. It is not clear (to the author) how these point conditions relate to the usual conditions of mapping points to the Poincaré dual homology cycles of the operators at the insertions.

37.10.3. Lagrangians and Bundles. Recovering the picture of mirror symmetry as T-duality can be seen by simply relating the structure sheaf of a point (the zero-brane) to a corresponding special Lagrangian three-brane. All moduli spaces of Lagrangian D-branes have torus fibrations, just as in the argument of section 4. Further, the deformations of the zero-brane moduli space are simply tangent vectors, and their two-point function is the metric. On the Fukaya side, the two-point function includes all the corrections from instantons.

More generally the computation of topological strings in the presence of D-branes, and the use of mirror symmetry in this context is rather natural. In fact recent work has led to enumerative predictions of disc instantons which have been verified (though not yet proven) through mathematical computations involving localization methods.

All told, Kontsevich's conjecture, suitably generalized, is likely to be valid. However, much remains to be explained or constructed: the role of tachyon condensation and how it relates to quasi-isomorphisms in the derived category; a functor from the two categories defined for compact, non-semi-flat Calabi–Yau manifolds; a proper definition of the moduli space of D-branes; a derivation and mathematical proof of boundary enumerative predictions involving maps from Riemann surfaces with boundaries, along the lines of classical mirror symmetry; etc.

Mirror symmetry may be somewhat well described at this stage, but many results – both mathematical and physical – remain to be uncovered through a unified understanding of this phenomenon.

CHAPTER 38

More on the Mathematics of D-branes: Bundles, Derived Categories, and Lagrangians

38.1. Introduction

So far in this book we have built up a good picture of mirror symmetry, with a dictionary of correspondences under the mirror map as follows, at least in the geometric regime near the large complex and Kähler limit points.

Complex n-fold M	Symplectic $2n$-manifold W
$\Omega = \Omega_M \in H^0(K_M) \cong H^{n,0}$	$\omega = \omega_W \in H^{1,1}$ s.t. $c_1 = 0$
B-model	A-model
Periods, VHS	GW-invariants
H^{ev}	H^n
$D^b(M)$	$\mathrm{Fuk}\,(W)$
$\omega = \omega_M \in H^{1,1}$	$\Omega = \Omega_W \in H^{n,0}$
even-dimensional branes	n-branes

(Note that if n is even we may have to modify the RHS by adding in all even-dimensional branes, and also that the GW-invariants of Ch. 26 really are symplectic invariants only; they do not depend on a complex structure (though one is often used to define or compute them). We are also ignoring the B-field for the purposes of this chapter – *all of our symplectic forms are real.*)

There is an important difference between the sixth and eighth lines which we have blurred until now, and which will be the main theme of this chapter. Kontsevich's conjecture, discussed in Sec. 37.7, involved *only the complex structure* on M, and just the symplectic (or "Kähler") structure on W, while to introduce the D-branes of Chs. 19 and 37 we need both types of structure on both sides. The conjecture deals just with holomorphic bundles (more generally, complexes of them up to quasi-isomorphism; this then includes the coherent sheaves of Sec. 37.6.2) on the B-model LHS and Lagrangians (up to Hamiltonian deformations, which we will discuss later, with a grading and a flat unitary connection on them) for the A-model RHS. (We are using Fukaya's original A^∞-category, not the modified sLag Fukaya category of Sec. 37.7.1; the close relationship between the two will become clearer by the end of this Chapter.) D-branes, however, are given on the LHS by connections on a bundle (perhaps supported on some subvariety) satisfying the MMMS (Sec. 37.3.1) or HYM equations (Eq. (38.2)) near the large Kähler limit (where the MMMS equations degenerate to the HYM equations). On the one hand this implies that the underlying bundle is holomorphic and so defines an object of $D^b(M)$, but on the other hand it contains more information, and, crucially, is *dependent on the introduction of a Kähler form* ω_M. Similarly A-model D-branes are special Lagrangian and so, in particular, are Lagrangian, but require a complex structure on W for their definition, and different choices will give different results. The sense in which D-branes can be identified with the objects of Kontsevich's conjecture, and the sense in which they are different, is subtle but important, and leads us to some interesting predictions. (We have been able to ignore it so far, for instance in Sec. 37.8, because, as we shall see, the subtlety does not really arise for T^2; almost all bundles are direct sums of bundles with HYM connections, for instance.) So we give an overview of the mathematics of stable bundles to compare and contrast with HYM connections. We will then give a rough outline of what the derived category $D^b(M)$ is, and explain why it is a much stronger invariant than say cohomology (of which it is a refinement via the Mukai vector) or D-branes; Kontsevich's idea is that one should be able to recover the whole B-model string theory from $D^b(M)$. Exploiting this on the symplectic (A-model) side will lead to a natural conjecture about the relationship between Lagrangians and special Lagrangians.

38.2. Holomorphic Bundles and Gauge Theory

We want to describe some of the now standard mathematics of gauge theory on holomorphic bundles, and the relationship between stable holomorphic bundles and HYM connections (all of these terms will be defined presently). In particular it is important that, for this chapter at least, the physics reader is *not to think of holomorphic bundles as instantons*, even if they are stable; for instance we will be interested in later sections in a pair of *different* holomorphic bundles whose associated HYM connections are the same in some non-degenerate limit of changing Kähler form. (Similar remarks will also apply to the MMMS equations.)

We briefly recall some facts from Sec. 1.3. Fix a complex manifold M. A rank r holomorphic bundle E on M is given by a collection of trivial bundles $U_i \times \mathbb{C}^r$ over an open cover $\{U_i\}_{i\in I}$ of M, patched together by *holomorphic* transition functions, i.e.,

$$\text{holomorphic maps } \phi_{ij} : U_i \cap U_j \to GL(r, \mathbb{C}). \tag{38.1}$$

Thus we can talk about holomorphic sections — locally these are r-tuples of holomorphic functions that patch under the ϕ_{ij}s; transforming a section by ϕ_{ij} does not affect holomorphicity by (38.1). Similarly the Cauchy-Riemann $\bar\partial$-operator patches together from the local operators (as $\bar\partial \circ \phi_{ij} = \phi_{ij} \circ \bar\partial$ by (38.1)) to give

$$\bar\partial_E : \Omega^0(E) \to \Omega^{0,1}(E),$$

whose kernel is precisely the holomorphic sections of E. Here $\Omega^{p,q}(E)$ denotes *smooth* (not necessarily holomorphic) forms of type (p,q) (i.e., $p\ dz_i$s and $q\ d\bar z_i$s in local coordinates) with values in E; thus $\Omega^0(E)$ is just smooth sections of E.

So a section s being holomorphic ($\bar\partial_E s = 0$) is predetermined in E, though not whether or not it is covariantly constant. Thus, the holomorphic structure defines *half a connection* on E; the $(0,1)$ (or $d\bar z_i$) part. That is, any connection

$$d_A : \Omega^0(E) \to \Omega^1(E) \cong \Omega^{1,0}(E) \oplus \Omega^{0,1}(E)$$

splits into $\partial_A \oplus \bar{\partial}_A$ according to the above decomposition, and for A to be compatible with the holomorphic structure of E we insist that $\bar{\partial}_A = \bar{\partial}_E$. If we put a Hermitian metric h on E, then a simple calculation shows that there exists a unique connection d_A compatible with both the metric and the holomorphic structure, that is $d_A(h) = 0$ and $\bar{\partial}_A = \bar{\partial}_E$; see Sec. 5.2.1. (If $(s_i)_{i=1}^r$ form a local holomorphic trivialisation of E in which $h = (h_{ij}) = h(s_i, s_j)$, then $d_A s_i = \partial_A s_i = \sum_{jk} \partial h_{ij}(h^{-1})_{jk} s_k$ and the Leibniz rule determine d_A.)

This connection has curvature

$$F_A = d_A^2 = F_A^{2,0} \oplus F_A^{1,1} \oplus F_A^{0,2} = \partial_A^2 \oplus (\partial_A \bar{\partial}_A + \bar{\partial}_A \partial_A) \oplus \bar{\partial}_A^2$$

with respect to the splitting of the two-forms (with values in $\operatorname{End} E$) $\Omega^2 = \Omega^{2,0} \oplus \Omega^{1,1} \oplus \Omega^{0,2}$. Of course $\bar{\partial}_E^2 = 0$ so that $F_A^{0,2} = 0$ (and so $F_A^{2,0} = 0$ by conjugation using the metric h).

There is also a converse to this (discussed in the more complicated case of the bundle being the tangent bundle in Sec. 5.2) — *a connection on any complex bundle satisfying $F_A^{0,2} = 0$ defines a holomorphic structure on that bundle*: $\bar{\partial}_A^2 = 0$ is the integrability condition for finding local bases of solutions of $\bar{\partial}_A s = 0$; these then define the local holomorphic trivialisation of the bundle, and transition functions between different patches are immediately holomorphic.

To get a closer link between holomorphic bundles and connections, we can try to fix the (arbitrary) metric by imposing an equation on the resulting connection. This is rather like uniformisation for Riemann surfaces: there one studies complex geometry by introducing a metric and doing Kähler geometry; this metric is arbitrary, but if we impose a constant scalar curvature condition (and fix the volume) it becomes unique and the study of the complex geometry and (constant scalar curvature) Kähler geometry are equivalent. Lengths of geodesics, for instance, are algebro-geometric invariants of the complex curve. In higher dimensions one can similarly consider the Hermitian-Einstein equations for a metric; this leads to Yau's celebrated theorem on Calabi–Yau manifolds; once only the cohomology class of a Kähler form is chosen, complex geometry is equivalent to Ricci-flat geometry.

So we now assume we have a Kähler form ω on M, and try to fix the mertic (and resulting connection A) using the Hermitian-Yang–Mills (HYM) equations

$$F_A^{2,0} = 0 = F_A^{0,2}, \tag{38.2}$$

$$F_A^{1,1} \wedge \omega^{n-1} = 0. \tag{38.3}$$

(We are restricting to $c_1(E) = 0$, but everything below applies to the more general case with $F_A^{1,1} \wedge \omega^{n-1} = C \operatorname{id} \omega^n$, $C = 2\pi i(c_1(E) \cdot \omega^{n-1}) / \int_M \omega^n$.)

One can also consider other similar equations on $F^{1,1}$ such as those of Conan Leung or the MMMS equation; we shall denote all of these by

$$\begin{gathered} F_A^{2,0} = 0 = F_A^{0,2}, \\ f(F_A^{1,1}) = 0, \end{gathered} \tag{38.4}$$

for some $2n$-form-valued function f.

Often, in particular for the $f(F) = F \wedge \omega^{n-1} - C \operatorname{id} \omega^n$ HYM equations, one can expect that for the *generic* holomorphic bundle there will exist a *unique* metric such that the associated connection satisfies the equation. (By generic we mean that if a single holomorphic bundle admits an HYM connection, then all but a collection of holomorphic subsets (of lower dimension) of the "space" of holomorphic bundles do; however no bundles at all may admit one, for instance if certain topological obstructions are not satisfied.)

Why might we expect this to be true ? Firstly, we can see it infinitesimally as follows. Linearising the equation $\bar{\partial}_A^2 = 0$ for a $(0,1)$-connection $A + a$ to define a holomorphic bundle about a holomorphic structure A, we get $\bar{\partial}_A a = 0$. But two $\bar{\partial}$-operators define isomorphic holomorphic structures if and only if they are conjugate: $\bar{\partial}_A \mapsto \phi \circ \bar{\partial}_A \circ \phi^{-1}$ by any general linear automorphism $\phi \in GL(E)$. ($GL(E)$ is the general linear gauge group of all smooth invertible automorphisms of E.) Infinitesimally, the Lie algebra $\Omega^0(\operatorname{End} E) \ni \psi$ acts by $\bar{\partial}_A \mapsto \bar{\partial}_A + \bar{\partial}_A \psi$. Thus, to first order, deformations of the isomorphism class of holomorphic structures are given by

$$\frac{\ker\{\bar{\partial}_A : \Omega^{0,1}(\operatorname{End} E) \to \Omega^{0,2}(\operatorname{End} E)\}}{\operatorname{im}\{\bar{\partial}_A : \Omega^{0}(\operatorname{End} E) \to \Omega^{0,1}(\operatorname{End} E)\}}, \tag{38.5}$$

i.e., by $H^{0,1}(\operatorname{End} E; \bar{\partial}_A)$ (meaning the sheaf, or Dolbeault, cohomology of $(\operatorname{End} E)$ endowed with the holomorphic structure $\bar{\partial}_A$).

Similarly given a HYM connection A, compatible with a fixed Hermitian metric h, first, order deformations satisfying the same are given by $A + i(a +$

a^*) (where $*$ is defined using h, and again $a \in \Omega^{0,1}(\operatorname{End} E)$) satisfying

$$\bar\partial_A a = 0 = \operatorname{Re}(\partial_A a \wedge \omega^{n-1}),$$
$$\text{i.e.,}\quad \bar\partial_A a = 0 = \operatorname{Re}(\bar\partial_A^* a).$$

We now have to divide by the smaller unitary gauge group $U(E)$ of automorphisms of E preserving h. Infinitesimally then, we divide by deformations $a = i\bar\partial_A\psi$ for ψ a real (self-adjoint) endomorphism of E; equivalently we can fix the gauge fix by insisting that $\operatorname{Im}\bar\partial_A^* a = 0$. The upshot is that deformations are given by

$$\text{(38.6)}\quad \ker\{\bar\partial_A : \Omega^{0,1}(\operatorname{End} E) \to \Omega^{0,2}(\operatorname{End} E)\} \cap \ker\{\bar\partial_A^* : \Omega^{0,1}(\operatorname{End} E) \to \Omega^{0}(\operatorname{End} E)\},$$

i.e., by $H^{0,1}(\operatorname{End} E; \bar\partial_A)$ again.

Thus the HYM equation provides, infinitesimally, a slice to the imaginary part of the $GL(E)$ gauge group action; instead of dividing by the whole complex Lie algebra action we can take the slice provided by HYM and divide by the real $U(E)$ part of the action.

This familiar linear trick of taking an orthogonal complement to a linear subspace (here the imaginary part of the infinitesimal gauge action), instead of the quotient by it (which is of course isomorphic), has a global nonlinear generalisation provided by the moment map. The space $\mathcal{A}^{1,1}$ of unitary connections A with $F_A^{0,2} = 0$ is formally an infinite dimensional complex manifold (since A is determined by its $(0,1)$ part, the tangent space $T_A\mathcal{A}^{1,1}$ is $\{a \in \Omega^{0,1}(\operatorname{End} E) : \bar\partial_A a = 0\}$ which is invariant under multiplication by i). It also has a symplectic form Ω, acting on tangent vectors (sections of $\Omega^{0,1}(\operatorname{End} E)$) by

$$\text{(38.7)}\qquad \Omega(a,b) = \int_M \operatorname{tr}(a \wedge b^* \wedge \omega^{n-1}),$$

making $\mathcal{A}^{1,1}$ Kähler. The gauge group $U(E)$ preserves Ω; in fact its action is *Hamiltonian*. That is, for each ψ in the Lie algebra $\mathfrak{u}(E)$ of *self-adjoint* endomorphisms of E, its infinitesimal action $A \mapsto A + i(\bar\partial_A\psi + \partial_A\psi)$ on $\mathcal{A}^{1,1}$ has a Hamiltonian (or momentum)

$$\text{(38.8)}\qquad m_\psi = -i\int_M \operatorname{tr}(\psi F_A \wedge \omega^{n-1}),$$

i.e., a real-valued function on $\mathcal{A}^{1,1}$ whose symplectic gradient (the vector field whose contraction with Ω is dm_ψ) is the vector field $i(\bar\partial_A\psi + \partial_A\psi)$ giving the infinitesimal action. All this means is that the vector field induced on $\mathcal{A}^{1,1}$ by the action of ψ is i times the gradient of m_ψ; $i \operatorname{grad} m_\psi = i(\bar\partial_A\psi + \partial_A\psi)$. That is, the one-parameter flow generated by ψ is motion on the symplectic manifold $\mathcal{A}^{1,1}$ according to Hamilton's equations with Hamiltonian m_ψ.

These Hamiltonians fit together to form a *moment map* m for the $U(E)$-action; that is a map m from $\mathcal{A}^{1,1}$ to the dual $\mathfrak{u}(E)^*$ (self-adjoint elements of $\Omega^{2n}(\operatorname{End} E)$) such that the pairing $\langle m, \psi\rangle = \int_M \operatorname{tr}(m\psi)$ is the Hamiltonian m_ψ (and m is equivariant with respect to the $U(E)$-action on $\mathcal{A}^{1,1}$ and the coadjoint action on $\mathfrak{u}(E)^*$).

From Eq. (38.8) we see that this moment map is

$$m : \mathcal{A}^{1,1} \to \Omega^{2n}(\operatorname{End}_{\mathbb{R}} E),$$
$$A \mapsto -iF_A \wedge \omega^{n-1},$$

whose zero set is precisely the HYM connections. (Recall that $-iF_A$ is self-adjoint.)

Now the action of $U(E)$ complexifies to an action of its complexification $GL(E)$ such that, at the infinitesimal level of Lie algebras $\operatorname{End} E = \mathfrak{u}(E)\otimes\mathbb{C}$, it is complex linear as a map to $T\mathcal{A}^{1,1}$. This gives the obvious $GL(E)$ action — conjugate a $\bar\partial$-operator, then take the corresponding connection compatible with this and the fixed metric h. This is *not* the same as conjugating the full connection, as $GL(E)$ does not preserve the metric; since conjugation of h with $GL(E)$ gives all metrics on E this is equivalent to fixing $\bar\partial_A$ and taking all metrics on E and their associated connections. (This is why we spoke of changing the metric to a HYM metric earlier; instead, and equivalently, we have fixed a metric and used gauge transformations that do not preserve it.)

Isomorphism classes of different holomorphic structures (or $\bar\partial$-operators) on the bundle E are parametrised by

$$\text{(38.9)}\qquad \mathcal{A}^{1,1}/GL(E);$$

the general theory of symplectic quotients (if $\mathcal{A}^{1,1}$ were finite dimensional) would identify this with

$$\text{(38.10)}\qquad m^{-1}(0)/U(E),$$

for the "generic" $\bar\partial$-operator. That is, the zero set of the moment map should provide a slice to the imaginary part of the action and be invariant under the real part; linearising this gives precisely the equality of Eqs. (38.5) and

(38.6). Moreover the "generic" $GL(E)$ orbit should intersect $m^{-1}(0)$ in a *unique* $U(E)$ orbit, where "generic" here means the *stable* orbits under the group action. Stability depends on the choice of the symplectic form and the moment map. We illustrate this equivalence of quotients with two examples.

The simple finite-dimensional example to have in mind is S^1 acting diagonally on $\mathbb{C}^m$ via $(z_1, \ldots, z_m) \mapsto (\lambda z_1, \ldots, \lambda z_m)$ for $\lambda \in S^1$. This complexifies to an action of $\mathbb{C}^\times$ with the same formula, and has moment map $\sum_i |z_i|^2 - a$ for any $a \in \mathbb{R}$. This provides a slice to the imaginary part of the $\mathbb{C}^\times$-action (the scaling part with orbits the radial straight lines) so that

$$\frac{m^{-1}(0)}{S^1} = \frac{S^{2m-1}}{S^1} \quad \text{is identified with} \quad \frac{\mathbb{C}^m \backslash \{0\}}{\mathbb{C}^\times} \cong \mathbb{P}^{m-1}$$

in the usual way. Here the unstable orbits, which we throw out, are just $\{0\}$ for $a > 0$ (and for $a \leq 0$ everything is unstable).

Secondly, what stability turns out to mean for bundles is the following.

DEFINITION 38.2.1. *A coherent sheaf E on a Kähler manifold M is (slope) stable if and only if for all proper subsheaves $0 \to F \to E$, we have $\mu(F) < \mu(E)$, where $\mu(E)$ is the* slope $c_1(E) \cdot \omega^{n-1}/r(E)$ *of E.*

Semistable bundles are ones for which the inequality is $\leq$ instead of $<$, and *polystable* bundles are direct sums of stable bundles of the same slope (and so are semistable). Stability is a generic ("Zariski open") condition in the sense mentioned briefly before, and *depends on a Kähler form.*

Using this infinite-dimensional picture and some very hard analysis, the analogue of the finite-dimensional results about symplectic quotients can be proved in this setting, namely Eqs. (38.9) and (38.10) can be identified for stable bundles.

THEOREM 38.2.2. *A holomorphic bundle E admits a compatible HYM connection if and only if it is polystable. This connection is unique up to isomorphism.*

This is a miraculous theorem, reducing an infinite-dimensional problem of solving a PDE to a finite-dimensional problem of linear algebra if M is projective algebraic (since then all holomorphic bundles can be described by algebraic, polynomial, data). It is similar in structure to Yau's theorem on Ricci-flat metrics, and both make physics very much easier, and both can be reinterpreted as moment map problems. In Sec. 38.4.2 we will try and

set up a similar structure for the similarly hard-to-solve sLag equations in the hope that they may be understood similarly.

One way to prove the theorem is to note that, in these moment map problems, the gradient flow of minus the norm square $-|m|^2$ of the moment map lies in the imaginary part of the $GL(E)$ orbit, and flows to a unique zero of the moment map (by convexity arguments) if the orbit is stable. Proving this in infinite dimensions involves a lot of analysis, but the formal picture suggests the right flow to find solutions (in this case it is the gradient flow of minus the Yang-Mills action $|m|^2 = \int_M |F|^2 +$ constant).

An important point is what happens to the flow in the strictly semistable, non-polystable, case. Then the limit does not lie in the same $GL(E)$ orbit — it defines a different holomorphic bundle (or sheaf, if singularities in the flow arise) lying in the orbit of a polystable bundle in the *closure* of the original bundle's orbit. This polystable bundle is the *Jordan-Hölder* decomposition of the original bundle; as a simple example suppose (as in the next section) that we have a semistable bundle E destabilised by a stable bundle E_1 of the same slope, with quotient another stable bundle E_2:

$$0 \to E_1 \to E \to E_2 \to 0.$$

Then this sequence need not split holomorphically (i.e., E need not be isomorphic to $E_1 \oplus E_2$; in fact extensions of the above form are naturally parametrised by a vector space of extensions $\mathrm{Ext}^1(E_2, E_1) = H^1(E_2^* \otimes E_1)$) so may not be polystable, but in the closure of its $GL(E)$ orbit is its polystable Jordan-Hölder decomposition $E_1 \oplus E_2$; it is to the direct sum HYM connection on this that the gradient flow converges.

So the gauge theory identifies semistable bundles whose Jordan-Hölder decomposition is the same; in fact so does the algebro-geometric theory of forming moduli of holomorphic bundles via geometric invariant theory.

Conjecturally, most of what we have said so far should apply to the Conan Leung/MMMS-type equations of the form $f(F_A) = 0$ (Eq. (38.4)) too. We replace the symplectic structure (Eq. (38.7)) by

$$\Omega(a, b) = \int_M \mathrm{tr}(a \wedge b^* \wedge f'(F_A)), \tag{38.11}$$

where f' denotes the derivative with respect to F_A only; for instance for $f = F_A \wedge \omega^{n-1}$, $f' = \omega^{n-1}$, yielding the standard symplectic structure Eq.

(38.7). Formally this gives the moment map

$$m(A) = f(F_A) \in \mathfrak{u}(E)^* \subset \Omega^{2n}(\operatorname{End} E)$$

for the gauge group action.

The case of $f = \exp(k\omega + F_A^{1,1})\,\mathrm{Td}\,(M)$ has been studied by Conan Leung and relates to Gieseker stability in algebraic geometry; more general equations have yet to be considered but it seems reasonable to conjecture that, for f sufficiently non-degenerate, Eq. (38.11) is non-degenerate at A (this is related to the ellipticity of the equation $f(F_A^{1,1}) = 0$ at A) and there should be a 1-1 correspondence as before. That is, modifying the notion of stability (Definition 38.2.1) by using the topological quantity

$$\mu(E) = \frac{\int_M \operatorname{tr} f(F_A)}{r(E)} \tag{38.12}$$

in place of the slope of a sheaf, (poly)stable holomorphic bundles should correspond to solutions of $f(F_A) = 0$, $F_A^{0,2} = 0$.

It would also be helpful to have algebro-geometric constructions of moduli of such bundles when ω is the first Chern class of a very ample line bundle $\mathcal{O}(1)$, and so Poincaré dual to a hyperplane section H. With $f = F_A \wedge \omega^{n-1}$, Eq. (38.12) is related to numbers of sections of a bundle on restriction to the generic curve that is the intersection H^{n-1} of $(n-1)$ hyperplanes in general position, and it is using just these sections that the moduli space of slope stable sheaves is constructed (on a complex surface, at least). Similarly for $f = \exp(k\omega + F_A)\,\mathrm{Td}\,(M)$, Eq. (38.12) measures numbers of sections of $E(k)$ on M, and this is what is used to construct the moduli space of Gieseker stable sheaves. Wherever Eq. (38.12) has such a Riemann-Roch-type interpretation it would be nice to find similar constructions.

In conclusion, a number of supersymmetric first-order physics equations come from moment maps, giving one a chance of solving them and getting a correspondence with purely holomorphic objects. This is one of the reasons for the relevance of purely holomorphic objects in physics, and it is to more of these that we turn now.

38.3. Derived categories

Since the appearance of Kontsevich's mirror symmetry conjecture (Sec. 37.7), derived categories have begun to infiltrate string theory, although perhaps not yet as much as they could have — their poorer cousin K-theory

gets more attention because it is more familiar, from index theory and supersymmetry, and easier to calculate with. But the derived category contains far more information than K-theory; in fact Kontsevich predicts the whole string theory can be recovered from it, as is more-or-less known for $K3$. The derived category is a very strong invariant of a complex variety — it determines the underlying variety in many cases (when the canonical bundle is ample or anti-ample; i.e., definitely not Calabi–Yau) and has very few autoequivalences (the right notion of automorphisms for categories). But, for instance, two varieties that differ by a flop have equivalent derived categories and also equivalent B-models (the variation of Hodge structure is unaffected, etc.). Other dualities in string theory also arise as autoequivalences of the derived category (Fourier–Mukai transforms) or equivalences between triangulated categories (Kontsevich's original mirror conjecture). Also the conjecture predicts that, in contrast with the (anti-)ample canonical bundle case, derived categories of sheaves on Calabi–Yau manifolds should admit many autoequivalences (mirror to the symplectomorphisms of the mirror symplectic manifold). This is easily checked at the level of K-theory, but the fact that this also holds at the level of the derived category is a significant check on Kontsevich's conjecture. Later (in Eq. (38.29)) we will give an example of an autoequivalence predicted by mirror symmetry that acts trivially on cohomology and K-theory but non-trivially on the derived category. This gives examples of two bundles Eqs. (38.15, 38.16), very different in the derived category, whose images in K-theory or cohomology are the same; this will be important to us in studying (special) Lagrangians in Sec. 38.4.

So derived categories are relevant in physics, and not just a piece of fancy mathematics for its own sake; here we will say a little about how to think of them. We are not actually going to say too much about their full theory (for that see the references), more how to think of their objects (rather than morphisms) in loose terms.

Roughly speaking, the objects of the bounded derived category of sheaves $D^b(M)$ on a complex manifold M should be thought of as some kind of set of all holomorphic bundles and coherent sheaves (Sec. 37.6.2) on M. However, it will also be important to be able to subtract sheaves from each other via a map between them, which in physics can be thought of in terms of annihilation of branes and anti-branes via a tachyon.

A simple example is given by the sheaf exact sequence

$$0 \to \mathcal{L}^{-1} \xrightarrow{s} \mathcal{O}_M \to \mathcal{O}_D \to 0. \tag{38.13}$$

Here $\mathcal{L}$ is the (sheaf of sections of) the line bundle $\mathcal{O}(D)$ defined by the divisor $D \subset M$, $\mathcal{L}^{-1}$ is its dual, and s is the section of $\mathcal{L}$ vanishing on D. This gives an isomorphism $\mathcal{L}^{-1}|_{M\setminus D} \xrightarrow{s} \mathcal{O}_{M\setminus D}$, but on D the section vanishes and so has image in $\mathcal{O}_M$ the functions vanishing on D; the quotient of this inclusion is precisely the functions on D, $\mathcal{O}_D$ (this is a skyscraper sheaf, zero away from D, called the structure sheaf of D).

In holomorphic K-theory, we identify $\mathcal{O}_D$ with $\mathcal{O}_M - \mathcal{L}^{-1}$ because of this sequence. The full definition is to take all sheaves, add in formal negatives, and identify the sum of the outer two terms in any short exact sequence with the central term. This is a refinement of the more familiar topological K-theory because the maps in exact sequences must be holomorphic; nonetheless there is an obvious map to topological K-theory (the charge map of D-branes), and to cohomology via the Chern character or Mukai vector.

If we pick another section s' of $\mathcal{L}$, with zero section D', this gives an identification of $\mathcal{O}_D$ with $\mathcal{O}_{D'}$ in holomorphic K-theory: both are $[\mathcal{O}_M] - [\mathcal{L}^{-1}]$. But in the derived category we keep the information of the arrow (in this case s); the sense in which we subtract $\mathcal{L}^{-1}$ from $\mathcal{O}_M$, via s, is kept track of, and we do not identify $\mathcal{O}_D$ with $\mathcal{O}_{D'}$ (in fact no two distinct coherent sheaves are ever identified in $D^b(M)$). Similarly the zero map $\mathcal{L}^{-1} \xrightarrow{0} \mathcal{O}_M$ is considered to be something different again (genuinely $\mathcal{O}_M - \mathcal{L}^{-1}$) and is not identified with any $\mathcal{O}_D$. While such a thing might not be considered in physics, as it is not stable or the stationary point of any action, it is important that in $D^b(M)$ all such unstable objects are kept — they may be the limit of stable objects, as for instance this one is as $s \to 0$.

Define a bounded *complex* of sheaves to be a sequence of sheaves and holomorphic maps (of $\mathcal{O}_M$-modules) between them

$$\{E^n \xrightarrow{d^{n+1}} \dots \xrightarrow{d^{m-2}} E^{m-2} \xrightarrow{d^{m-1}} E^{m-1} \xrightarrow{d^m} E^m\} \tag{38.14}$$

(where $n \leq m$ can be positive or negative) *such that* $d^k \circ d^{k-1} = 0$ for all k. Then the right statement to make is that the sequence Eq. (38.13) makes the complexes $\{\mathcal{L}^{-1} \xrightarrow{s} \mathcal{O}_M\}$ and $\{\mathcal{O}_D\}$ (thought of as a complex in only one degree) *quasi-isomorphic* — there is a map of complexes

$$\begin{array}{ccc} \{\mathcal{L}^{-1} & \xrightarrow{s} & \mathcal{O}_M\} \\ \downarrow & & \downarrow \\ \{0 & \longrightarrow & \mathcal{O}_D\} \end{array}$$

such that all maps commute, and induce an isomorphism on cohomology (the kth cohomology sheaf of the complex Eq. (38.14) is $\ker d^{k+1}/\operatorname{im} d^k$; thus the cohomology of Eq. (38.13) is zero — the complex is exact — and the cohomology of both of the above complexes is just $\mathcal{O}_D$ in degree 0). This *is not the same as saying the complexes have the same cohomology*, though clearly it implies it. If it were we could do without the arrows and use just sums and differences of cohomology sheaves; the strength of the derived category is in keeping track of the arrows.

In the derived category quasi-isomorphisms are inverted. Clearly, as in the above example, there need not be an actual inverse map for any quasi-isomorphism (any map of sheaves $\mathcal{O}_D \to \mathcal{O}_M$ has the property that when multiplied by a function vanishing on D, it vanishes; the only section of $\mathcal{O}_M$ with this property is 0, and so all maps $\mathcal{O}_D \to \mathcal{O}_M$ vanish). So we just add one in formally, as we think of two quasi-isomorphic objects as isomorphic. This is *not* a map of complexes, just a map in the derived category — an abstract arrow between two objects. There is then an obvious way of composing all such maps of complexes and arrows ϕ^{-1} (where ϕ is a quasi-isomorphism) such that $\phi\phi^{-1} = \mathrm{id}$, etc., and each arrow starts where the last one leaves off. This then defines the bounded derived category $D^b(M)$ of M, with a map to both K-theory and cohomology (since quasi-isomorphic complexes map to the same class: $\sum_k (-1)^k [E^k]$ for Eq. (38.14)).

As a continuation of our simple example, consider points (divisors) p_i on a Riemann surface, represented by their structure sheaves $\oplus_i \mathcal{O}_{p_i}$. In cohomology two such collections have the same class if and only if there is the same number n of points in the collection; they then define the class $n \in H^2$. The same is true of topological K-theory (which is isomorphic to cohomology in this example). In holomorphic K-theory, two points p_1, p_2 are identified if and only if they are linearly equivalent: if and only if there is a meromorphic function vanishing only at p_1 and with a pole only at p_2; if and only if they define the same line bundle $\mathcal{L} \cong \mathcal{O}(p_1) \cong \mathcal{O}(p_2)$. In this case the two sections s_i of $\mathcal{L}$ and the sequences

$$0 \to \mathcal{L}^{-1} \xrightarrow{s_i} \mathcal{O} \to \mathcal{O}_{p_i} \to 0$$

identify the two $\mathcal{O}_{p_i}$ with $\mathcal{O}_M - \mathcal{L}^{-1}$ in K-theory, and $s_1 \cdot s_2^{-1}$ is the required meromorphic function.

In the derived category, however, the $\mathcal{O}_{p_i}$ are complexes with different cohomology $\mathcal{O}_{p_i}$ so they remain distinct. Thus the derived category sees the points, holomorphic K-theory sees the linear equivalence class of the points, and cohomology sees only their total degree. (In higher dimensions cohomology detects the numerical equivalence class of a divisor, holomorphic K-theory sees its linear equivalence class, and the derived category sees the divisor itself.)

The structure of $D^b(M)$ is in general very complicated in dimensions greater than 1. This is both very useful for mathematics and homological algebra (for instance maps between objects in $D^b(M)$ turn out to give all their Ext groups and gives an easy way to understand these) and very difficult to deal with.

But Kontsevich's conjecture suggests that unfortunately these complications are necessary for physics. We really must consider all of these unstable objects as well as the D-branes. There are many instances (Fourier–Mukai transforms, etc.) where a duality of string theory may generically take D-branes to other D-branes, but in some cases takes them to more exotic unstable objects. Also, recent work of Douglas et al, mentioned in the references, has shown that away from the large Kähler limit these unstable complexes can become stable in a suitable sense, and so become part of the spectrum of D-branes.

As we have already mentioned, considering $D^b(M)$ also takes the Kähler dependence out of the B-model, gives us a way of identifying different HYM connections for different Kähler forms via the underlying holomorphic bundles, and keeps track of a bundle as it becomes unstable when the Kähler form "crosses a wall" (making one sub-bundle have larger slope than another — this has nothing to do with walls of the Kähler cone or degenerate Kähler forms) and is no longer represented by an HYM connection.

We give a simple example of such wall-crossing that will prove useful later on as mirror to a construction with sLags. Suppose we have two bundles (or

coherent sheaves) E_1 and E_2 with

$$\operatorname{Ext}^1(E_2, E_1) \cong \mathbb{C},$$

for simplicity. (This is $H^1(E_2^* \otimes E_1)$ in the case of E_2 being a bundle, and parametrises extensions (38.15).) We then form E from this unique non-trivial extension class

$$0 \to E_1 \to E \to E_2 \to 0. \tag{38.15}$$

Take a family of Kähler forms ω^t on M (*for fixed complex structure*) such that $\mu^t(E_1) - \mu^t(E_2)$ has the same sign as t. Suppose also that the E_i are stable for all $t \in (-\epsilon, \epsilon)$; then it is possible to show that E is stable for sufficiently small $t < 0$, while it is destabilised by E_1 for $t \geq 0$.

In the two-dimensional $K3$ case, Serre duality gives

$$\operatorname{Ext}^1(E_1, E_2) \cong \operatorname{Ext}^1(E_2, E_1)^* \cong \mathbb{C},$$

so that for $t > 0$ we can instead form an extension

$$0 \to E_2 \to E' \to E_1 \to 0, \tag{38.16}$$

to give a new bundle E which is stable for small $t > 0$, and has the same Mukai vector

$$v(E') = v(E_1) + v(E_2) = v(E) \quad (:= \operatorname{ch}(E)\sqrt{\operatorname{Td} M}).$$

In the threefold case, Serre duality gives

$$\operatorname{Ext}^2(E_1, E_2) \cong \operatorname{Ext}^1(E_2, E_1)^* \cong \mathbb{C} \tag{38.17}$$

instead, and so no extension (38.16). In both dimensions, at $t = 0$ the polystable bundle

$$E_1 \oplus E_2 \tag{38.18}$$

admits an HYM connection, while E and E' do not, and the limit of the HYM connections on E for $t < 0$ (E' for $t > 0$) is this direct sum HYM connection on the different holomorphic bundle $E_1 \oplus E_2$.

Therefore, following the HYM connection as t (and so the Kähler form ω^t) varies, the underlying bundle has exactly three different holomorphic structures, undergoing a wall crossing at $t = 0$, moving into a different gauge group orbit, but one which intersects the *closure* of the gauge group orbits of the bundles E (Eq. (38.15)) and E' (Eq. (38.16)):

(38.19)

	3-fold case	**$K3$ case**
• $t < 0$	E HYM $\forall t$	E HYM $\forall t$
• $t = 0$	$E_1 \oplus E_2$ HYM	$E_1 \oplus E_2$ HYM
• $t > 0$	–	E' HYM $\forall t$.

What in fact we do get on a threefold from the one-dimensional Ext^2 (Eq. (38.17)) is a complex E' in the derived category $D^b(M)$ fitting into an exact sequence of complexes

$$0 \to E_2 \to E' \to E_1[-1] \to 0, \tag{38.20}$$

where $E_1[-1]$ is E_1 shifted in degree by one place to the right as a complex. This has Mukai vector

$$v(F) = v(E_2) - v(E_1). \tag{38.21}$$

Thus as we pass through $t = 0$ the natural object on the other side has the wrong homology class in odd dimensions (but the right one in even dimensions) and so an element of the appropriate moduli space disappears.

Changing the semistable E into the polystable $E_1 \oplus E_2$ at $t = 0$, or the unstable E into the sum of stable bundles $E_1 \oplus E_2$ of different degree for $t > 0$, are examples of the Jordan–Hölder filtration, or decomposition, of an unstable sheaf into stable parts. Its relevance to us is that the gradient flow of the Yang–Mills action mentioned in the last section tends, in an appropriate sense (after some rescaling) to the direct sum of the HYM connections on this decomposition if the bundle E is not stable. If $t = 0$ (the semistable case) this is itself HYM, and when $t > 0$ (the unstable case) it is *not* HYM, but still a stationary point of the Yang–Mills action (satisfying the second-order Yang–Mills equation $d_A^* F_A = 0$).

38.4. Lagrangians

D-branes on the A-model side are given by special Lagrangians, but this requires a complex structure to be fixed on this symplectic side. The analogue of considering holomorphic bundles, rather than HYM connections, is to consider Lagrangians up to Hamiltonian deformations, thus using only

the symplectic geometry. This is the data that goes into the Fukaya category (roughly speaking this is built from Lagrangians modulo Hamiltonian deformations, with flat unitary connections on them, rather like, but more complicated than, the way $D^b(M)$ is built up from sheaves on M). In this section we will explain why this should generically (again modulo issues of stability) correspond to considering sLags. But first we need to recall some definitions from Sec. 37.3.1.

A *Lagrangian* submanifold L of a symplectic manifold (W^{2n}, ω) is a n-dimensional submanifold such that the restriction $\omega|_L \equiv 0$ is identically zero. In the presence of a complex structure on W making it a Calabi–Yau manifold, we may pick a nowhere zero holomorphic n-form Ω, and note (by a local calculation in differential geometry) that

$$\Omega|_L = e^{i\theta} \mathrm{vol}\,, \tag{38.22}$$

where vol is the Riemannian volume form on L induced by Yau's Ricci-flat metric on W, and θ is some (multiple-valued) real function on L. vol provides an orientation for L, and reversing its sign alters the phase θ by π. A *special Lagrangian* (sLag) is a Lagrangian with *constant* phase θ, which may be set to zero, without loss of generality, by replacing Ω by $e^{-i\theta}\Omega$.

The winding class $\pi_1(L) \to \pi_1(S^1)$ of the phase map

$$L \xrightarrow{e^{i\theta}} S^1,$$

is called the Maslov class of the Lagrangian (and is clearly zero for a sLag). *We only consider Lagrangians of vanishing Maslov class*, for which $e^{i\theta}$ can then be lifted to a global, single-valued function $\theta : L \to \mathbb{R}$. Such a lift is called a *grading* of L and is equivalent, but perhaps simpler, than the more general definition of grading we gave in Sec. 37.7.1 using only the symplectic structure. There are therefore a $\mathbb{Z}$ of gradings $\theta \mapsto \theta + 2n\pi$ (corresponding under mirror symmetry to shifts $[2n]$ of a complex in the derived category). Adding π to the function θ corresponds to reversing L's orientation, and shifting by $[1]$ on the mirror. (Grading and Maslov class can be defined purely topologically without reference to a complex structure, of course.) From now on *we will only consider graded Lagrangians*, i.e., when talking about a Lagrangian L we will have implicitly chosen a pair (L, θ), where θ is a lift of $e^{i\theta}$. Thus we will have a natural orientation for L too.

Similarly we can define a kind of average phase $\phi = \phi(L)$ of a submanifold (or even just homology class) $L \subset W$ by

$$\int_L \Omega = A\, e^{i\phi(L)},$$

for some real number A; we then use $\mathrm{Re}\,(e^{-i\phi(L)}\Omega|_L)$ to orient L. Reversing the sign of A alters the phase by π and reverses the orientation. Again for a *graded* Lagrangian $L = (L, \theta)$, $\phi(L)$ is canonically a well-defined real number (rather than S^1-valued). Shifting the grading $[2n] : \theta \mapsto \theta + 2n\pi$ gives a similar shift to the phase $\phi(L)$.

The terminology comes from the fact that if there is a submanifold in the same homology class as L that is sLag, then the constant phase θ of that sLag is $\phi(L)$. Slope, which we define as

$$\mu(L) := \tan(\phi(L)) = \int_L \mathrm{Im}\,\Omega \Big/ \int_L \mathrm{Re}\,\Omega,$$

is independent of grading, monotonic in ϕ in the range $(-\pi/2, \pi/2)$, and invariant under change of orientation $\phi \mapsto \phi \pm \pi$. This agrees with the slope of a straight line sLag in the case of T^2.

First-order deformations of a Lagrangian are given by normal vector fields v preserving the $\omega|_L \equiv 0$ condition, i.e., such that the Lie derivative $\mathcal{L}_v\omega = d(v \lrcorner \omega)$ is zero. Thus, under the natural isomorphism $v \mapsto v \lrcorner \omega$ between the normal bundle to L and T^*L, deformations correspond to closed 1-forms $Z^1(L)$: an infinite-dimensional space.

Hamiltonian deformations of L are given by those closed one-forms that are exact, i.e., by normal vector fields $dh \lrcorner \omega^{-1}$ for some "Hamiltonian" function h on L. For instance, two curves on a Riemann surface are Hamiltonian deformations of one another if and only if the area between them is zero (when measured with sign). In general a global deformation of a Lagrangian is Hamiltonian if and only if the *flux* of each loop $\gamma \in H_1(L)$ is zero; that is, if the integral of ω over the two-chain swept out by γ under the deformation is zero.

Thus, to first order, deformations of a Lagrangian up to Hamiltonian deformations are given by

$$Z^1(L)/d\Omega^0(L) = \ker d/\mathrm{im}\, d = H^1(L;\mathbb{R}),$$

compare Eq. (38.5).

On the other hand those closed one-forms σ whose corresponding deformation preserves the special condition (which is $\mathrm{Im}\,\Omega|_L \equiv 0$ for Ω rotated such that L has phase 0) are just those for which

$$0 = \mathcal{L}_{\sigma \lrcorner \omega^{-1}} \mathrm{Im}\,\Omega = d((\sigma \lrcorner \omega^{-1}) \lrcorner \mathrm{Im}\,\Omega) = d(*\sigma),$$

for L sLag. Thus deformations of sLags are given by

$$\ker d \cap \ker d^* = H^1(L;\mathbb{R}),$$

compare Eq. (38.6).

So we see that we are in a situation analogous to that for connections; just as the HYM condition provided a slice to the imaginary part of the complex gauge group action (which infinitesimally means that deformations of HYM connections are the same as deformations of holomorphic bundles up to complex gauge transformations), the special condition provides, infinitesimally at least, a slice to the Hamiltonian deformations of a sLag. No such first-order Hamiltonian deformation of a sLag is sLag, and deformations of a sLag agree with deformations of the underlying Lagrangian modulo Hamiltonian deformations. Thus we would like to find a moment map formalism to globalise this, with zeros the sLags, unique in a Hamiltonian deformation orbit of a generic Lagrangian.

However, the space of Lagrangian submanifolds is not complex, and we do not yet have a real group action whose complexification has imaginary part the Hamiltonian deformations. So first we introduce flat $U(1)$ connections on our Lagrangians. Deformations of flat $U(1)$ connections are given by closed one-forms, so when combined with the Lagrangians we find the formal tangent space to the space $\mathcal{Z}$ of all such is naturally complex with a complex structure J described by:

$$T_L\mathcal{Z} \cong Z^1(L) \oplus Z^1(L), \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{38.23}$$

We then extend the table at the start of the chapter as follows; the rest of this section will be taken up with explaining all of its entries.

While the bottom half of the table will hopefully work in all dimensions, the top half applies only in dimension 3, and is included to make more contact with Sec. 37.5. It need not be exact in any physical sense except in some appropriate limit, but can be thought of as a list of nice analogies.

Complex n-fold M	Symplectic $2n$-manifold W
Connections A on a fixed C^∞ complex bundle E	Submanifolds L with a $U(1)$ connection on $\mathbb{C} \times L$
$v(E) = \mathrm{ch}(E)\sqrt{\mathrm{Td}\, M} \in H^{\mathrm{ev}}$	$[L] \in H^3$
$CS_{\mathbb{C}}(A = A_0 + a) =$ $\frac{1}{4\pi^2}\int_M \mathrm{tr}\left(\bar\partial_{A_0} a \wedge a + \frac{2}{3}a^{\wedge 3}\right) \wedge \Omega$	$f_{\mathbb{C}}(A, L) = \int_{L_0}^{L}(F+\omega)^2$ $= \int_{L_0}^{L}(F^2+\omega^2) + 2\int_{L_0}^{L}\omega \wedge F$
Critical points: $F_A^{0,2} = 0$ holomorphic bundles	Critical points: $\omega\vert_L = 0,\ F_A = 0$ Lagrangians + flat line bundles
Holomorphic Casson invariant	Counting sLags (Joyce)
Gauge group	$U(1)$ gauge group on L
Complexified gauge group	Hamiltonian deformations
Moment map $F_A \wedge \omega^{n-1}$	Moment map $\mathrm{Im}\,\Omega\vert_L$
Stability, slope $\mu =$ $\frac{1}{\mathrm{rk}\, E}\int_M \mathrm{tr} F_A \wedge \omega^{n-1}$	Stability, slope $\mu =$ $\frac{1}{\int_L \mathrm{Re}\,\Omega}\int_L \mathrm{Im}\,\Omega$

One of the reasons for the disclaimer is that the picture is a deliberate simplification of mirror symmetry, and avoids the derived category and Fukaya category. These are undoubtedly necessary in general — the mirror of a Lagrangian can in examples be a non-trivial complex of sheaves rather than a single sheaf, and vice-versa — Kontsevich was not just showing off by introducing these constructions into mirror symmetry. Fourier–Mukai transforms show that in some sense the derived category is the smallest structure that we can use that contains all stable holomorphic vector bundles (and so, using extensions and resolutions, all coherent sheaves). The table will not hold in many cases, but the analogy does. Also, as before, we should consider the MMMS equations as well as HYM, but since we have seen that these too can come from a moment map the analogy still holds.

Another complication that ought to be mentioned is the issue of obstructions. While Lagrangians (and sLags) L have unobstructed deformations governed by $H^1(L)$, holomorphic bundles (and HYM gauge fields) E do not. In Kontsevich's proposal the deformation space $\mathrm{Ext}^1(E,E)$ (this is just $H^1(\mathrm{End}\, E)$ if E is a vector bundle) is related not to $H^1(L)$, but to the Floer cohomology $HF^1(L,L)$. This is approximated by $H^1(L)$, but has corrections due to holomorphic disks (disk instantons) with boundary on L. The interpretation is that we should only consider Lagrangians whose Floer cohomology can be defined as possibly mirror to objects in the derived category (in physics terms this means those for which the gradient of the superpotential vanishes), and these are obstructed. Again, we can ignore this technicality in our analogies.

Firstly, then, we are considering a bundle E on a threefold M, with Mukai vector $v(E) = \mathrm{ch}(E)\sqrt{\mathrm{Td}\, M} \in H^{\mathrm{ev}}(M)$ mirror to a class $[L] \in H^3(W)$ (E need *not* be holomorphic, nor L Lagrangian, at this stage). Considering the space $\mathcal{A}$ of all $(0,1)$-connections on E as in Sec. 38.2 (with a Hermitian metric on E, this can be identified with the space of unitary connections), Witten's holomorphic Chern–Simons functional

$$CS_{\mathbb{C}}(A = A_0 + a) = \frac{1}{4\pi^2}\int_M \mathrm{tr}\left(\bar\partial_{A_0} a \wedge a + \frac{2}{3} a \wedge a \wedge a\right) \wedge \Omega$$

on $\mathcal{A}$ is invariant under all $GL(E)$ gauge transformations connected to the identity, and so descends to a multi-valued function on $\mathcal{A}/GL(E)$. Its critical points are those connections with $\bar\partial^2 = 0$, i.e., those that define a holomorphic structure on E. Thus we expect the space of holomorphic bundles of a fixed topological type on M to have virtual dimension 0 (the critical points of a functional, i.e., the zeros of a one-form (its exterior derivative) gives as many equations as there are unknowns) and we might hope to count them using virtual cycles as in Sec. 26.1 — this can be proved to work, defining a holomorphic Casson invariant. (In fact what the holomorphic Casson invariant counts is *stable* bundles, i.e., HYM connections, plus some sheaves in the compactification of the moduli space[1].)

[1]A nice example of the holomorphic Casson invariant is given by the intersection X of a quadric and a quartic in $\mathbb{P}^5$; the quadric is the Grassmannian of two-planes in $\mathbb{C}^4$ and so has two tautological two-plane bundles on it; the universal sub- and quotient- bundles A and B. Then $A|_X$ and $B^*|_X$ lie in the same moduli space of stable bundles on X, and are the only semistable sheaves of the same Chern classes; thus the appropriate holomorphic

This functional can be written more topologically in three different ways, which will be more clearly analogous to the A-model functionals below. They all reflect the fact that $\bar{\partial}$ of the three-form $\bar{\partial}_{A_0}a \wedge a + \frac{2}{3}a \wedge a \wedge a$ is the four-form $F \wedge F$, where $F = F^{0,2}_{A_0+a}$ is the (0,2)-part of the curvature of $A_0 + a$. The differential form $\frac{1}{4\pi^2}\operatorname{tr} F \wedge F$ is the Chern–Weil representative of minus the first Pontrjagin characteristic class of a bundle, $-p_1 = 2c_2 - c_1^2$.

Let π denote the projection $X \times [0,1] \to X$, and let $\mathbb{A}$ be any connection on π^*E restricting to A_0 on $X \times \{0\}$ and A on $X \times \{1\}$. Then the first alternative formula for Witten's holomorphic Chern–Simons functional is

$$CS_{\mathbb{C}}(A) = \frac{1}{4\pi^2}\int_{X\times[0,1]} \operatorname{tr}(F_{\mathbb{A}} \wedge F_{\mathbb{A}}) \wedge \pi^*\Omega.$$

Secondly, and more in keeping with the holomorphic nature of $CS_{\mathbb{C}}$, suppose that the Calabi–Yau threefold X is a smooth anti-canonical divisor in a fourfold Y; that is, the zero set of a holomorphic section $\eta \in H^0(K_Y^{-1})$. We think of η^{-1} as a meromorphic complex volume (4,0)-form on Y with first-order poles along X. Suppose that (E, A) extends to a bundle with (non-integrable) connection $(\mathbb{E}, \mathbb{A})$ on Y. Then, up to an additive constant,

$$CS_{\mathbb{C}}(A) = \frac{1}{4\pi^2}\int_Y \operatorname{tr}(F_{\mathbb{A}} \wedge F_{\mathbb{A}}) \wedge \eta^{-1}.$$

Finally, there is a third formula in the special case that $A = A_0 + a$ is integrable, admits a holomorphic section s, and has $c_1 = 0$, so that $-p_1(E)$ is now just $2c_2$, twice the Euler class of E, and so is represented by Poincaré duality by twice the zero set of s. That is, fix once and for all a two-cycle $Z_0 \subset X$ representing $c_2(E)$. Then the zero set Z of s is homologous to Z_0, so that $Z - Z_0$ is the boundary $\partial\Delta$ of a three-chain Δ, and (up to an additive constant),

$$CS_{\mathbb{C}}(A) = \frac{1}{2\pi^2}\int_\Delta \Omega.$$

Since we are interested in more than just holomorphic bundles, but all (complexes of) coherent sheaves, it will come as no surprise that there is also a functional (on the space of two-cycles in a fixed homology class in X) with critical points that are holomorphic curves in X (integrate Ω over a three-chain bounding the two-cycle minus a fixed cycle) and a functional (on the space of four-cycles with $U(1)$ connections A over them) with critical

Casson invariant is 2. It would be nice to find two distinguished isolated sLag cycles on the mirror, which can be constructed by Batyrev's method.

points that are holomorphic surfaces in X with integrable $F_A^{0,2} = 0$ connections on them (integrate $\Omega \wedge F_{\mathbb{A}}$ over a bounding five-cycle). The value of these functionals at such a holomorphic object (critical point) can be easily related to the holomorphic Chern–Simons functional of a resolution by holomorphic bundles of the associated coherent sheaf ($CS_{\mathbb{C}}$ is additive on exact sequences, so can be defined on complexes). This comes from using transgression formulae like those above for the Chern classes of the bundles (wedged with Ω); the point is that the homology class of the curve or surface can be recovered from the Chern classes of the bundles in a resolution.

On the symplectic side there is also a functional on submanifolds L, whose critical points constitute the space of *Lagrangian submanifolds*; it is

$$f(L) = \int_{L_0}^{L} \omega^2.$$

Here we have fixed a submanifold L_0 in the same homology class, and integrated over a four-chain bounding $L - L_0$. This, in good analogy to the gauge invariance of $CS_{\mathbb{C}}$, is invariant under Hamiltonian deformations of L and choices of the four-cycle, up to some periods from changing the homology class of the four-cycle.

However, as mentioned before, we should complexify by adding in $U(1)$ connections A (*not* necessarily flat at this stage), then consider the functional

$$f_{\mathbb{C}}(A, L) = \int_{L_0}^{L} (F_{\mathbb{A}} + \omega)^2 = \int_{L_0}^{L} (F_{\mathbb{A}}^2 + \omega^2 + 2\omega \wedge F_{\mathbb{A}}).$$

Here we have chosen a connection $\mathbb{A}$ on the four-chain bounding L_0, L, with restrictions A_0, A to L_0, L respectively, for some fixed A_0. (Recall that the curvature F is imaginary, so that $f_{\mathbb{C}}$ is complex-valued.) The critical points of this functional are the set of Lagrangians with flat connections on them, and the functional is formally holomorphic with respect to Eq. (38.23) (suitably extended to submanifolds close to Lagrangians). The functional is also invariant under $U(1)$ gauge transformations of A (*not* $GL(1)$ gauge transformations as on the B-model side above) and Hamiltonian deformations of L (if A is carried along with L appropriately). So this ties in three well-known moduli problems of virtual dimension 0 (i.e., with deformation theories whose Euler characteristic vanishes) — flat bundles on three-manifolds, holomorphic bundles on Calabi–Yau threefolds, and Lagrangians up to Hamiltonian deformation (or sLags) in symplectic sixfolds.

So as mirror to the holomorphic Casson invariant one would like to count Lagrangians up to Hamiltonian deformations, or sLags, plus flat line bundles on them, and this is what some recent work of Dominic Joyce's has begun (in the rigid case of L being a homology sphere).

The functional $f_{\mathbb{C}}$ is of course the real Chern–Simons functional $\int_{L_0}^{L} F_{\mathbb{A}}^2 = \int_L a \wedge da$ of $A = A_0 + a$, plus some other terms. If we do not vary (L, A) outside the set of Lagrangian submanifolds, these extra terms remain constant and mirror symmetry relates to the real and holomorphic Chern–Simons functionals in Sec. 37.9. This explains how a real and a holomorphic functional can be mirror. If we extend the real functional to $f_{\mathbb{C}}$ and allow additional deformations in the directions that are J times the deformations of the connection A (Eq. (38.23)), i.e., in the direction of non-Lagrangian submanifolds L, the functionals both take complex values and remain equal.

In fact writing, in a neighbourhood of L, $\omega = dB$ (or, more invariantly, assume $\omega/2\pi$ is integral and pick a connection iB on a line bundle with curvature $i\omega$), we may rewrite

$$f_{\mathbb{C}}(A, L) = \int_L (B + iA) \wedge d(B + iA) = \int_L C dC$$

for the "complexified connection" $C = B+iA$, (i.e., a $\mathbb{C}^{\times}$-connection, instead of a $U(1)$-connection). Now, by the methods of Sec. 37.9 (or of the paper [**176**]), one finds that the holomorphic Chern–Simons functional is equal to $f_{\mathbb{C}}$, for smooth sLag T^n-fibrations with L a section.

Having hopefully motivated this whole set-up, we now restrict to the above critical points — integrable $\bar{\partial}$-operators ($\bar{\partial}^2 = 0$) on the complex side, and Lagrangians with flat $U(1)$ connections on them on the symplectic side; we forget all about the functionals, and work in arbitrary dimensions.

Fixing a metric on the bundle E in the B-model, we then have two natural real groups acting — the unitary gauge group $U(E)$ of E, and the unitary gauge group $C^{\infty}(L; U(1))$ of smooth $U(1)$-valued functions on L. We saw in Sec. 38.2 that the natural complexification of $U(E)$'s action, with respect to the obvious complex structure on $\mathcal{A}$, is the complex gauge group $GL(E)$. But with respect to the complex structure Eq. (38.23) it is easy to see that the complex group action on (L, A) is a combination of Hamiltonian deformations of L and unitary gauge transformations of A. That is, the real Lie algebra $\Omega^0(L;\mathbb{R})$ acts only on A as $A \mapsto A + idh$, while

the imaginary part

$$i\Omega^0(L;\mathbb{R}) \subset \Omega^0(L;\mathbb{C})$$

of its complexification deforms L by the Hamiltonian h, and carries A along with it.

We are left with explaining the moment maps and stability conditions on the RHS of the table. Put a symplectic structure on the space of Lagrangians plus flat $U(1)$-connections on them using the metric

$$\langle a, b \rangle = \int_L a \wedge ((b \lrcorner \omega^{-1}) \lrcorner \operatorname{Im} \Omega)$$

on $\Omega^1(L;\mathbb{R})$ and the complex structure Eq. (38.23). For Lagrangians whose phase function is bounded inside $(-\pi/2, \pi/2)$ this is non-degenerate. We leave aside questions of integrability of these structures — there are many problems with this infinite-dimensional set-up anyway such as ignoring disk instantons, and the fact that the group $C^{\infty}(L; U(1))$ has not been invariantly identified on different Lagrangians. These can be overcome under certain conditions discussed in the references, so we shall blithely continue, not letting them interfere with our general picture.

Then a formal calculation shows that the $U(1)$ gauge group action has a moment map (in the dual $\Omega^n(L;\mathbb{R})$ of the Lie algebra $\Omega^0(L;\mathbb{R})$) indeed given by

$$(L, A) \mapsto \operatorname{Im} \Omega|_L \in \Omega^n(L;\mathbb{R}).$$

Thus we have the ingredients for a Hitchin–Kobayashi conjecture, that gradient flow of minus the norm square of the moment map $\operatorname{Im} \Omega|_L$, which gives a Hamiltonian deformation of L, should converge to a zero of the moment map — a special Lagrangian — if and only if some stability condition is satisfied. Firstly note that of course we should satisfy the topological constraint $\int_L \operatorname{Im} \Omega = 0$ (i.e., $\phi(L) = 0 \bmod 2\pi\mathbb{Z}$) for there to be solutions, and this can be done by rotating Ω. Secondly, picking the right metric on the Lie algebra we can choose the gradient flow to be the mean curvature flow of the Lagrangian, which is the flow with Hamiltonian the phase function θ (at least if the variation of θ is not too large over L, which is all we might expect the theory to work for anyway; as $\sup |\theta|$ decreases under mean curvature flow this is not a problem). Mean curvature flow is a much studied flow to minimise the volume of submanifolds, but its convergence properties are not at all well understood. With the stability condition on Lagrangians motivated by the next section, we will be led to a conjecture about how it behaves and what it converges to for Lagrangians in a Ricci-flat Calabi–Yau manifold.

38.4.1. Monodromy. To describe an example studied by Joyce in terms of Lagrangian geometry (up to Hamiltonian deformations), instead of his sLag description, we need the notion of Lagrangian surgery, or *Lagrangian connect sum*. Given two Lagrangian submanifolds L_1, L_2 intersecting transversely in a finite number of points (which we may assume after a Hamiltonian deformation), there is a Lagrangian

$$L_1\#L_2$$

constructed by gluing in a fixed local model in Darboux charts around the intersection points. Firstly we are interested in the case where the intersection is just one point p; then *the class of $L_1\#L_2$ up to Hamiltonian deformation is uniquely defined.* Topologically the construction gives the usual connect sum — removing a small ball B^n neighbourhood of p from each L_i and gluing together the resulting S^{n-1} boundaries, *using the relative orientation of the tangent spaces T_pL_i* given by the canonical orientation of

$$T_pW \cong T_pL_1 \oplus T_pL_2,$$

which are induced by the symplectic form. But symplectically the construction is much more delicate; for instance in even dimensions, where reversing the order of L_1, L_2 does not change the topological connect sum,

$L_1\#L_2$ *is not Hamiltonian deformation equivalent to* $L_2\#L_1$.

Choosing orientations such that $L_1 \cdot L_2 = +1$, we have, at the level of homology, in n complex dimensions,

$$[L_1\#L_2] = [L_1] + [L_2],$$

$$\pm[L_2\#L_1] = [L_2] + (-1)^n[L_1]. \tag{38.24}$$

The $\pm$ sign arises because there is no natural orientation on $L_2\#L_1$ coming from those on the L_i. What we are interested in, however, is *graded connect sums* — here the connect sum is given a grading *restricting on the L_is to their given grading*. Given a fixed grading on L_1 there will therefore be a unique grading on L_2 such that $L_1\#L_2$ exists as a graded connect sum (i.e., such that the corresponding Lagrangian can be compatibly graded), and, in

the case of multiple intersection points of the L_i, there may be no graded connect sum at all. From now on $\#$ will mean graded connect sum of graded Lagrangians. Since a grading gives a Lagrangian an orientation, there is no $\pm$ in Eq. (38.24).

Given these preliminaries, we can now write Joyce's examples in the following suggestive manner, to fit our general picture. We have a family of Calabi–Yau threefolds W^t, $t \in (-\epsilon, \epsilon)$ such that the complex structure (and so the holomorphic form Ω^t) on W^t varies with t *but the Kähler form remains fixed*; thus each W^t can be thought of as the same symplectic manifold. We also have a family of sLag homology three-spheres L_1^t, $L_2^t \subset W^t$ such that L_1^t and L_t^2 intersect at a point, and the phases of the L_i^ts vary as follows (after rotating Ω^t suitably):

$$\phi(L_1^t) = t, \qquad \phi(L_2^t) = 0, \ \forall t.$$

Also, for $t < 0$ there is a sLag L^t (of some phase $\phi^t \in (t, 0)$) in the homology class $[L^t] = [L_1^t] + [L_2^t]$, such that as $t \uparrow 0$, this degenerates to the singular union $L^0 = L_1^0 \cup L_2^0$ (which is sLag as L_1^0 and L_2^0 have the same phase at $t = 0$) and then disappears for $t > 0$. We claim that L_i^t is in the *same* Hamiltonian deformation class L_i for all $t > 0$, and that L^t is in fact

$$L^t \approx L_1\#L_2,$$

where we use $\approx$ to denote equality up to Hamiltonian deformation, and we have chosen appropriate gradings of the L_is. Thus, for $t < 0$, as Ω^t varies, the symplectic structure and Hamiltonian deformation classes of the Lagrangians never vary; we are thinking of this as mirror to a bundle like E (Eq. (38.15)) above, with fixed holomorphic structure but varying HYM connection as the mirror Kähler form varies.

As we reach $t = 0$, L^t becomes the singular union $L^0 = L_1^0 \cup L_2^0$, which *is not in the same Hamiltonian deformation class as $L_1\#L_2$*, but *in the closure of the Hamiltonian deformation orbit.* This should be thought of as mirror to the polystable bundle $E_1 \oplus E_2$ of Eq. (38.18) in the closure of the gauge group orbit of E, which has a reducible HYM connection at $t = 0$ (the sum of the HYM connections of the same slope on the E_i) just as $L_1^0 \cup L_2^0$ is the sLag union of two sLags of the same slope or phase.

Most importantly, where L^t exists as a smooth sLag ($t < 0$) we have the slope (and phase) inequality

$$\mu_1^t < \mu_2^t, \qquad (\phi_1^t < \phi_2^t); \tag{38.25}$$

at $t = 0$, L^t becomes the singular union of L_1^t and L_2^t, with $\mu_1^t = \mu_2^t$ ($\phi_1^t = \phi_2^t$); then for $\mu_1^t > \mu_2^t$ ($\phi_1^t > \phi_2^t$) there is no sLag in L's homology class, *even though there is a Lagrangian.* Though we have been using the slope μ in order to strengthen the analogy with the mirror, bundle, situation, from now on we shall use only the phase (lifted to $\mathbb{R}$ using the grading). While both are monotonic in the other for small phase (as $\tan\phi = \mu$), slope does not see orientation as phase does; reversing orientation adds $\pm\pi$ to the phase but leaves μ unchanged. This is related to the fact that we should really be working with *complexes* of bundles on the mirror side (the bundle analogy is too narrow) and changing orientation has no mirror analogue in terms of only stable bundles; it corresponds to shifting (complexes of) bundles by one place in the derived category. Also, we are not claiming that slope should be exactly mirror to the slopes of bundles; the relationship is probably more complicated, but as we saw in Sec. 38.2, it is possible to find set-ups that involve any number of perturbations of the slope parameter in stability definitions, so again we shall simply use the analogy.

In the analogous $K3$ case, however, a hyper-Kähler rotation (to make sLags algebraic curves) and some algebro-geometric deformation theory show that such an obstruction to deforming a sLag does not occur, and that the corresponding picture for L^t (which is sLag, remember) is the following.

Hamiltonian def. class of sLag L^t	**3-fold case**	**$K3$ case**
• $t < 0$	Constant $\forall t$ $L^t \approx L_1 \# L_2$ $(*)$	Constant $\forall t$ $L^t \approx L_1 \# L_2$ $(*)$
• $t = 0$	In closure of orbit $(*)$ $L^t \approx L_1 \cup L_2$	In closure of orbits $(*, \dagger)$ $L^t \approx L_1 \cup L_2$
• $t > 0$	–	Constant $\forall t$ $L^t \approx L_2 \# L_1$ $(\dagger)$

Notice the different order of the connect sum in the last entry, and compare (38.19). Wherever we change Hamiltonian deformation class, the 2 corresponding classes have intersecting closure.

To see why the Hamiltonian deformation class of L^t changes from $L_1 \# L_2$ to $L_2 \# L_1$ in the $K3$ case, we can sometimes use monodromy around a point in the complex structure moduli space where the cycle L_1 collapses completely (as its volume $\left|\int_{L_1}\Omega\right| \to 0$). Suppose then that L_1 is the Lagrangian S^2 vanishing cycle of an ordinary double point degeneration of (M, Ω), that is, lies in a Kähler family of manifolds M_u, $u \in D^2$, with smooth fibers except that M_0 has a singularity modelled on

$$\sum_{i=0}^{2} x_i^2 = u,$$

at $u = 0$. Then away from $u = 0$ we have a symplectic fiber bundle (while the complex structure on the fibers M_u changes with u, the symplectic structure remains fixed), whose monodromy therefore lies in the symplectomorphism group $\mathrm{Aut}(M, \omega)$ of the generic fiber. This symplectomorphism is in fact the generalised Dehn twist T_{L_1} about the Lagrangian sphere L_1. Doing it twice, e.g., by double covering the above family under $u \mapsto u^2$ to give the family with local model at $u = 0$,

$$\sum_{i=0}^{2} x_i^2 = u^2$$

gives monodromy $T_{L_1}^2$.

The advantage with working with this second family is that it preserves the homology class $[L_1]$. $(T_{L_1})_*[L_1] = -[L_1]$, so it does not make sense to talk about the homology class $[L_1]$ globally in the original family; $\int_{L_1}\Omega$ is only defined up to sign and vanishes to order $1/2$ at $u = 0$. In this double cover $[L_1]$ is well-defined globally, and $\int_{L_1}\Omega$ has a simple zero at $u = 0$. Then it is a result of Paul Seidel that

$$T_{L_1}^2(L_2 \# L_1) \approx L_1 \# L_2, \tag{38.26}$$

as graded Lagrangians. So for different values (and phases of) $\int_{L_1}\Omega$, the situation is as shown in Fig. 1 (in which we have chosen Ω^t such that $\phi(L_2) \equiv 0$; this is possible since its winding is zero as we go round $u = 0$). Also, the winding number of $\phi(L_1)$ is $+1$ (since we observed that $\int_{L_1}\Omega$ has a simple zero at $u = 0$), so we may take it to be equal to u. We plot the image of the projection of complex structure moduli space to $\mathbb{C}$ via $u = \int_{L_1} \Omega$.

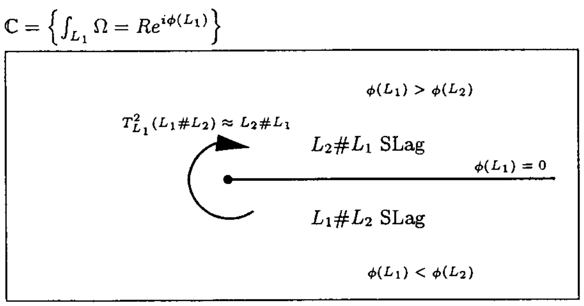


FIGURE 1. $\left(\int_{L_1} \Omega\right)$-space in a $K3$, with polar coordinates $(R,\, \phi(L_1))$

In all dimensions the deformation of the sLag, as the complex structure is varied, is unobstructed while it is a smooth submanifold. In our family it becomes singular on the $\phi = 0$ axis as the union $L_1 \cup L_2$, but let us assume for the purposes of this example that nowhere else can the sLag be split up in such a way — certainly it cannot using the homology classes $[L_1]$, $[L_2]$. (Note that the $\phi = 0$ axis really does end at the origin — when $\phi = \pi$ the union $L_1 \cup L_2$ is *not* sLag; only in the homology class of $-[L_1]$ is there a sLag of phase 0, not in $[L_1]$.) So assuming no singularities develop, we get a sLag in the Hamiltonian deformation class of $L_2\#L_1$ when we come back round to $\phi \downarrow 0$. This is in the same homology class Eq. (38.24) but a different Hamiltonian deformation class as $L_1\#L_2$,

$$L_1\#L_2 \not\approx L_2\#L_1,$$

and both have $L_1 \cup L_2$ in their closures, L^t changing class as it passes through $\phi = 0$.

The analogue with the bundle case Eq. (38.16) is clear, as is the inequality Eq. (38.25). At first sight this inequality might appear to be violated by the sLag $L_2\#L_1$ if we write it in the form

$$L_2\#L_1 \approx T^2_{L_1} L_1 \# T^2_{L_1} L_2,$$

suggested by the monodromy, as the connect sum of Lagrangians of phases $\epsilon > 0$ and zero, respectively. But following the monodromy around, we see

that in fact the phase of $T^2_{L_1} L_1$ is $-2\pi + \epsilon$ rather than ϵ, and Eq. (38.25) is not violated *so long as we keep track of the grading*; the graded connect sum with $T^2_{L_1} L_1$ of phase ϵ does not exist — any attempt to give a grading θ to the resulting Lagrangian would be discontinuous. That is, $T^2_{L_1} L_1$ is *not* L_1 as a graded Lagrangian, but

$$T^2_{L_1} L_1 \approx L_1[-2], \tag{38.27}$$

i.e., L_1 with its grading shifted by -2π, so that

$$L_2\#L_1 \approx L_1[-2]\# T^2_{L_1} L_2$$

does not violate the phase inequality Eq. (38.25) in either presentation as a connect sum.

Returning to the threefold case, the picture is different; see Fig. 2.

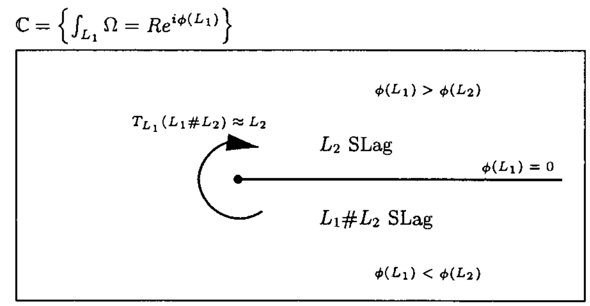


FIGURE 2. $\left(\int_{L_1} \Omega\right)$-space in a threefold, with polar coordinates $(R,\, \phi(L_1))$

$T_{L_1}[L_1]$ is now $+[L_1]$, $\int_{L_1} \Omega$ is well defined and winds once round 0 in the family $\sum_{i=0}^3 x_i^2 = u$ (with vanishing cycle the Lagrangian three-sphere L_1) without squaring the monodromy T_{L_1}. This is because in three dimensions $T_{L_1} L_1 \approx L_1[-2]$, as opposed to $T_{L_1} L_1 \approx L_1[-1]$ and Eq. (38.27) in two dimensions. We now have

$$T_{L_1}(L_1\#L_2) \approx L_2. \tag{38.28}$$

So there is a sLag on the other side of the $\phi(L_1) = 0$ wall, it is just in a different homology class than $L_1\#L_2$. Similarly this can be decomposed into a number of (non-destabilising) connect sums via monodromy:

$$T_{L_1}(L_1\#L_2) \;\approx\; L_1[-2]\#(L_2\#L_1[-1]) \;\approx\; L_2 \;\approx\; (L_1\#L_2)\#L_1[1].$$

Mirror to these geometric monodromy transformation is, according to Kontsevich's conjecture, the rather more abstract action of an autoequivalence of $D^b(M)$ (usually *not* induced by an automorphism of M). Here it is a certain "twist" T_{E_1} on $D^b(M)$, mirror to the Dehn twist T_{L_1}, easily described in the references in terms of homological algebra, whose action on homology reduces to the familiar Picard–Lefschetz monodromy of Sec. 18.2, and whose square's action on the extension E (Eq. (38.15)) can be computed to be the extension E' (Eqs. (38.16, 38.20)),

$$T_{E_1}^2\, E = E', \tag{38.29}$$

compare Eq. (38.26), and, on a CY threefold,

$$T_{E_1}\, E = E_2, \tag{38.30}$$

compare Eq. (38.28).

38.4.2. Stability. Interpreting the above example in terms of a constant *stable* Hamiltonian deformation class of Lagrangians becoming semistable at $\phi(L_1) = 0$ (and so a sLag representative existing only on a Lagrangian in the closure of this class) and unstable for $\phi(L_1) > 0$ (and so admitting no sLag representative), sheds some light on what the stability condition for arbitrary Lagrangians (and not just sLags) should be.

We also have to consider connect sums of Lagrangians intersecting at a number of points. Then the connect sum is not unique up to Hamiltonian deformation — H_1 is added to the resulting Lagrangian as loops between the intersection points, giving additional deformations of its Hamiltonian isotopy class. The upshot is that we assign a nonzero number at each intersection point, which scales the size of the neck of the connect sum at that point; we denote any such Lagrangian by $L_1\#L_2$. There is also a notion of a *relative*, or family, Lagrangian connect sum for *clean* intersections $L_1 \cap L_2$; roughly speaking these intersections are transverse normal to a smooth submanifold $L_1 \cap L_2$, and the connect sum is a family version of the usual one, fibred over $L_1 \cap L_2$.

DEFINITION 38.4.1. *Fix graded Lagrangians (L_1, θ_1) and (L_2, θ_2), Hamiltonian isotoped to intersect cleanly, and such that the (relative) Lagrangian connect sums $L_1\#L_2$ can be graded. Then a Lagrangian L of Maslov class zero is said to be destabilised by the L_i if it is Hamiltonian isotopic to such a graded connect sum $L_1\#L_2$, and the phases (*real *numbers, induced by the gradings) satisfy*

$$\phi(L_1) \geq \phi(L_2).$$

If L is not destabilised by any such L_i, then it is called stable.

It turns out that a graded Lagrangian (L, θ) can only be destabilised by L_1, L_2 with phases $\phi(L_i)$ satisfying

$$\sup\theta \geq \phi(L_1) \geq \phi(L_2) \geq \inf\theta,$$

so this narrows down the necessary checks (analogously to proper subsheaves of vector bundles having rank between 0 and the rank of the bundle). Also, one may check that if two Lagrangians are Hamiltonian deformations of one another by checking that the flux between them (the integral of ω over the two-chains swept out by closed loops in the Lagrangians under the deformation) is zero, so long as we know they they are deformations of each other as Lagrangians. This last condition is far harder to check, and currently poorly understood, unfortunately.

There is no notion of stability for all objects of the derived category of coherent sheaves on the mirror; there can only be one involving subclasses of objects such as "abelian subcategories" (for instance the subcategory of sheaves, thought of as one-step complexes). What we are proposing here is that in fact this can be extended to all objects mirror to a single Lagrangian (rather than a general object of the derived Fukaya category).

The natural hope then would be for a Hitchin–Kobayashi-type correspondence for Lagrangians of Maslov class zero and phase $\phi(L) = 0$ (with loss of generality) and pointwise phase θ taking values in some fixed interval $(-\delta, \delta)$ over the Lagrangian (for sufficiently small δ). That is, one would like Hamiltonian deformation classes of Lagrangians to admit a (unique) sLag representative if and only if the Lagrangian is stable. This should be reached by mean curvature flow, which is the gradient flow of minus the norm square of the moment map $\operatorname{Im}\Omega|_L$, and is Hamiltonian.

Proving this conjecture seems a long way off, as mean curvature flow is hard to control. But in this context of Lagrangians in Calabi–Yau manifolds there are reasons to hope it may behave better; for instance, under the flow θ satisfies a heat equation and so a maximum principle, for Lagrangians whose phase starts inside $(-\pi/2, \pi/2)$, the forms $\operatorname{Re}\Omega|_L$ and vol_L remain boundedly comparable by Eq. (38.22); this helps the analysis of what singularities may develop. Also, knowing what the conjectural limit should be (even in the unstable case) helps: i.e., the analogue of the Jordan–Hölder decomposition; in the example studied above this would just be the union of sLags

$$L_1^t \cup L_2^t$$

for $\phi(L_1) > \phi(L_2)$; this is not itself sLag but is still stationary (it is a critical point for the volume functional on submanifolds) and in fact minimal in odd dimensions as it satisfies the "angle criterion." By "converge" we actually expect that in finite time the flow would split the Lagrangian into destabilising pieces; these should then flow to sLags.

If the conjecture were true we would get a way of dealing with sLags in terms of symplectic geometry and Lagrangians, just as HYM connections are studied in terms of algebraic geometry and holomorphic bundles. While the HYM and sLag equations seem not be *explicitly* soluble, an analogue of Theorem 38.2.2 for Lagrangians would give a powerful existence result for sLags. See the references for examples where the conjecture has been proved; one trivial case is the two-torus, which is a good illustrative example.

38.4.3. The Two-torus. We already understand mirror symmetry for T^2, at least at the level of sLags and *stable* bundles, from Sec. 37.8.1. The extension to graded Lagrangians works out simply: mean curvature flow for curves of Maslov class zero converges to straight lines. This is a mirror symmetric analogue of Atiyah's classification of sheaves on an elliptic curve — they are basically all sums of stable sheaves. The only exceptions are the non-trivial extensions of certain sheaves by themselves; these correspond to thickenings of the corresponding special Lagrangian; such singular fat sLags do not concern us here.

We give in Fig. 3 an example to demonstrate why one cannot form smooth *unstable* Lagrangians on T^2. First, giving L_1 and L_2 the gradings such that their phases are 0 and $\pi/4$, we expect $L_1\#L_2$ to be stable, and indeed we see it is Hamiltonian deformation equivalent to the slope 1/2 sLag shown. This is mirror to the slope 1/2 stable extension E of $\mathcal{O}$ by $\mathcal{O}(p)$ (where p is a basepoint of T^2 with corresponding line bundle mirror to the diagonal sLag drawn); recall from Sec. 37.8.1 that slope corresponds to slope under the mirror map.

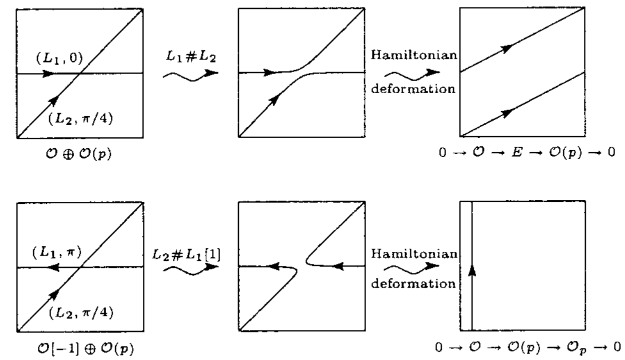


FIGURE 3. $L_1\#L_2$, $L_2\#L_1[1]$, equivalent sLags, and mirror sheaves

If one then tries to form an unstable sLag $L_2\#L_1$, the *graded* connect sum does not exist — the phase would become discontinuous. To form $L_2\#L_1$ we see from the diagram that we have to take the phase of L_1 to be π, thus reversing its orientation. Then the stability inequality Eq. (38.25) is not violated, and in fact this Lagrangian is stable and Hamiltonian deformation equivalent to the sLag in T^2 represented by the vertical edge of the square (and so drawn with a little artistic license in Fig. 3). Under the mirror map this corresponds to replacing the extension Ext^1 class by a Hom (adding π to the phase is mirror to shifting complexes of sheaves by one place) and taking its cone in the derived category; this is the infinite slope cokernel $\mathcal{O}_p$ of the Fig., mirror to the vertical line.

In fact this example can be extended to show that we cannot form the graded connect sum $L_1\#L_2$ of any two Lagrangians (via a class in $HF^*(L_1, L_2)$) if $\phi(L_1) > \phi(L_2)$. So unstable smooth Lagrangians do not exist on T^2, and our conjecture is true in this simple case.

CHAPTER 39

Boundary $\mathcal{N} = 2$ Theories

In this chapter, we formulate and study quantum field theory on $1 + 1$ dimensional manifolds with boundary. Such a system appears, among others, as the theory on the worldsheet of an open string. In some cases the open string propagates freely in the target space-time, but in some other cases an end point may be constrained in some submanifold. The end point may or may not be charged under some gauge potential. Various cases correspond to various types of D-branes in the target space-time. From the point of view of worldsheet quantum field theory, the distinction is made by a boundary condition on the worldsheet fields or boundary interactions. We will study such a system purely from the worldsheet point of view.

In earlier chapters, we learned a lot about quantum field theories in $1 + 1$ dimensions. However, we have not learned much about field theories on a manifold with boundary. Such a system is indeed difficult to study in general, and the current understanding is very poor compared to the theory on a manifold without boundary. In such a situation, some of the procedures we have taken in learning about QFT are useful. There are two main approaches toward a better understanding. One is the study of free field theories where one can explicitly quantize the system. This will be useful to develop our intuitions, and also to provide a basis for the study of more general interacting systems. The other approach is to impose supersymmetry. Some properties and quantities are protected from quantum corrections, and one can determine them exactly. We will indeed take both of these approaches.

In Sec. 39.1, we study free field theories with linear boundary conditions that keep the system free. We will also introduce the notion of boundary states. In Sec. 39.2, we determine the condition under which half of the bulk $(2, 2)$ supersymmetry is conserved. We will find A-type and B-type D-branes, which are objects of interest from symplectic geometry and complex geometry respectively. In Sec. 39.3, we study the axial anomaly induced from the boundary conditions. The Maslov index plays a role analogous to that of the first Chern class in the bulk theory. In Sec. 39.4, we determine the spectrum of supersymmetric ground states of various open string systems. In Sec. 39.5, we study the properties of boundary states and extract some exact information. This will also be useful in learning about the D-brane charge in the target space-time.

39.1. Open Strings — Free Theories

In this section, we study the free quantum field theory formulated on worldsheets with boundaries. In particular, we consider linear boundary conditions so that the boundary theory is also free. For such boundary conditions, we will be able to explicitly quantize open strings. This study can be used to develop intuitions for theories formulated on worldsheets with boundaries. Along the way, we introduce notions such as boundary states and boundary entropy. Free boundary theories also provide starting points for perturbation theory, where interactions are introduced both in the bulk and also in the boundary.

39.1.1. Boundary Conditions. The theories we study in this section are from Ch. 11: (i) massless scalar theory (the sigma model on $\mathbb{R}$ or S^1 and also on $\mathbb{C}$ or T^2), (ii) massless Dirac fermion. Let us formulate them on the "left half-plane" $\Sigma = (-\infty, 0] \times \mathbb{R}$ parametrized by the space coordinate $-\infty < s \leq 0$ and the time coordinate $-\infty < t < \infty$. The worldsheet has the time-like boundary $\partial\Sigma = \{s = 0\} \times \mathbb{R}$. We shall determine the boundary conditions of the fields at this boundary.

(i) Massless Scalar Field. Let us first consider the theory of a massless scalar field $x(t,s)$, which is either single-valued or periodic, $x \equiv x + 2\pi R$. The action is given by

$$S = \frac{1}{4\pi}\int_{\Sigma}\left\{(\partial_t x)^2 - (\partial_s x)^2\right\} dtds. \tag{39.1}$$

The variation of the action with respect to $x \to x + \delta x$ is

$$\delta S = \frac{1}{2\pi}\int_{\Sigma}\delta x\left(-\partial_t^2 x + \partial_s^2 x\right) dtds - \frac{1}{2\pi}\int_{\partial\Sigma}\delta x \partial_s x dtds. \tag{39.2}$$

For the equation of motion to be the same as $(\partial_t^2 - \partial_s^2)x = 0$, we need to require $\delta x \partial_s x = 0$ along the boundary $\partial\Sigma$. There are two kinds of solutions to this. One is the Neumann boundary condition

$$\partial_s x = 0 \ \text{ along } \ \partial\Sigma. \tag{39.3}$$

The other is the Dirichlet boundary condition

$$x = x_* \ \text{(fixed)} \ \text{ along } \ \partial\Sigma, \tag{39.4}$$

where x_* is an arbitrary fixed value. For the Neumann boundary condition, we can consider adding the boundary term

$$S_{\text{boundary}} = -\int_{\partial\Sigma} a\partial_t x dt \tag{39.5}$$

to the action.

D-branes. One can also consider a theory of many scalar fields where a Neumann or Dirichlet boundary condition is imposed on each of them. Such a system corresponds to the theory on the worldsheet of a string ending on a space-time object called a *D-brane* (or a *Dirichlet-brane*). Let us consider the theory of $d+1$ scalar fields $X^0, X^1, \ldots, X^d$. Without a boundary, the $1+1$ dimensional theory can be considered as the theory on the worldsheet of a string freely propagating in $(d+1)$-dimensional space-time. (We consider X^0 as the time and $X^1, \ldots, X^d$ as the space coordinates.) Let us consider the worldsheet with boundary, where the boundary can be considered as the worldline of an end point of an open string. If we impose a Neumann boundary condition on $X^0, X^1, \ldots, X^p$ and a Dirichlet boundary condition on $X^{p+1}, \ldots, X^d$, the worldline can be anywhere in the first $p+1$ coordinates but is confined at a fixed position $(X_*^{p+1}, \ldots, X_*^d)$ in the last $(d-p)$ directions. This means that the string end point is fixed at $(X^I) = (X_*^I)$ (for $I = p+1, \ldots, d$). In other words, the string ends on a p-dimensional object located at $(X^I) = (X_*^I)$. It is this object that is called a Dp-brane. For the Neumann directions, $X^0, \ldots, X^p$, one can add a term like Eq. (39.5) to the action:

$$S_{\text{boundary}} = -\int_{\partial\Sigma}\sum_{\mu=0}^{p} a_\mu \partial_t X^\mu dt. \tag{39.6}$$

This addition is identical to the coupling of a point particle to the flat gauge field $a = a_\mu dX^\mu$ in $p+1$ dimensions. That is, introduction of the term shown in Eq. (39.6) corresponds to introducing a gauge field on the Dp-brane with respect to which the open string end point is charged. This term is called the *Wilson line* term.

In what follows, we will sometimes mention the space-time picture including D-branes. In such a case, we assume that we have a time direction where a Neumann boundary condition is imposed, even though we will not mention it explicitly.

(ii) Massless Dirac Fermion. Let us next consider the system of a Dirac fermion with the action

$$S = \frac{1}{2\pi}\int_\Sigma \left(\frac{i}{2}\overline{\psi}_-(\overleftrightarrow{\partial}_t + \overleftrightarrow{\partial}_s)\psi_- + \frac{i}{2}\overline{\psi}_+(\overleftrightarrow{\partial}_t - \overleftrightarrow{\partial}_s)\psi_+\right) dtds, \tag{39.7}$$

where we define

$$\overline{\psi}\overleftrightarrow{\partial}_\mu\psi := \overline{\psi}\partial_\mu\psi - (\partial_\mu\overline{\psi})\psi. \tag{39.8}$$

The integrand is different from the ones in Ch. 11 by a total derivative that is irrelevant if Σ has no boundary. Here Σ has a boundary and the total derivative cannot be ignored. The above integrand is chosen so that it is manifestly real. The variation of the action is

$$\begin{aligned} \delta S &= \frac{i}{2\pi}\int_\Sigma \left(\delta\overline{\psi}_-\partial_+\psi_- + \delta\psi_-\partial_+\overline{\psi}_- + \delta\overline{\psi}_+\partial_-\psi_+ + \delta\psi_+\partial_-\overline{\psi}_+\right) dtds \\ &\quad + \frac{i}{4\pi}\int_{\partial\Sigma}\left(\overline{\psi}_-\delta\psi_- + \psi_-\delta\overline{\psi}_- - \overline{\psi}_+\delta\psi_+ - \psi_+\delta\overline{\psi}_+\right) dt, \end{aligned} \tag{39.9}$$

where $\partial_\pm = \frac{1}{2}(\partial_t \pm \partial_s)$. For the equation of motion $(\partial_t + \partial_s)\psi_- = 0$ etc. not to be altered, the boundary term must vanish. There are two kinds of solutions. One is

$$\psi_+ = e^{-i\beta}\psi_-, \quad \overline{\psi}_+ = e^{i\beta}\overline{\psi}_-, \tag{39.10}$$

which we call the B_β boundary condition. The other is

$$\psi_+ = e^{-i\alpha}\overline{\psi}_-, \quad \overline{\psi}_+ = e^{i\alpha}\psi_-, \tag{39.11}$$

which we call the A_α boundary condition. Recall that the bulk theory has vector and axial rotation symmetries

$$V : \psi_\pm \mapsto e^{-i\gamma}\psi_\pm, \tag{39.12}$$

$$A : \psi_\pm \mapsto e^{\mp i\gamma}\psi_\pm. \tag{39.13}$$

The B_β (resp. A_α) boundary condition is not preserved by the axial (resp. vector) rotation except for those with $e^{i\gamma} = \pm 1$. (Others would change the boundary condition to $B_{\beta-2\gamma}$ (resp. $A_{\alpha-2\gamma}$).) On the other hand vector (resp. axial) rotation remains as a symmetry of the system with the B_β (resp. A_α) boundary condition.

39.1.2. Quantization of Open Strings. Let us now formulate the free systems on the strip

$$\Sigma = [0,\pi] \times \mathbb{R}, \tag{39.14}$$

where the spatial coordinate s spans the segment $[0,\pi]$ while the time runs from $-\infty$ to ∞. At the two boundary lines, we impose the boundary conditions introduced above. We may consider Σ as the worldsheet of an open string where the boundary $\partial\Sigma = (\{s=\pi\}\times\mathbb{R}) \cup (\{s=0\}\times\mathbb{R})$ may be regarded as the worldlines of the two end points of the open string. Here we quantize this open string system. We first consider the scalar field theory where we choose the Neumann (N) and/or Dirichlet (D) boundary condition, and next consider the Dirac fermion system with the B_β or A_α boundary condition. In any of these cases, the system has time translation symmetry generated by the Hamiltonian of the system. However, by the presence of the boundary, space-translation symmetry of the bulk theory is lost. Thus, there will be no worldsheet momentum.

39.1.2.1. *NN.* We first consider the massless scalar system where the Neumann boundary condition is imposed at both boundary lines.

$$\partial_s x = 0 \ \text{ at } s = 0, \pi. \tag{39.15}$$

This may be regarded as the theory of an open string ending on a D1-brane. We first consider the case where x takes values in $\mathbb{R}$. The space of functions obeying the condition from Eq. (39.15) is spanned by $\cos(ns)$ with $n = 0,1,2,\ldots$. Thus, the field $x(t,s)$ can be expanded as

$$x(t,s) = x_0(t) + \sum_{n=1}^{\infty} x_n(t) 2\cos(ns), \tag{39.16}$$

where $x_n(t)$ are *real* fields. The Lagrangian is expressed as

$$L = \frac{1}{4}\dot{x}_0^{\,2} + \sum_{n=1}^{\infty}\left(\frac{1}{2}\dot{x}_n^2 - \frac{n^2}{2}x_n^2\right). \tag{39.17}$$

Thus, the system is decomposed into the sum of infinitely many systems. Since they are decoupled from one another, one can quantize each system separately. The zero mode sector has the Lagrangian $L_0 = \dot{x}_0^2/4$, and is the single-variable quantum mechanics without a potential. The conjugate momentum is $p_0 = \dot{x}_0/2$ and the Hamiltonian is $H_0 = p_0^2$. The nth sector has the Lagrangian $L_n = \frac{1}{2}\dot{x}_n^2 - \frac{n^2}{2}x_n^2$, and is the harmonic oscillator system with frequency n. As usual, we introduce the annihilation and creation operators

$a_n = (p_n/\sqrt{n} - i\sqrt{n}x_n)/\sqrt{2}$ and $a_n^\dagger = (p_n/\sqrt{n} + i\sqrt{n}x_n)/\sqrt{2}$ obeying the commutation relation $[a_n, a_n^\dagger] = 1$. The Hamiltonian is then expressed as $H_n = n(a_n^\dagger a_n + 1/2)$. The total system is their sum, and operators in different sectors commute with each other. Let us introduce the notaion $\alpha_n = \sqrt{n}a_n$ and $\alpha_{-n} = \sqrt{n}a_n^\dagger$. Then these operators obey the commutation relations

$$[\alpha_n, \alpha_m] = n\delta_{n+m,0}. \tag{39.18}$$

The total Hamiltonian is the sum

$$\begin{aligned} H_o &= \sum_{n=0}^{\infty} H_n = p_0^2 + \sum_{n=1}^{\infty}\left(\alpha_{-n}\alpha_n + \frac{n}{2}\right), \\ &= p_0^2 + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - \frac{1}{24}, \end{aligned} \tag{39.19}$$

where the zeta function regularization is used to sum up the ground state oscillation energies, $\sum_{n=1}^{\infty} n/2 = \zeta(-1)/2 = -1/24$. The Hamiltonian H_o and the operators x_0, p_0, α_n have the commutation relations $[H_o, x_0] = -2ip_0$, $[H_o, p_0] = 0$ and $[H_o, \alpha_n] = -n\alpha_n$. Thus, we find $x_0(t) = x_0 + 2tp_0$ and $\alpha_n(t) = \mathrm{e}^{-int}\alpha_n$, which yield

$$x(t,s) = x_0 + 2tp_0 + \frac{i}{\sqrt{2}}\sum_{n\neq 0}\frac{\alpha_n}{n}\mathrm{e}^{-int}2\cos(ns). \tag{39.20}$$

The boundary condition given by Eq. (39.15) preserves the target space translation symmetry δx =constant. The corresponding conserved charge is

$$p = \frac{1}{2\pi}\int_0^\pi \dot{x}ds = \frac{1}{2}\dot{x}_0 = p_0, \tag{39.21}$$

which indeed commutes with H_o. The partition function of the system factorizes into the product of such for the zero mode and the oscillator modes:

$$\begin{aligned} \mathrm{Tr}\,\mathrm{e}^{-2\pi T H_o} &= \mathrm{Tr}\,\mathrm{e}^{-2\pi T\left[(p_0-\Delta a)^2 + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - \frac{1}{24}\right]} \\ &= \mathrm{Tr}_0\,\mathrm{e}^{-2\pi T p_0^2} \times \prod_{n=1}^{\infty}\mathrm{Tr}_n\,\mathrm{e}^{-2\pi T\alpha_{-n}\alpha_n} \times \mathrm{e}^{-2\pi T(-\frac{1}{24})}. \end{aligned} \tag{39.22}$$

The zero mode partition function is evaluated as

$$\mathrm{Tr}_0\,\mathrm{e}^{-2\pi T p_0^2} = V\int_{-\infty}^{\infty}\frac{dk}{2\pi}\mathrm{e}^{-2\pi T k^2} = \frac{V}{2\pi}\frac{1}{\sqrt{2T}}, \tag{39.23}$$

where V is the cut-off volume that is introduced in order to make the partition function finite. For the oscillator mode, we have

$$\mathrm{Tr}_n\,\mathrm{e}^{-2\pi T\alpha_{-n}\alpha_n} = \sum_{k=0}^{\infty} q_o^{nk} = \frac{1}{1-q_o^n}, \tag{39.24}$$

where we introduced

$$q_o = \mathrm{e}^{-2\pi T}. \tag{39.25}$$

Collecting things together, we find

$$\mathrm{Tr}\,q_o^{H_o} = \frac{V}{2\pi}\frac{1}{\sqrt{2T}} \times \prod_{n=1}^{\infty}\frac{1}{1-q_o^n} \times q_o^{\frac{1}{24}} = \frac{V}{2\pi}\frac{1}{\sqrt{2T}}\frac{1}{\eta(iT)} \tag{39.26}$$

where $\eta(iT) = q_o^{\frac{1}{24}}\prod_{n=1}^{\infty}(1-q_o^n)$ is the Dedekind eta function.

Let us next consider the sigma model on S^1 of radius R where x is a periodic variable, $x \equiv x + 2\pi R$. We also turn on the boundary term

$$S_{\text{boundary}} = -\frac{1}{2\pi}\int_{s=\pi} a_\pi \dot{x}dt + \frac{1}{2\pi}\int_{s=0} a_0\dot{x}dt \tag{39.27}$$

where a_0 and a_π are real numbers defined modulo $2\pi/R$. This alters the expression of the Lagrangian in Eq. (39.17) to

$$L = \frac{1}{4}\dot{x}_0^2 - \frac{\Delta a}{2\pi}\dot{x}_0 + \sum_{n=1}^{\infty}\left(\frac{1}{2}\dot{x}_n^2 - \frac{1}{\pi}(a_\pi(-1)^n - a_0)\dot{x}_n - \frac{n^2}{2}x_n^2\right), \tag{39.28}$$

where

$$\Delta a = a_\pi - a_0. \tag{39.29}$$

The linear term in $\dot{x}_n$ for $n \geq 1$ does not affect the final result for the Hamiltonian $H_n = n(a_n^\dagger a_n - 1/2)$. However, for the zero mode, we find $H_0 = (p_0 + \frac{\Delta a}{2\pi})^2$ where $p_0 = \partial L/\partial\dot{x}_0 = \frac{\dot{x}_0}{2} - \frac{\Delta a}{2\pi}$ is the conjugate momentum. Furthermore, since x_0 is a periodic variable, the momentum has to be quantized; $p_0 = l/R$ with $l \in \mathbb{Z}$ just as in the closed string theory. However, unlike in the closed string, there is no winding number since the open string boundary can move freely. Thus, the space of states has only a single grading (by momentum); $\mathcal{H} = \bigoplus_{l\in\mathbb{Z}}\mathcal{H}_l$. The partition function of the system is given by

$$\begin{aligned}\operatorname{Tr} q_o^{H_o} &= \operatorname{Tr} q_o^{[(p_0+\Delta a/2\pi)^2+\sum_{n=1}^{\infty}\alpha_{-n}\alpha_n-\frac{1}{24}]}\\ &= q_o^{-\frac{1}{24}}\sum_{l\in\mathbb{Z}} q_o^{(l/R+\Delta a/2\pi)^2}\prod_{n=1}^{\infty}\frac{1}{1-q_o^n}\\ &= \frac{1}{\eta(iT)}\sum_{l\in\mathbb{Z}} q_o^{(l/R+\Delta a/2\pi)^2}. \qquad (39.30)\end{aligned}$$

Note that we recover Eq. (39.26) in the $R \gg 1$ limit, where $V = 2\pi R$.

39.1.2.2. *DD.* We next consider the case where the Dirichlet boundary condition is imposed at both boundary lines.

$$x = x_{*0} \text{ at } s = 0 \text{ and } x = x_{*\pi} \text{ at } s = \pi. \qquad (39.31)$$

This may be regarded as the theory of an open string stretched from a D0-brane at $x = x_0$ to a D0-brane at $x = x_\pi$. The field $x(t,s)$ obeying this condition is expanded as

$$x(t,s) = x_{*0} + \Delta x\frac{s}{\pi} + \sum_{n=1}^{\infty} x_n(t) 2i\sin(ns), \qquad (39.32)$$

where

$$\Delta x = x_{*\pi} - x_{*0}. \qquad (39.33)$$

The Lagrangian is the sum of $-(\Delta x/2\pi)^2$ and $(\dot{x}_n^2 - n^2x_n^2)/2$ and can be quantized as before. We then obtain the Hamiltonian

$$H_o = \left(\frac{\Delta x}{2\pi}\right)^2 + \sum_{n=1}^{\infty}\alpha_{-n}\alpha_n - \frac{1}{24}, \qquad (39.34)$$

where α_n obey the same commutation relation as in Eq. (39.18). The field x is expressed as

$$x(t,s) = x_{*0} + \Delta x\frac{s}{\pi} + \frac{i}{\sqrt{2}}\sum_{n\neq 0}\frac{\alpha_n}{n}\mathrm{e}^{-int}2i\sin(ns). \qquad (39.35)$$

(Target space) momentum is lost since the boundary condition from Eq. (39.31) breaks the target translation symmetry δx =constant. The partition function is given by

$$\operatorname{Tr} q_o^{H_o} = \frac{q_o^{(\Delta x/2\pi)^2}}{\eta(iT)}. \qquad (39.36)$$

If we consider the sigma model on S^1 of radius R, we can have configurations with non-trivial winding number. In the sector with winding number $m \in \mathbb{Z}$, the above formulae are modified so that

$$\Delta x \to \Delta x + 2\pi R m. \qquad (39.37)$$

The partition function of the system is then

$$\operatorname{Tr} q_o^{H_o} = \frac{1}{\eta(iT)}\sum_{m\in\mathbb{Z}} q_o^{(mR+\Delta x/2\pi)^2}. \qquad (39.38)$$

39.1.2.3. *DN.* As the final example, we impose the Dirichlet boundary condition at one end while the Neumann boundary condition is imposed at the other end;

$$x = x_* \text{ at } s = 0 \text{ and } \partial_s x = 0 \text{ at } s = \pi. \qquad (39.39)$$

The theory may be regarded as that of an open string stretched between a D0-brane at $x = x_*$ and a D1-brane. The field $x(t,s)$ obeying this condition is expanded as

$$x(t,s) = x_* + \sum_{n=0}^{\infty} x_{n+\frac{1}{2}}(t) 2i\sin\left(\left(n+\frac{1}{2}\right)s\right). \qquad (39.40)$$

The Lagrangian is the sum of $\frac{1}{2}\dot{x}_{n+\frac{1}{2}}^2 - \frac{1}{2}(n+\frac{1}{2})^2 x_{n+\frac{1}{2}}^2$. Quantization is as before and we obtain the Hamiltonian

$$H_o = \sum_{n=0}^{\infty}\left(\alpha_{-n-\frac{1}{2}}\alpha_{n+\frac{1}{2}} + \frac{1}{2}\left(n+\frac{1}{2}\right)\right) = \sum_{n=0}^{\infty}\alpha_{-n-\frac{1}{2}}\alpha_{n+\frac{1}{2}} + \frac{1}{48}, \qquad (39.41)$$

where α_r obey the commutation relation $[\alpha_r, \alpha_{r'}] = r\delta_{r+r',0}$. To sum up the ground state oscillation energies, we have used the zeta function regularization $\sum_{n=0}^{\infty}(n+1/2) = \zeta(-1,\frac{1}{2}) = 1/24$. The field $x(t,s)$ is expressed as

$$x(t,s) = x_* + \frac{i}{\sqrt{2}}\sum_{r\in\mathbb{Z}+\frac{1}{2}}\frac{\alpha_r}{r}\mathrm{e}^{-irt}2i\sin(rs). \qquad (39.42)$$

The partition function is given by

$$\operatorname{Tr} q_o^{H_o} = q_o^{\frac{1}{48}}\prod_{n=0}^{\infty}\frac{1}{1-q_o^{n+\frac{1}{2}}}. \qquad (39.43)$$

This holds both for the case where x is a single-valued field (the sigma model on $\mathbb{R}$) and also for the case where it is a periodic field (the sigma model on a circle).

39.1.2.4. *Intersecting D1-Branes in* $\mathbb{R}^2$. Let us consider a slightly more complicated but interesting example. The bulk theory is the non-linear sigma model on $\mathbb{R}^2 \cong \mathbb{C}$, namely the free theory of a massless complex scalar field z. We consider one D1-brane located at the real axis ($\mathrm{Im}(z)=0$) and another D1-brane at its angle θ rotation ($\mathrm{Im}(\mathrm{e}^{-i\theta}z)=0$). We consider an open string stretched from the first D1-brane to the second D1-brane. The boundary condition on the field is

$$\partial_s \mathrm{Re}(z)=0, \quad \mathrm{Im}(z)=0, \quad \text{at } s=0, \tag{39.44}$$

$$\partial_s \mathrm{Re}(\mathrm{e}^{-i\theta}z)=0, \quad \mathrm{Im}(\mathrm{e}^{-i\theta}z)=0, \quad \text{at } s=\pi. \tag{39.45}$$

The fields obeying these boundary conditions can be expanded as

$$z(t,s)=\sum_{r\in\mathbb{Z}+\frac{\theta}{\pi}} z_r\,\mathrm{e}^{irs}, \tag{39.46}$$

where

$$z_r\in\mathbb{R}, \quad \forall r\in\mathbb{Z}+\frac{\theta}{\pi}.$$

In terms of the variables z_r, the Lagrangian is expressed as

$$L=\sum_{r\in\mathbb{Z}+\frac{\theta}{\pi}}\left\{\frac{1}{4}(\dot{z}_r)^2-\frac{r^2}{4}(z_r)^2\right\}. \tag{39.47}$$

If θ/π is not an integer, the system consists of infinitely many harmonic oscillators that are decoupled from each other. For the rth system, the Hamiltonian is $H_r=|r|(a_r^\dagger a_r+1/2)$ where a_r and $a_r^\dagger$ are annihilation and creation operators obeying $[a_r,a_r^\dagger]=1$. (The relation to the original variable z_r and its conjugate momentum $p_r=\dot{z}_r/2$ is $a_r=(p_r-i|r|z_r/2)/\sqrt{|r|}$ and $a_r^\dagger=(p_r+i|r|z_r/2)/\sqrt{|r|}$. The ground state is the tensor product of the ground states of these oscillator systems. In particular, the ground state oscillation energy is the (regularized) sum:

$$\begin{aligned} E_0 &= \sum_{r\in\mathbb{Z}+\frac{\theta}{\pi}}\frac{|r|}{2} \\ &= \sum_{n=0}^{\infty}\frac{1}{2}\left(n+\frac{\theta}{\pi}\right)+\sum_{n=1}^{\infty}\frac{1}{2}\left(n-\frac{\theta}{\pi}\right) \\ &= \frac{1}{2}\zeta(-1,\frac{\theta}{\pi})+\frac{1}{2}\zeta(-1,1-\frac{\theta}{\pi}) \\ &= \frac{1}{24}-\frac{1}{2}\left(\frac{\theta}{\pi}-\left[\frac{\theta}{\pi}\right]-\frac{1}{2}\right)^2. \end{aligned} \tag{39.48}$$

In the intermediate step of the above manipulation, we assumed $0\le\frac{\theta}{\pi}<1$, but the final expression is written in a way valid for any θ. The partition function is thus

$$\begin{aligned} \mathrm{Tr}\,q_o^{H_o} &= q_o^{\frac{1}{24}-\frac{1}{2}(\frac{\theta}{\pi}-[\frac{\theta}{\pi}]-\frac{1}{2})^2}\prod_{r\in\mathbb{Z}+\frac{\theta}{\pi}}\frac{1}{1-q^{|r|}} \\ &= \mathrm{e}^{\pi i(\frac{\theta}{\pi}-[\frac{\theta}{\pi}]-\frac{1}{2})}\frac{\eta(iT)}{\vartheta\left[\begin{smallmatrix}\frac{\theta}{\pi}-\frac{1}{2}\\ \frac{1}{2}\end{smallmatrix}\right](0,iT)}. \end{aligned} \tag{39.49}$$

(For the definition and properties of the theta functions $\vartheta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right]$, see Sec. 11.4.2.) It diverges as $\theta/\pi\to n\in\mathbb{Z}$ — the case where the two D-branes become parallel. The divergence comes from the infinite volume (real line) along which the open string can freely move.

39.1.2.5. *BB.* We move on to the massless Dirac fermion. We first impose the B-boundary condition at both boundary lines, B_0 at $s=0$ and B_β at $s=\pi$;

$$\begin{aligned} \psi_- &= \psi_+ \quad \text{at } s=0, \\ \psi_- &= \mathrm{e}^{i\beta}\psi_+ \quad \text{at } s=\pi. \end{aligned} \tag{39.50}$$

Then by the equation of motion we also have $\partial_s(\psi_-+\psi_+)=0$ at $s=0$ and $\partial_s(\psi_-+\mathrm{e}^{i\beta}\psi_+)=0$ at $s=\pi$. The fields obeying these boundary conditions can be expanded as

$$\psi_-=\sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}}\psi_r(t)\,\mathrm{e}^{irs}, \quad \overline{\psi}_-=\sum_{r'\in\mathbb{Z}-\frac{\beta}{2\pi}}\overline{\psi}_{r'}(t)\,\mathrm{e}^{ir's}, \tag{39.51}$$

$$\psi_+=\sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}}\psi_r(t)\,\mathrm{e}^{-irs}, \quad \overline{\psi}_+=\sum_{r'\in\mathbb{Z}-\frac{\beta}{2\pi}}\overline{\psi}_{r'}(t)\,\mathrm{e}^{-ir's}, \tag{39.52}$$

where $\psi_r^\dagger=\overline{\psi}_{-r}$. The Lagrangian is expressed in terms of ψ_r as

$$L=\sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}}\overline{\psi}_{-r}(i\partial_t-r)\psi_r. \tag{39.53}$$

Now we recognize that this system is half of the Dirac fermion formulated on $S^1\times\mathbb{R}$ that we studied in Ch. 11. In particular, it is the "right half" obeying the twisted boundary condition with the twist parameter $a=\beta/2\pi$ (Sec. 11.3.3). Thus, we have already had experience with the quantization of the system. We have the anti-commutation relation

$$\{\psi_r, \overline{\psi}_{r'}\} = \delta_{r+r',0}, \quad \{\psi_{r_1}, \psi_{r_2}\} = \{\overline{\psi}_{r'_1}, \overline{\psi}_{r'_2}\} = 0. \tag{39.54}$$

The Hamiltonian is

$$H_o = \sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}} r\overline{\psi}_{-r}\psi_r = \sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}} r{:}\overline{\psi}_{-r}\psi_r{:} + \frac{1}{2}\left(\frac{\beta}{2\pi} - \left[\frac{\beta}{2\pi}\right] - \frac{1}{2}\right)^2 - \frac{1}{24}, \tag{39.55}$$

where the normal ordering : : is defined with respect to a ground state $|0,\beta\rangle$ annihilated by ψ_r $(r \geq 0)$ and $\overline{\psi}_{r'}$ $(r' > 0)$. We have used the zeta function regularization to sum up the ground state oscillation energies $E_0 = -\zeta(-1, 1 - \frac{\beta}{2\pi} + [\frac{\beta}{2\pi}])$. This family of vacua is discontinuous at $\beta \in 2\pi\mathbb{Z}$ at which there are two ground states ($|0, 2\pi n - \epsilon\rangle$ and $|0, 2\pi n + \epsilon\rangle$ as $\epsilon \to +0$ become the two states). As noted before, the vector rotation is a symmetry of the system with B-boundary conditions. The associated conserved charge is

$$\begin{aligned} F_V &= \frac{1}{2\pi}\int_0^\pi (\overline{\psi}_-\psi_- + \overline{\psi}_+\psi_+)ds \\ &= \sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}} \overline{\psi}_{-r}\psi_r = \sum_{r\in\mathbb{Z}+\frac{\beta}{2\pi}} {:}\overline{\psi}_{-r}\psi_r{:} + \frac{\beta}{2\pi} - \left[\frac{\beta}{2\pi}\right] - \frac{1}{2}, \end{aligned} \tag{39.56}$$

where again we have used the zeta function regularization to sum up the ground state charges. Note that the charge of the vacuum $|0,\beta\rangle$ is discontinuous at $\beta \in 2\pi\mathbb{Z}$. This is because the family $|0,\beta\rangle$ is discontinuous, and F_V itself is smooth as a function of β. The partition function of the system, weighted by $e^{-2\pi i(a-\frac{1}{2})F_V}$, is given by

$$\begin{aligned} \mathrm{Tr}\, e^{-2\pi i(a-\frac{1}{2})F_V} q_o^{H_o} &= e^{-2\pi i(a-\frac{1}{2})(\frac{\beta}{2\pi}-\frac{1}{2})} q_o^{-\frac{1}{24}+\frac{1}{2}(\frac{\beta}{2\pi}-\frac{1}{2})^2} \\ &\quad \times \prod_{n=1}^{\infty}(1 - q_o^{n-1+\frac{\beta}{2\pi}} e^{-2\pi i a})(1 - q_o^{n-\frac{\beta}{2\pi}} e^{2\pi i a}) \\ &= \frac{\vartheta\left[\begin{smallmatrix}\frac{\beta}{2\pi}-\frac{1}{2} \\ -(a-\frac{1}{2})\end{smallmatrix}\right](0, iT)}{\eta(iT)}, \end{aligned} \tag{39.57}$$

where we assumed $0 \leq \frac{\beta}{2\pi} < 1$ in the intermediate step, but the final result is independent of this choice since $\vartheta\left[\begin{smallmatrix}a+1 \\ b\end{smallmatrix}\right] = \vartheta\left[\begin{smallmatrix}a \\ b\end{smallmatrix}\right]$.

39.1.2.6. *AA*. We next consider the A-boundary condition at both boundary lines;

$$\begin{aligned} \psi_- &= \overline{\psi}_+ \ \text{ at } s = 0, \\ \psi_- &= e^{-i\alpha}\overline{\psi}_+ \ \text{ at } s = \pi. \end{aligned} \tag{39.58}$$

Since the system is essentially the same as BB case (under the exchange of ψ_- and $\overline{\psi}_-$) we will be brief. The fields have the mode expansion

$$\psi_- = \sum_{r\in\mathbb{Z}-\frac{\alpha}{2\pi}} \psi_r(t)\, e^{irs}, \quad \overline{\psi}_- = \sum_{r'\in\mathbb{Z}+\frac{\alpha}{2\pi}} \overline{\psi}_{r'}(t)\, e^{ir's}, \tag{39.59}$$

$$\psi_+ = \sum_{r'\in\mathbb{Z}+\frac{\alpha}{2\pi}} \overline{\psi}_{r'}(t)\, e^{-ir's}, \quad \overline{\psi}_+ = \sum_{r\in\mathbb{Z}-\frac{\alpha}{2\pi}} \psi_r(t)\, e^{-irs}, \tag{39.60}$$

where the modes obey the anti-commutation relations $\{\psi_r, \overline{\psi}_{r'}\} = \delta_{r+r',0}$ and $\{\psi_{r_1}, \psi_{r_2}\} = \{\overline{\psi}_{r'_1}, \overline{\psi}_{r'_2}\} = 0$. The Hamiltonian is

$$H_o = \sum_{r\in\mathbb{Z}-\frac{\alpha}{2\pi}} r\overline{\psi}_{-r}\psi_r = \sum_{r\in\mathbb{Z}-\frac{\alpha}{2\pi}} r{:}\overline{\psi}_{-r}\psi_r{:} + \frac{1}{2}\left(\frac{\alpha}{2\pi} - \left[\frac{\alpha}{2\pi}\right] - \frac{1}{2}\right)^2 - \frac{1}{24}, \tag{39.61}$$

where the normal ordering : : is defined with respect to a ground state $|0,\alpha\rangle$ annihilated by ψ_r $(r > 0)$ and $\overline{\psi}_{r'}$ $(r' \geq 0)$. This family of ground states is discontinuous at $\alpha \in 2\pi\mathbb{Z}$ where two energy levels meet. The conserved charge associated with the axial rotation symmetry is

$$F_A = -\sum_{r\in\mathbb{Z}-\frac{\alpha}{2\pi}} {:}\overline{\psi}_{-r}\psi_r{:} + \frac{\alpha}{2\pi} - \left[\frac{\alpha}{2\pi}\right] - \frac{1}{2}. \tag{39.62}$$

The weighted partition function is

$$\begin{aligned} \mathrm{Tr}\, e^{2\pi i(a-\frac{1}{2})F_A} q_o^{H_o} &= e^{2\pi i(a-\frac{1}{2})(\frac{\alpha}{2\pi}-\frac{1}{2})} q_o^{\frac{1}{2}(\frac{\alpha}{2\pi}-\frac{1}{2})^2-\frac{1}{24}} \\ &\quad \times \prod_{n=1}^{\infty}(1 - q_o^{n-1+\frac{\alpha}{2\pi}} e^{2\pi i a})(1 - q_o^{n-\frac{\alpha}{2\pi}} e^{-2\pi i a}) \\ &= \frac{\vartheta\left[\begin{smallmatrix}\frac{\alpha}{2\pi}-\frac{1}{2} \\ a-\frac{1}{2}\end{smallmatrix}\right](0, iT)}{\eta(iT)}. \end{aligned} \tag{39.63}$$

39.1.3. T-duality. Let us consider again the theories of open strings in S^1 with Neumann or Dirichlet boundary conditions. We notice that the partition function given by Eq. (39.30) for the open string with NN boundary is identical to the one given by Eq. (39.38) for DD boundary under the replacement

$$R \to 1/R, \quad l \to m, \quad \Delta a \to \Delta x. \tag{39.64}$$

We recall that the bulk theory has the equivalence under $R \to 1/R$ called T-duality, where momenta and winding numbers are exchanged. The above observation suggests that T-duality extends to the theories formulated on the worldsheet with boundary: *It maps the D1-brane wrapped on S^1 to the D0-brane at a point of the dual circle, where the Wilson line of the D1-brane is mapped to the position of the D0-brane.*

This can be understood if we recall that the field x and the T-dual variable $\widehat{x}$ are related by Eq. (11.83). In particular,

$$\partial_s x \to \partial_t \widehat{x} \tag{39.65}$$

so that the Neumann boundary condition $\partial_s x = 0$ is mapped to a Dirichlet boundary condition $\partial_t \widehat{x} = 0$.

Finally, this T-duality between Neumann and Dirichlet boundary conditions can also be understood from the path-integral. As in the derivation of bulk T-duality, we first consider a system of an auxiliary one-form field B and a periodic scalar field $\vartheta \equiv \vartheta + 2\pi$, both defined on the bulk of the worldsheet Σ, with the action

$$S_{\text{bulk}} = \frac{1}{2\pi}\int_\Sigma \left(\frac{R^2}{2} B \wedge *B + i d\vartheta \wedge B\right). \tag{39.66}$$

This time, however, we also introduce a boundary periodic scalar field $u \equiv u + 2\pi$, a field defined only on the boundary $\partial\Sigma$. We introduce the boundary action

$$S_{\text{boundary}} = \frac{i}{2\pi}\int_{\partial\Sigma}(a - \vartheta)du, \tag{39.67}$$

where a is a parameter. Let us first integrate out the bulk scalar field ϑ. This requires that

$$B = d\varphi \text{ on } \Sigma, \tag{39.68}$$

$$B|_{\partial\Sigma} = du, \tag{39.69}$$

where φ is a periodic scalar field (of period 2π if ϑ has period 2π). In particular, we have $du = d\varphi$, and thus we obtain the system of a single periodic variable $\varphi \equiv \varphi + 2\pi$ with the action

$$S = \frac{1}{4\pi}\int_\Sigma R^2 d\varphi \wedge *d\varphi + \frac{ia}{2\pi}\int_{\partial\Sigma} d\varphi. \tag{39.70}$$

This corresponds to a D1-brane with Wilson line a. On the other hand, let us consider integrating out the field B first. Then, we can solve for B as

$$B = -\frac{i}{R^2} * d\vartheta, \tag{39.71}$$

and if we plug this back into the action, we have

$$\widetilde{S} = \frac{1}{4\pi}\int_\Sigma \frac{1}{R^2} d\vartheta \wedge *d\vartheta + \frac{i}{2\pi}\int_{\partial\Sigma}(a - \vartheta)du. \tag{39.72}$$

After integrating out u, we obtain the boundary condition

$$\vartheta|_{\partial\Sigma} = a. \tag{39.73}$$

This corresponds to a D0-brane located at $\vartheta = a$.

39.1.4. Boundary States. Let us consider a quantum field theory formulated on a (Euclidean) Riemann surface Σ with boundary circles. We choose an orientation of each component S^1 of the boundary and we call it an *incoming* (resp. *outgoing*) component if the 90° rotation of the positive tangent vector of S^1 is an inward (outward) normal vector at the boundary. We choose the metric on Σ such that it is a flat cylinder near each boundary component. Suppose Σ has a single outgoing boundary, $S^1 = \partial\Sigma$. The partition function on Σ depends on the boundary condition a on the fields at $\partial\Sigma$ and we denote it by $Z^a(\Sigma)$. On the other hand, the path-integral over the fields on Σ defines a state $|\Sigma\rangle$ that belongs to the quantum Hilbert space $\mathcal{H}_{S^1}$ of the boundary circle. We define the *boundary state* $\langle a|$ corresponding to the boundary condition a by the property

$$Z^a(\Sigma) = \langle a|\Sigma\rangle. \tag{39.74}$$

If Σ has a single incoming boundary $\partial\Sigma = S^1$, we have a state $\langle\Sigma|$ that belongs to the dual space $\mathcal{H}_{S^1}^\dagger$. For a boundary condition b at S^1, we define the boundary state $|b\rangle$ by

$$Z_b(\Sigma) = \langle\Sigma|b\rangle, \tag{39.75}$$

where $Z_b(\Sigma)$ stands for the partition function on Σ with the boundary condition b. In general, the boundary state $\langle a|$ (resp. $|b\rangle$) does not belong to $\mathcal{H}_{S^1}^\dagger$ (resp. $\mathcal{H}_{S^1}$) but is a formal sum of elements therein. If $\partial\Sigma$ consists of several incoming components S_i^1 and outgoing components S_j^1, we have a map $f_\Sigma : \otimes_i \mathcal{H}_{S^1_i} \to \otimes_j \mathcal{H}_{S^1_j}$. Let us consider the partition function on Σ with

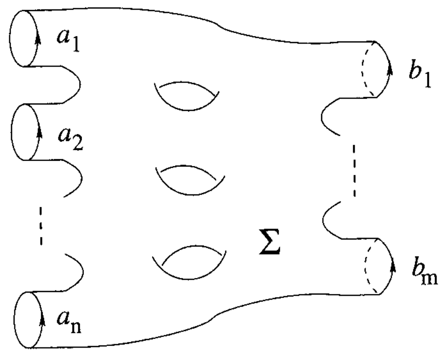


FIGURE 1. The partition function $Z^{\{a_j\}}_{\{b_i\}}(\Sigma)$

the boundary conditions $\{a_j\}$ and $\{b_i\}$ (See Fig. 1). It can be expressed using the boundary states as

$$Z^{\{a_j\}}_{\{b_i\}}(\Sigma) = \left(\bigotimes_j \langle a_j|\right) f_\Sigma \left(\bigotimes_i |b_i\rangle\right). \tag{39.76}$$

For instance, let us consider a flat finite-size cylinder Σ of length ℓ and circumference β. With a choice of orientation in the circle direction, we have one incoming boundary and one outgoing boundary. We choose the boundary conditions b and a there. Then, the partition function is given by $Z^a_b(\Sigma) = \langle a|\,\mathrm{e}^{-\ell H_c(\beta)}|b\rangle$, where $H_c(\beta)$ is the Hamiltonian of the theory on the circle of circumference β. This is the interpretation of the partition function from the closed string viewpoint. On the other hand, one can interpret it from the point of view of open strings. Let $\mathcal{H}_{ab}$ be the space of states on the interval of length ℓ with a and b as the left and the right boundary conditions and let $H_o(\ell)$ be the Hamiltonian generating the evolution in the circle direction. Thus we have

$$\mathrm{Tr}_{\mathcal{H}_{ab}}\, \mathrm{e}^{-\beta H_o(\ell)} = \langle a|\,\mathrm{e}^{-\ell H_c(\beta)}|b\rangle. \tag{39.77}$$

If the theory has spin 1/2 fermions and if the spin structure is periodic (anti-periodic) along the circle direction, the partition function is the trace

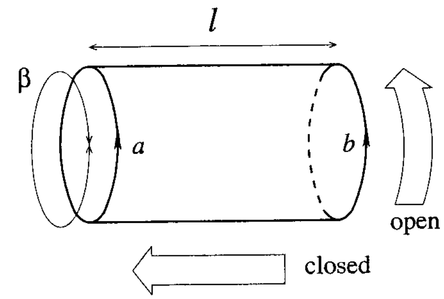


FIGURE 2. Open and closed string channels

of $(-1)^F\,\mathrm{e}^{-\beta H_o(\ell)}$ $(\mathrm{e}^{-\beta H_o(\ell)})$ over $\mathcal{H}_{ab}$. Thus, we have

$$\mathrm{Tr}_{\mathcal{H}_{ab}}(-1)^F\,\mathrm{e}^{-\beta H_o(\ell)} = {}_{\text{R-R}}\langle a|\,\mathrm{e}^{-\ell H_c(\beta)}|b\rangle_{\text{R-R}}, \tag{39.78}$$

$$\mathrm{Tr}_{\mathcal{H}_{ab}}\,\mathrm{e}^{-\beta H_o(\ell)} = {}_{\text{NS-NS}}\langle a|\,\mathrm{e}^{-\ell H_c(\beta)}|b\rangle_{\text{NS-NS}}, \tag{39.79}$$

where R-R (NS-NS) indicates that the fermions on the circle are periodic (anti-periodic).

In what follows, we determine the boundary states for the boundary conditions of free field theories specified in Sec. 39.1.1. For this, we first have to rewrite the boundary conditions on a Euclidean Riemann surface with boundary circles. Let us consider the left half-plane with the coordinates (s,t). We perform the Wick rotation $t = -i\tau$ so that we have a Euclidean left-half plane with the complex coordinate $w = s + i\tau$. If we compactify the Euclidean time direction as $\tau \equiv \tau + 2\pi$, we obtain a semi-infinite cylinder $(-\infty, 0] \times S^1$ which has a single incoming boundary if we orient the circle in the positive "time" direction. We next perform the 90° rotation of the coordinate

$$(s', \tau') = (\tau, -s), \quad \text{or} \quad w' = \mathrm{e}^{-\pi i/2} w. \tag{39.80}$$

The new coordinates s' and τ' can be regarded as the space and (Euclidean) time coordinates from the closed string point of view. A useful fact to remember is that the derivatives $\partial_\pm = \frac{1}{2}(\partial_t \pm \partial_s)$ are expressed, after Wick rotation and 90° rotation, as

$$\partial_+ \xrightarrow{\text{Wick}} \partial_{\overline{w}} \xrightarrow{90^\circ} i\partial_{\overline{w}'}, \tag{39.81}$$

$$\partial_- \xrightarrow{\text{Wick}} -\partial_w \xrightarrow{90^\circ} i\partial_{w'}. \tag{39.82}$$

Wick rotation of the boundary conditions for the free scalar theory is straightforward. The Neumann boundary condition from Eq. (39.3), which can be written as $\partial_+ x = \partial_- x$ at $s=0$, is translated to

$$\partial_{\overline{w}'} x = \partial_{w'} x \quad \text{at } \overline{w}' = w'. \tag{39.83}$$

The Dirichlet boundary condition given by Eq. (39.4) simply translates to

$$x = x_* \quad \text{at } \overline{w}' = w'. \tag{39.84}$$

Since $\partial_{\overline{w}'} + \partial_{w'}$ is tangent to the boundary, this implies $\partial_{\overline{w}'} x + \partial_{w'} x = 0$ at the boundary. For the Dirac fermion, we have to be careful since the fields transform non-trivially under coordinate change. After Wick rotation the fields ψ_- and $\overline{\psi}_-$ become independent sections $\psi_-\sqrt{dw}$ and $\overline{\psi}_-\sqrt{dw}$ of the holomorphic spinor bundle $\sqrt{K}$. Likewise, we also have independent sections $\psi_+\sqrt{d\overline{w}}$ and $\overline{\psi}_+\sqrt{d\overline{w}}$ of $\sqrt{K^*}$. After 90° rotation, $w' = \mathrm{e}^{-i\pi/2}w$, we have $\psi_-\sqrt{dw} = \psi_-\,\mathrm{e}^{i\pi/4}\sqrt{dw'}$ etc, and it is appropriate to introduce the notation

$$\psi'_- = \mathrm{e}^{i\pi/4}\psi_-, \quad \overline{\psi}'_- = \mathrm{e}^{i\pi/4}\overline{\psi}_-, \tag{39.85}$$

$$\psi'_+ = \mathrm{e}^{-i\pi/4}\psi_+, \quad \overline{\psi}'_+ = \mathrm{e}^{-i\pi/4}\overline{\psi}_+, \tag{39.86}$$

for the fields with respect to the coordinate system $(w', \overline{w}')$. We can now write down the boundary conditions in this coordinate system. The B-boundary condition as shown by Eq. (39.10) is translated to

$$\mathrm{e}^{i\pi/4}\psi'_+ = \mathrm{e}^{-i\beta}\,\mathrm{e}^{-i\pi/4}\psi'_-, \quad \mathrm{e}^{i\pi/4}\overline{\psi}'_+ = \mathrm{e}^{i\beta}\,\mathrm{e}^{-i\pi/4}\overline{\psi}'_-. \tag{39.87}$$

The A-boundary condition as shown by Eq. (39.11) is written as

$$\mathrm{e}^{i\pi/4}\psi'_+ = \mathrm{e}^{-i\alpha}\,\mathrm{e}^{-i\pi/4}\overline{\psi}'_-, \quad \mathrm{e}^{i\pi/4}\overline{\psi}'_+ = \mathrm{e}^{i\alpha}\,\mathrm{e}^{-i\pi/4}\psi'_-. \tag{39.88}$$

Without any change, the boundary conditions written above apply also to the semi-infinite cylinder $[0,\infty)\times S^1$ which has a single outgoing boundary. Moreover, they essentially apply also to any Riemann surface Σ with several incoming and outgoing boundaries. At each boundary circle S^1_i, we choose a flat complex coordinate w_i such that the boundary is given by $\mathrm{Re}(w_i)=0$ and such that Σ is in the region $\mathrm{Re}(w_i)\le 0$ (resp. $\mathrm{Re}(w_i)\ge 0$) if S^1_i is an incoming (resp. outgoing) boundary. Then, we impose the same boundary condition as above. For the free fermion system, the fields are considered as sections of the spinor bundles; $\psi_-, \overline{\psi}_- \in \Gamma(\Sigma, \sqrt{K})$ and $\psi_+, \overline{\psi}_+ \in \Gamma(\Sigma, \sqrt{K^*})$. We have a unitary isomorphism

$$\tau : \sqrt{K}|_{\partial\Sigma} \to \sqrt{K^*}|_{\partial\Sigma} \tag{39.89}$$

mapping $\sqrt{dw_i}$ to $\sqrt{d\overline{w}_i}$. [1] Then, the A- and B-boundary conditions are

$$\text{(A)} \qquad \psi_+ = \mathrm{e}^{-i\alpha}\tau\overline{\psi}_-, \quad \overline{\psi}_+ = \mathrm{e}^{i\alpha}\tau\psi_-, \tag{39.90}$$

$$\text{(B)} \qquad \psi_+ = \mathrm{e}^{-i\beta}\tau\psi_-, \quad \overline{\psi}_+ = \mathrm{e}^{i\beta}\tau\overline{\psi}_-, \tag{39.91}$$

39.1.4.1. *Massless Scalar Field.* We shall determine the boundary states $|N\rangle$ and $|D\rangle$ for the Neumann and Dirichlet boundary conditions of the free scalar field theory. One constraint for the boundary state is that the equations of the fields determining the boundary condition should hold as equations for operators acted on the boundary state. In particular, Eq. (39.83) (resp. (39.84)) should hold on the boundary state $|N\rangle$ (resp. $|D\rangle$).

We recall from Sec. 11.1 that the free scalar field formulated on the circle has the mode expansion

$$x = x_0 - \frac{1}{2}(w' - \overline{w}')p_0 + \frac{i}{\sqrt{2}}\sum_{n\neq 0}\frac{1}{n}\left(\alpha_n\,\mathrm{e}^{inw'} + \widetilde{\alpha}_n\,\mathrm{e}^{-in\overline{w}'}\right). \tag{39.92}$$

Here w' is the complex coordinate of the cylinder with $w' \equiv w' + 2\pi$ (the notation is adjusted to the discussion above). Then the Neumann boundary condition from Eq. (39.83) can be expressed in terms of the Fourier modes as

$$p_0 = 0, \quad \alpha_n + \widetilde{\alpha}_{-n} = 0 \ \ \forall n. \tag{39.93}$$

These equations must hold on the boundary state $|N\rangle$. A solution is given by

$$|N\rangle = c_N \exp\left(-\sum_{n=1}^{\infty}\frac{1}{n}\alpha_{-n}\widetilde{\alpha}_{-n}\right)|0\rangle, \tag{39.94}$$

[1] Note that there is an ambiguity in taking the square root, but the above map has an unambiguous meaning: Choose either one of the two square roots of dw_i and call it $\sqrt{dw_i}$. Then, we define $\sqrt{d\overline{w}_i}$ as the conjugate of $\sqrt{dw_i}$, where "conjugation" is the anti-linear map $\sqrt{K} \to \sqrt{K^*}$ defined over the entire Σ via the metric. The map $\sqrt{dw_i} \mapsto \sqrt{d\overline{w}_i}$ is well defined for the following reason. We first note that the holonomy of $\sqrt{K}$ is ± 1 along $\partial\Sigma$. $\sqrt{dw_i}$ is a local parallel section and therefore, if we make a tour once around S^1_i, $\sqrt{dw_i}$ comes back with the multiplication of the holonomy ± 1. Then, $\sqrt{d\overline{w}_i}$ also comes back with the same holonomy $(\pm 1)^* = \pm 1$.

where $|0\rangle$ is the ground state (with $p_0 = 0$) and c_N is a number. The boundary state $\langle N|$ for the outgoing boundary (of circumference 2π) is the conjugate state

$$\langle N| = \overline{c_N}\langle 0| \exp\left(-\sum_{n=1}^{\infty}\frac{1}{n}\alpha_n\widetilde{\alpha}_n\right). \tag{39.95}$$

Indeed, Eq. (39.93) holds when the operators act from the right. Now, let us compute the partition function on the finite size cylinder Σ. We take the cylinder to be the product of our circle S^1 of circumference 2π and the interval $[0, \pi L]$ of length πL. Then the operator f_Σ is

$$f_\Sigma = \mathrm{e}^{-\pi L H_c} \tag{39.96}$$

where H_c is the Hamiltonian shown in Eq. (11.21) for the scalar field theory formulated on our circle S^1;

$$H_c = \frac{1}{2}p_0^2 + \sum_{n=1}^{\infty}(\alpha_{-n}\alpha_n + \widetilde{\alpha}_{-n}\widetilde{\alpha}_n) - \frac{1}{12}. \tag{39.97}$$

Thus the partition function is

$$Z_N^N(\Sigma) = \langle N| q_c^{H_c} |N\rangle, \tag{39.98}$$

where we introduced the notation

$$q_c = \mathrm{e}^{-\pi L}. \tag{39.99}$$

This is evaluated as follows;

$$\begin{aligned} Z_N^N(\Sigma) &= |c_N|^2\langle 0|\,\mathrm{e}^{\sum_{n=1}\frac{-1}{n}\alpha_n\widetilde{\alpha}_n} q_c^{H_c}\,\mathrm{e}^{\sum_{n=1}^{\infty}\frac{-1}{n}\alpha_{-n}\widetilde{\alpha}_{-n}}|0\rangle \\ &= |c_N|^2 q_c^{-\frac{1}{12}}\langle 0|\,\mathrm{e}^{\sum_{n=1}\frac{-1}{n}\alpha_n\widetilde{\alpha}_n}\,\mathrm{e}^{\sum_{n=1}^{\infty}\frac{-1}{n}q_c^{2n}\alpha_{-n}\widetilde{\alpha}_{-n}}|0\rangle \\ &= |c_N|^2 V q_c^{-\frac{1}{12}}\prod_{n=1}^{\infty}(1 + q_c^2 + q_c^4 + \cdots) \\ &= |c_N|^2 V q_c^{-\frac{1}{12}}\prod_{n=1}^{\infty}\frac{1}{1-q_c^{2n}} = |c_N|^2 V\frac{1}{\eta(iL)} \end{aligned} \tag{39.100}$$

where we have used $\langle 0|0\rangle = 2\pi\delta(0) = V$. Now, we note that our cylinder is that of $(\beta, \ell) = (2\pi, \pi L)$ while the open string partition function from Eq. (39.26) can be considered as the one on the cylinder with $(\beta, \ell) = (2\pi T, \pi)$. The two cylinders are conformally equivalent if β/ℓ matches, that is, if

$$T = L^{-1}. \tag{39.101}$$

Since the scalar field theory is conformally invariant, and since our cylinders are flat, the two partition functions must agree. To examine this we use the modular transformation property in Eq. (11.58) of the eta function $1/\eta(iL) = \sqrt{L}/\eta(iL^{-1})$ from which we find, for $T = 1/L$,

$$|c_N|^2 V\frac{1}{\eta(iL)} = |c_N|^2 V\frac{1}{\sqrt{T}}\frac{1}{\eta(iT)}. \tag{39.102}$$

This is equal to the open string partition function given by Eq. (39.26) when $|c_N|^2 = \frac{1}{\sqrt{2}\,2\pi}$. Namely, we have fixed the normalization as

$$c_N = \frac{1}{\sqrt{\sqrt{2}\,2\pi}}, \tag{39.103}$$

up to a phase which can be absorbed by the redefinition of the ground state $|0\rangle$.

The above is the story for the case where x is a single-valued field. Let us now consider the case where x is a periodic scalar field of period $2\pi R$ (i.e., sigma model on the circle of radius R). The Eq. (39.93) remains the same. The momentum $p_0 = l/R$ is constrained to be zero, but the winding number $w = mR$ is not. Thus, the state $|0\rangle$ in Eq. (39.94) is replaced by a sum $\sum_{m\in\mathbb{Z}} c_m|0, m\rangle$. Moreover, when the target space is a circle, we can also consider turning on the Wilson line $(a/2\pi)\oint dx$. Since $\oint dx = 2\pi R m$ in the sector with winding number m, we find that the boundary state for the Neumann boundary condition with the Wilson line a is

$$|N_a\rangle = \sum_{m\in\mathbb{Z}} c_m\,\mathrm{e}^{-iRam}\exp\left(-\sum_{n=1}^{\infty}\frac{1}{n}\alpha_{-n}\widetilde{\alpha}_{-n}\right)|0, m\rangle. \tag{39.104}$$

Likewise, the outgoing boundary state is

$$\langle N_a| = \sum_{m\in\mathbb{Z}}\overline{c_m}\,\mathrm{e}^{iRam}\langle 0, m|\exp\left(-\sum_{n=1}^{\infty}\frac{1}{n}\alpha_n\widetilde{\alpha}_n\right). \tag{39.105}$$

The partition function on the $(\beta, \ell) = (2\pi, \pi L)$ cylinder Σ with Wilson lines a_0 and a_π at the two boundary circles is

$$\begin{aligned} Z_{N_{a_\pi}}^{N_{a_0}}(\Sigma) &= \langle N_{a_0}| q_c^{H_c} |N_{a_\pi}\rangle \\ &= \frac{1}{\eta(iL)}\sum_{m\in\mathbb{Z}}|c_m|^2 q_c^{\frac{m^2R^2}{2}}\,\mathrm{e}^{-iR\Delta a m}, \end{aligned} \tag{39.106}$$

where $\Delta a = a_\pi - a_0$. One can show that, if $|c_m|^2 = |c|^2$ for all m (i.e., independent of m), we have

$$\langle N_{a_0}|q_c^{H_c}|N_{a_\pi}\rangle = |c|^2 \frac{\sqrt{2}}{R}\frac{1}{\eta(i/L)}\sum_{l\in\mathbb{Z}} e^{-\frac{2\pi}{L}(l/R+\Delta a/2\pi)^2}. \tag{39.107}$$

Under the relation $T = 1/L$, this agrees with the open string partition function as shown by Eq. (39.30) if $|c|^2 = R/\sqrt{2}$. Thus, we have fixed the coefficients as

$$c_m = \sqrt{\frac{R}{\sqrt{2}}} \quad \forall m \in \mathbb{Z}. \tag{39.108}$$

up to phases that can be absorbed by the redefinition of the states $|0,m\rangle$.

Let us next consider the Dirichlet boundary condition given by Eq. (39.4). In terms of the oscillator modes, the condition can be written as

$$x_0 = x_*, \quad \alpha_n - \widetilde{\alpha}_{-n} = 0 \ \forall n. \tag{39.109}$$

These equations must hold when the operators are acting on the incoming boundary state $|D_{x_*}\rangle$ from the left and the outgoing boundary state $\langle D_{x_*}|$ from the right. The correctly normalized boundary states turn out to be

$$|D_{x_*}\rangle = \sqrt{\frac{2\pi}{\sqrt{2}}}\int \frac{dk}{2\pi}\exp\left(\sum_{n=1}^{\infty}\frac{1}{n}\alpha_{-n}\widetilde{\alpha}_{-n}\right) e^{ik(x_0-x_*)}|0\rangle \tag{39.110}$$

$$\langle D_{x_*}| = \sqrt{\frac{2\pi}{\sqrt{2}}}\int \frac{dk}{2\pi}\langle 0|\, e^{-ik(x_0-x_*)}\exp\left(\sum_{n=1}^{\infty}\frac{1}{n}\alpha_n\widetilde{\alpha}_n\right). \tag{39.111}$$

It is a simple exercise to show that $\langle D_{x_{*0}}|q_c^{H_c}|D_{x_{*\pi}}\rangle$ agrees with the open string partition function from Eq. (39.36) if $T = 1/L$.

For the sigma model on the circle of radius R, the correctly normalized boundary states are

$$|D_{x_*}\rangle = \sqrt{\frac{1}{R\sqrt{2}}}\sum_{l\in\mathbb{Z}}\exp\left(\sum_{n=1}^{\infty}\frac{1}{n}\alpha_{-n}\widetilde{\alpha}_{-n}\right) e^{-i(l/R)x_*}|l,0\rangle \tag{39.112}$$

$$\langle D_{x_*}| = \sqrt{\frac{1}{R\sqrt{2}}}\sum_{i\in\mathbb{Z}}\langle l,0|\, e^{i(l/R)x_*}\exp\left(\sum_{n=1}^{\infty}\frac{1}{n}\alpha_n\widetilde{\alpha}_n\right). \tag{39.113}$$

It is a simple exercise to show that $\langle D_{x_{*0}}|q_c^{H_c}|D_{x_{*\pi}}\rangle$ agrees with the open string partition function from Eq. (39.38) if $T = 1/L$.

For both sigma models on $\mathbb{R}$ and on S^1, one can show that $\langle D_{x_*}|q_c^{H_c}|N_a\rangle$ agrees with the open string partition function from Eq. (39.43) under $T = 1/L$.

39.1.4.2. *Massless Dirac Fermion.* We next determine the boundary states $|B\rangle$ and $|A\rangle$ for the B- and A-boundary conditions for the Dirac fermion. We recall that the fermion fields have the mode expansion

$$\psi'_- = \sum_{r\in\mathbb{Z}+a}\psi_r\, e^{irw'}, \quad \overline{\psi}'_- = \sum_{r'\in\mathbb{Z}-a}\overline{\psi}_{r'}\, e^{ir'w'}, \tag{39.114}$$

$$\psi'_+ = \sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}}\widetilde{\psi}_{\widetilde{r}}\, e^{-i\widetilde{r}w'}, \quad \overline{\psi}'_+ = \sum_{\widetilde{r}'\in\mathbb{Z}-\widetilde{a}}\overline{\widetilde{\psi}}_{\widetilde{r}'}\, e^{-i\widetilde{r}'w'}, \tag{39.115}$$

where again we have adjusted the notation to the present set-up. a and $\widetilde{a}$ are parameters (defined modulo $\mathbb{Z}$) that determine the periodicity along the circle. The Hamiltonian is

$$\begin{aligned} H_c = &\sum_{r\in\mathbb{Z}+a} r{:}\overline{\psi}_{-r}\psi_r{:} \\ &+\sum_{\widetilde{r}\in\mathbb{Z}+\widetilde{a}}\widetilde{r}{:}\overline{\widetilde{\psi}}_{-\widetilde{r}}\widetilde{\psi}_{\widetilde{r}}{:} - \tfrac{1}{12} + \tfrac{1}{2}(a-[a]-\tfrac{1}{2})^2 - \tfrac{1}{2}(\widetilde{a}-[\widetilde{a}]-\tfrac{1}{2})^2. \end{aligned} \tag{39.116}$$

Here : : is the normal ordering defined in Eq. (11.159), with respect to the ground state $|0\rangle_{a,\widetilde{a}}$ annihilated by ψ_r ($r \geq 0$), $\overline{\psi}_{r'}$ ($r' > 0$) $\widetilde{\psi}_{\widetilde{r}}$ ($\widetilde{r} \geq 0$) and $\overline{\widetilde{\psi}}_{\widetilde{r}'}$ ($\widetilde{r}' > 0$). It is the unique ground state if $a \notin \mathbb{Z}$ and $\widetilde{a} \notin \mathbb{Z}$. But the family of vacua $|0\rangle_{a,\widetilde{a}}$ is discontinuous at $a \in \mathbb{Z}$ or $\widetilde{a} \in \mathbb{Z}$ where two states become degenerate vacua.

Let us first consider a B-type boundary condition as shown by Eq. (39.87). It is easy to see that the condition is compatible with the above mode expansion only if $a = -\widetilde{a}$ (mod $\mathbb{Z}$). Then the condition can be expressed as

$$\psi_r = i\, e^{i\beta}\widetilde{\psi}_{-r} \quad \forall r \in \mathbb{Z}+a, \tag{39.117}$$

$$\overline{\psi}_{r'} = i\, e^{-i\beta}\overline{\widetilde{\psi}}_{-r'} \ \forall r' \in \mathbb{Z}-a. \tag{39.118}$$

These have to hold on the boundary state. A solution for the incoming boundary is

$$|B_\beta\rangle_a = \exp\left(i\, e^{-i\beta}\sum_{r'\geq 0}\psi_{-r'}\overline{\widetilde{\psi}}_{-r'} + i\, e^{i\beta}\sum_{r>0}\overline{\psi}_{-r}\widetilde{\psi}_{-r}\right)|0\rangle_{a,-a}. \tag{39.119}$$

For the outgoing boundary,

$$ {}_a\langle B_\beta| = {}_{a,-a}\langle 0|\exp\left(i\,e^{i\beta}\sum_{r'\geq 0}\widetilde{\psi}_{r'}\overline{\psi}_{r'} + i\,e^{-i\beta}\sum_{r>0}\overline{\widetilde{\psi}}_r\psi_r\right). \tag{39.120} $$

Note that Eq. (39.120) is not just the conjugate state of the state given by Eq. (39.119) but of the one with $e^{i\beta}$ replaced by $-e^{i\beta}$. This is because the orientation of the outgoing boundary is opposite from the one for the incoming boundary, and 180° rotation acts on ψ_r and $\widetilde{\psi}_{-r}$ with different signs. We have tacitly chosen the normalization, but that will be justified shortly. The family of boundary states $|B_\beta\rangle_a$ can jump by a phase at $a\in\mathbb{Z}$. This discontinuity is milder than the one for the family of ground states where $|0\rangle_{n+0,-(n+0)}$ is not even proportional to $|0\rangle_{n-0,-(n-0)}$.

Let us consider the $(\beta,\ell)=(2\pi,\pi L)$ cylinder with B_0 condition at the boundary $\mathrm{Im}(w')=\pi L$ and B_β condition at the boundary $\mathrm{Im}(w')=0$. The partition function is given by

$$ \begin{aligned} {}_a\langle B_0|q_c^{H_c}|B_\beta\rangle_a &= q_c^{-\frac{1}{12}+(a-[a]-\frac{1}{2})^2}\prod_{r'\geq 0}(1-e^{-i\beta}q_c^{2r'})\prod_{r>0}(1-e^{i\beta}q_c^{2r}) \\ &= e^{2\pi i(-\frac{\beta}{2\pi}-\frac{1}{2})(a-\frac{1}{2})}\frac{\vartheta\left[\begin{smallmatrix} a-\frac{1}{2} \\ \frac{\beta}{2\pi}+\frac{1}{2}\end{smallmatrix}\right](0,iL)}{\eta(iL)} \\ &= \frac{\vartheta\left[\begin{smallmatrix} \frac{\beta}{2\pi}-\frac{1}{2} \\ -(a-\frac{1}{2})\end{smallmatrix}\right](0,iL^{-1})}{\eta(iL^{-1})}, \end{aligned} \tag{39.121} $$

where we have assumed $0<a<1$. There is actually a discontinuity at $a\in\mathbb{Z}$, $\langle\cdots\rangle|_{a=n}=-e^{-i\beta}\langle\cdots\rangle|_{a=n+0}$, as can be seen by the first expression. This has its origin in the discontinuity in the phase of the boundary states. Now the periodicity with $a=-\widetilde{a}$, such as $\psi_\pm(w'+2\pi)=e^{2\pi i a}\psi_\pm(w')$, means that the fields obey

$$ \Psi(w'+2\pi) = e^{-2\pi i a F_V}\,\Psi(w')\,e^{2\pi i a F_V}. \tag{39.122} $$

Thus, the cylinder partition function should be equal to the open string partition function $\mathrm{Tr}\,e^{-2\pi i(a-\frac{1}{2})F_V}\,e^{-2\pi L^{-1}H}$ which was computed in Sec. 39.1.2.5. Indeed, our present result, Eq. (39.121), agrees with the open string partition function as shown by Eq. (39.57) in the range $0<a\leq 1$. This justifies our choice of the normalization of the boundary states Eqs. (39.119)-(39.120). The region $0<a\leq 1$ can be extended by making the family $|B_\beta\rangle_a$ continuous.

Let us next consider the A-boundary condition given by Eq. (39.88). It is compatible with the mode expansion with $a=\widetilde{a}$ (mod $\mathbb{Z}$). The condition for the modes is

$$ \psi_r = i\,e^{-i\alpha}\overline{\widetilde{\psi}}_{-r}\quad \forall r\in\mathbb{Z}+a, \tag{39.123} $$

$$ \overline{\psi}_{r'} = i\,e^{i\alpha}\widetilde{\psi}_{-r'}\quad \forall r'\in\mathbb{Z}-a. \tag{39.124} $$

The boundary states are

$$ |A_\alpha\rangle_a = \exp\left(i\,e^{i\alpha}\sum_{r'>0}\psi_{-r'}\widetilde{\psi}_{-r'} + i\,e^{-i\alpha}\sum_{r\geq 0}\overline{\psi}_{-r}\overline{\widetilde{\psi}}_{-r}\right)|\widetilde{0}\rangle_{a,a}, \tag{39.125} $$

$$ {}_a\langle A_\alpha| = {}_{a,a}\langle\widetilde{0}|\exp\left(i\,e^{-i\alpha}\sum_{r'>0}\widetilde{\overline{\psi}}_{r'}\overline{\psi}_{r'} + i\,e^{i\alpha}\sum_{r\geq 0}\widetilde{\psi}_r\psi_r\right). \tag{39.126} $$

Here $|\widetilde{0}\rangle_{a,a}$ is the ground state annihilated by ψ_r $(r\geq 0)$, $\overline{\psi}_{r'}$ $(r'>0)$ $\widetilde{\psi}_{\widetilde{r}}$ $(\widetilde{r}\geq 0)$ and $\overline{\widetilde{\psi}}_{\widetilde{r}'}$ $(\widetilde{r}'>0)$. (It is equal to $|0\rangle_{a,a}$ if $a\neq 0$ but is $\psi_0|0\rangle_{0,0}$ if $a=0$.) One can show that the cylinder partition function is given by

$$ {}_a\langle A_0|q_c^{H_c}|A_\alpha\rangle_a = \frac{\vartheta\left[\begin{smallmatrix} \frac{\alpha}{2\pi}-\frac{1}{2} \\ a-\frac{1}{2}\end{smallmatrix}\right](0,i/L)}{\eta(i/L)} \tag{39.127} $$

which agrees (for $T=1/L$) with the open string partition function shown in Eq. (39.63) in the range $0\leq a<1$ (which can be extended to $a\in\mathbb{R}$).

39.1.5. Boundary Entropy. Let us consider the partition function of a quantum mechanical system

$$ \mathrm{Tr}\,e^{-\beta H} = \int dE\,\rho(E)\,e^{-\beta E}, \tag{39.128} $$

where $\rho(E)$ is the density of the number of states at energy E. This formula is written as if the energy spectrum is continuous, but it applies also to the case of discrete spectrum $\rho(E)=\sum_n D_n\delta(E-E_n)$, where D_n is the number of states with $E=E_n$. In fact, we only consider this case here. The asymptotic density $\rho(E)$ at $E\to\infty$ is a measure of the number of degrees of freedom of the theory. This can be studied by looking at the behaviour of the partition function in the limit $\beta\to 0$. For instance, let us consider a conformal field theory formulated on the circle of circumference 2π. Then Eq. (39.128) can be identified as the partition function on the square torus of size $(2\pi,\beta)$. Since this torus is conformally equivalent to the square torus of size $(\frac{(2\pi)^2}{\beta},2\pi)$, the partition function can also be written as $\mathrm{Tr}\, e^{-\frac{(2\pi)^2}{\beta}H}$. Thus, the small β behavior of $\mathrm{Tr}\, e^{-\beta H}$ is determined by the large $1/\beta$ behavior of this last quantity. This in turn is determined by the low energy spectrum of the theory. The Hamiltonian of the conformally invariant field theory can be written as $H = L_0 + \widetilde{L}_0 - c/12$, where c is the central charge of the theory. Thus, if the theory has a unique ground state with $L_0 = \widetilde{L}_0 = 0$ (corresponding to the identity operator), the partition function behaves as

$$\mathrm{Tr}\, e^{-\beta H} = \mathrm{Tr}\, e^{-\frac{(2\pi)^2}{\beta}H} \sim \exp\left(\frac{(2\pi)^2}{\beta}\frac{c}{12}\right), \qquad \beta \to 0. \tag{39.129}$$

This shows that c determines the small β behaviour of $\mathrm{Tr}\, e^{-\beta H}$ or the large E behaviour of $\rho(E)$. One can in fact show from this that the density of states behaves as

$$\rho(E) \sim \frac{1}{2}\left(\frac{c}{3E^3}\right)^{\frac{1}{4}} \exp\left(2\pi\sqrt{\frac{cE}{3}}\right), \qquad E \to \infty. \tag{39.130}$$

Thus, the central charge is a measure of the number of degrees of freedom of the theory. In a sense, the central charge can be considered as the entropy of the theory. Under renormalization group flow, the central charge of the starting CFT is larger than the one for the end point (if both have a ground state with $L_0 = \widetilde{L}_0 = 0$). This is Zamolodchikov's *c-theorem.*

Now let us formulate a conformal field theory on a segment of length π, and impose conformally invariant boundary conditions a and b at the left and the right boundaries. Then $\mathrm{Tr}\, e^{-\beta H_o}$ is the partition function on a cylinder of size (π, β) (where H_o is the Hamiltonian of the system formulated on the segment). Since this cylinder is conformally equivalent to a cylinder of size $(\frac{2\pi^2}{\beta}, 2\pi)$, as we have seen, the partition function can be written as

$$\mathrm{Tr}_{ab}\, e^{-\beta H_o} = \langle a|\, e^{-\frac{2\pi^2}{\beta}H} |b\rangle. \tag{39.131}$$

Since $H = L_0 + \widetilde{L}_0 - c/12$, if the theory has a unique ground state $|0\rangle$ with $L_0 = \widetilde{L}_0 = 0$ the partition function behaves as

$$\mathrm{Tr}\, e^{-\beta H_o} \sim \langle a|0\rangle\langle 0|b\rangle\, e^{\frac{2\pi^2}{\beta}\frac{c}{12}}, \qquad \beta \to 0. \tag{39.132}$$

Here the ground state is assumed to be normalized as $\langle 0|0\rangle = 1$. The density of states now behaves as

$$\rho(E) \sim \langle a|0\rangle\langle 0|b\rangle \frac{1}{2}\left(\frac{c}{6E^3}\right)^{\frac{1}{4}} \exp\left(2\pi\sqrt{\frac{cE}{6}}\right), \qquad E \to \infty. \tag{39.133}$$

We note that there is an extra factor $\langle a|0\rangle\langle 0|b\rangle$ compared to Eqs. (39.129) or (39.130). This can be considered as the contribution of the boundaries to the number of degrees of freedom. The individual factors $\langle a|0\rangle$ and $\langle 0|b\rangle$ can be considered as the contributions of the left and the right boundaries respectively. We shall thus call

$$g_a = \langle 0|a\rangle \tag{39.134}$$

the *boundary entropy* associated with the boundary condition a. It has been conjectured that, for a given bulk CFT, the boundary entropy decreases under the renormalization group flow of the boundary conditions. This is called the (conjectural) *g-theorem*, and it has been proved for some cases.

In what follows, we will compute the boundary entropy of the boundary conditions studied in the previous subsections.

Sigma Model on S^1. Let us first consider the sigma model on the circle S^1 of radius R. The theory formulated on the circle has a unique ground state $|0,0\rangle$ with $L_0 = \widetilde{L}_0 = 0$. The boundary state for the Neumann boundary condition is obtained in Eq. (39.104) with $c_m = \sqrt{R/\sqrt{2}}$ for any m (as determined in Eq. (39.108). The boundary entropy is therefore

$$g_{N_a} = \langle 0,0|N_a\rangle = \sqrt{\frac{R}{\sqrt{2}}}. \tag{39.135}$$

The boundary state for the Dirichlet boundary condition is obtained in Eq. (39.112). Thus, the boundary entropy is

$$g_{D_{x_*}} = \langle 0,0|D_{x_*}\rangle = \sqrt{\frac{1}{R\sqrt{2}}}. \tag{39.136}$$

We note that $g_N > g_D$ if $R > 1$. Thus, if the radius R is larger than the self-dual radius, there can be an RG flow from Neumann to Dirichlet but the opposite is impossible, if the g-theorem holds. For R smaller than the self-dual radius, RG flow is possible from Dirichlet to Neumann but the opposite is impossible. This is consistent with T-duality $R \to 1/R$ where Neumann and Dirichlet boundary conditions are exchanged.

In string theory, the boundary entropy also has a meaning in space-time physics. Let us consider a string theory on $(d+1)$-dimensional space-time $\mathbb{R}^{d+1}$ times some internal space, which is abstractly given by some worldsheet quantum field theory — let us call it the internal theory. We will consider a D0-brane in $\mathbb{R}^{d+1}$ that is specified by some boundary condition a in the internal theory. Then the boundary entropy $g_a = \langle 0|a\rangle$ is proportional to the mass of the D0-brane

$$m_a = C g_a, \tag{39.137}$$

where C is a constant independent of a. To see this, we first note that the mass of an object can be measured by looking at what happens to the gravitons thrown at such an object. To be more specific, it is measured by the disc amplitude with the insertion of the graviton vertex operator — the worldsheet operator corresponding to graviton emission/absorption. See Fig. 3. On the other hand, the graviton vertex is trivial in the internal

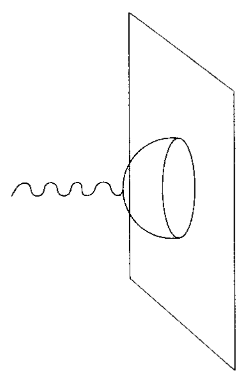


FIGURE 3. Disc diagram for graviton emission from D-brane

theory. Thus, the mass is proportional to the disc amplitude of the internal theory with no insertion of operator. This is nothing but the overlap $g_a = \langle 0|a\rangle$.

As an application, one can compute the ratio of the tension of a D1-brane and a D0-brane (or D$(p+1)$ and Dp-branes). We take the circle S^1 of radius R as the internal space, and consider particles in the $(d+1)$-dimensional space-time coming from a D1-brane wrapped on S^1 and a D0-brane at a point of S^1. The mass of the wrapped D1-brane is given by the tension of the D1-brane $T_{\rm D1}$ times the circumference $2\pi R$ of the circle, $m_{\rm D1} = 2\pi R T_{\rm D1}$. The mass of the D0-brane is its tension itself, $m_{\rm D0} = T_{\rm D0}$. On the other

hand, by the relation given by Eq. (39.137), we have

$$m_{\rm D1} = C\cdot\sqrt{\frac{R}{\sqrt{2}}},\quad m_{\rm D0} = C\cdot\sqrt{\frac{1}{R\sqrt{2}}}, \tag{39.138}$$

where we have used the result Eqs. (39.135)-(39.136). Comparing with $m_{\rm D1} = 2\pi R T_{\rm D1}$ and $m_{\rm D0} = T_{\rm D0}$, we find

$$\frac{T_{\rm D1}}{T_{\rm D0}} = \frac{1}{2\pi}. \tag{39.139}$$

Of course we also have $T_{{\rm D}(p+1)}/T_{{\rm D}p} = 1/2\pi$, or $T_{{\rm D}p} = (2\pi)^{-p}T_{\rm D0}$.

Dirac Fermion. Let us next consider the massless Dirac fermion. Since we are interested in the ordinary partition function $\mathrm{Tr}\, e^{-\beta H_o}$ (with neither $(-1)^F$ nor $e^{i\gamma F}$), we should consider an anti-periodic boundary condition for the closed string channel. This requires taking $a = \widetilde{a} = 1/2$ (NS-NS sector). There is a unique ground state $|0\rangle_{\frac{1}{2},\frac{1}{2}}$ with $L_0 = \widetilde{L}_0 = 0$ in this sector. The boundary state for the B-boundary is obtained in Eq. (39.119) and thus the boundary entropy is

$$g_{B_\beta} = {}_{\frac{1}{2},\frac{1}{2}}\langle 0|B_\beta\rangle_{\frac{1}{2}} = 1. \tag{39.140}$$

The boundary state for the A-boundary is obtained in Eq. (39.125) and the boundary entropy is

$$g_{A_\alpha} = {}_{\frac{1}{2},\frac{1}{2}}\langle 0|A_\alpha\rangle_{\frac{1}{2}} = 1. \tag{39.141}$$

39.2. Supersymmetric Boundary Conditions in $\mathcal{N}=2$ Theories

In the rest of this Chapter, we will consider theories with $(2,2)$ supersymmetry in the bulk of the worldsheet. We will focus on the boundary conditions and boundary interactions that preserve a half of the supersymmetry. As the starting point, in this section, we classify supersymmetric D-brane boundary conditions in the classical regime of non-linear sigma models and Landau–Ginzburg models.

39.2.1. A-type and B-type Supersymmetry. Recall that a $(2,2)$ supersymmetric field theory has four supercharges $Q_\pm, \overline{Q}_\pm$ that are all nilpotent and obey the anti-commutation relations $\{Q_\pm, \overline{Q}_\pm\} = H \pm P$, where the Hamiltonian H and the momentum P are the time and space translation operators. Correspondingly, there are four supercurrents $G^\mu_\pm$, $\overline{G}^\mu_\pm$ that satisfy the conservation equations $\partial_\mu G^\mu_\pm = \partial_\mu \overline{G}^\mu_\pm = 0$. The $(2,2)$ supersymmetry algebra also has two $U(1)$ R-symmetries. One is the vector

R-symmetry generated by F_V with $[F_V, Q_\pm] = -Q_\pm$, $[F_V, \overline{Q}_\pm] = Q_\pm$ while the other is the axial R-symmetry generated by F_A with $[F_A, Q_\pm] = \mp Q_\pm$, $[F_A, \overline{Q}_\pm] = \pm\overline{Q}_\pm$. Whether F_V and F_A are realized as symmetries depends on the theory.

Let us now formulate the theory on the strip $\Sigma = [0,\pi] \times \mathbb{R}$, where $[0,\pi]$ and $\mathbb{R}$ are parametrized by the space and time coordinates respectively. The presence of the boundary breaks the space translation symmetry and we lose the worldsheet momentum P. This in particular means that we cannot have all four supercharges. But we still have time translation symmetry H if we choose a time-independent boundary condition. Thus, one may hope to have some combination Q and $Q^\dagger$ so that we have a supersymmetry algebra $\{Q, Q^\dagger\} \propto H$. There are two kinds of such combinations: One is

$$Q_A = \overline{Q}_+ + e^{i\alpha} Q_-, \quad Q_A^\dagger = Q_+ + e^{-i\alpha}\overline{Q}_-. \tag{39.142}$$

The other is

$$Q_B = \overline{Q}_+ + e^{i\beta} \overline{Q}_-, \quad Q_B^\dagger = Q_+ + e^{-i\beta} Q_-. \tag{39.143}$$

Indeed, if we use the bulk supersymmetry algebra Eqs. (12.70)–(12.73) with trivial central charges $Z = \widetilde{Z} = 0$, both $Q = Q_A$ and $Q = Q_B$ obey the $(0+1)$-dimensional supersymmetry algebra

$$\begin{aligned} Q^2 = (Q^\dagger)^2 &= 0, \\ \{Q, Q^\dagger\} &= 2H. \end{aligned} \tag{39.144}$$

We call Eqs. (39.142)–(39.143) A-type and B-type supersymmetries respectively. Conservation of Q and $Q^\dagger$ means that the space components of the corresponding supercurrents vanish at the boundary:

$$\overline{G}^1_+ + e^{i\alpha} G^1_- = G^1_+ + e^{-i\alpha}\overline{G}^1_- = 0 \quad \text{for } Q = Q_A, \tag{39.145}$$

$$\overline{G}^1_+ + e^{i\beta} \overline{G}^1_- = G^1_+ + e^{-i\beta} G^1_- = 0 \quad \text{for } Q = Q_B. \tag{39.146}$$

As we have seen in free theories, a boundary condition typically relates the fermions of opposite chirality. Thus, the relation between supercurrents on the boundary should not come as a surprise.

A-type supersymmetry is compatible with the axial R-symmetry, in the sense that $[F_A, Q_A] = -Q_A$ and $[F_A, Q_A^\dagger] = Q_A^\dagger$. If the bulk theory has axial $U(1)$ R-symmetry, the theory with boundary may also have it as an R-symmetry (although it is possible that the boundary breaks it quantum

mechanically; see Sec. 39.3). Even if the bulk theory has only a discrete subgroup of the axial $U(1)$ R-symmetry, that subgroup may serve as the R-symmetry of the boundary theory. On the other hand, the vector R-rotat ion $e^{i\delta F_V}$ rotates the phase of the A-type supersymmetry as $e^{i\alpha} \to e^{i(\alpha - 2\delta)}$. In any case, the mod 2 fermion number

$$(-1)^F \propto e^{i\pi F_V} \propto e^{i\pi F_A} \tag{39.147}$$

is unbroken and preserves the phase $e^{i\alpha}$. Similarly, B-type supersymmetry is compatible with the vector R-symmetry $[F_V, Q_B] = Q_B$, $[F_V, Q_B^\dagger] = -Q_B^\dagger$. The axial R-rotatiom $e^{\delta F_A}$ rotates the phase as $e^{i\beta} \to e^{i(\beta - 2\delta)}$. The mod 2 fermion number $(-1)^F$ is always unbroken and preserves the phase $e^{i\beta}$.

We note that the combinations Eqs. (39.142)–(39.143) are the ones that become scalar supersymmetries in the A- and B-twisted topological field theories. In fact, if the bulk theory is A-twistable (i.e., the vector R-symmetry is conserved and has integral charges), one can consider A-twisting the boundary theory as well. Suppose that the anti-commutation relations as shown by Eq. (39.144) do indeed hold in the boundary theory. Then, as in Ch. 16, we obtain boundary topological field theory by declaring that the physical observables are Q_A cohomology elements (which are in one-to-one correspondence with the supersymmetric ground states of the "open string" boundary theory). A similar remark holds for $Q = Q_B$ for which we consider the B-twist.

The $(2,2)$ supersymmetry algebra has a $(1,1)$ subalgebra generated by $Q^1_\pm = Q_\pm + \overline{Q}_\pm$ that satisfy the anti-commutation relations $(Q^1_\pm)^2 = 0$ and $\{Q^1_\pm, Q^1_\pm\} = 2(H \pm P)$. Thus, a $(2,2)$ theory can be regarded as a $(1,1)$ supersymmetric theory. Then a boundary condition preserving A-type or B-type supersymmetry (with $e^{i\alpha} = e^{i\beta} = 1$) preserves a half of the $(1,1)$ supersymmetry:

$$Q^1 = Q^1_+ + Q^1_-, \tag{39.148}$$

that obeys $\{Q^1, Q^1\} = 4H$. (Even for $e^{i\alpha} \neq 1$ and $e^{i\beta} \neq 1$, this holds by the following redefinition of $Q^1_\pm$: $Q^1_\pm = e^{i\frac{\alpha}{2}} Q_\pm + e^{-i\frac{\alpha}{2}}\overline{Q}_\pm$ for A-type and $Q^1_\pm = e^{\pm i\frac{\beta}{2}} Q_\pm + e^{\mp i\frac{\beta}{2}}\overline{Q}_\pm$ for B-type.) Thus, if we want to find a boundary condition preserving A- or B-type supersymmetry, the first thing to do is to find a boundary condition preserving the $\mathcal{N}=1$ boundary supersymmetry Q^1.

39.2.2. D-Brane Boundary Condition. In this subsection, we determine the $\mathcal{N}=1$ preserving boundary condition on the fields of the supersymmetric sigma models and Landau–Ginzburg models. We will consider boundary conditions corresponding to D-branes. Since we only consider here $\mathcal{N}=1$ supersymmetry, although we will be talking about $(2,2)$ theories, the final result Eqs. (39.158)-(39.161), Eqs. (39.164)-(39.167) and Eqs. (39.169)-(39.172) hold also for more general $(1,1)$ theories such as the non-linear sigma model on a Riemannian manifold.

Let us consider a supersymmetric sigma model on a Kähler manifold X of dimension n with a superpotential W. We denote the Kähler metric with respect to local complex coordinates z^i by $g_{i\bar{j}}$. Let us write down the action of the model. If the worldsheet Σ has a boundary, one cannot ignore the total derivative terms that frequently appeared and were ignored in the discussion in Ch. 13. Thus, it is extremely important to write down the action explicitly. It is best done, in this situation , in terms of the component fields:

$$\begin{aligned} S \;=\; \int_\Sigma d^2x \Big\{ & -g_{i\bar{j}}\partial^\mu\phi^i\partial_\mu\overline{\phi^{j}} + \frac{i}{2}g_{i\bar{j}}\overline{\psi^{j}_{-}}(\overleftrightarrow{D}_0+\overleftrightarrow{D}_1)\psi^i_- \\ & +\frac{i}{2}g_{i\bar{j}}\overline{\psi^{j}_{+}}(\overleftrightarrow{D}_0-\overleftrightarrow{D}_1)\psi^i_+ - \frac{1}{4}g^{\bar{j}i}\partial_{\bar{j}}\overline{W}\partial_iW - \frac{1}{2}(D_i\partial_jW)\psi^i_+\psi^j_- \\ & -\frac{1}{2}(D_{\bar{i}}\partial_{\bar{j}}\overline{W})\overline{\psi^{i}_{-}}\,\overline{\psi^{j}_{+}} + R_{i\bar{k}j\bar{l}}\psi^i_+\psi^j_-\overline{\psi^{k}_{-}}\,\overline{\psi^{l}_{+}} \Big\}, \end{aligned} \tag{39.149}$$

where $\overline{\psi^{j}}\overleftrightarrow{D}_\mu\psi^i = \overline{\psi^{j}}(D_\mu\psi)^i - (D_\mu\overline{\psi})^{\bar{j}}\psi^i$. See Ch. 13 for other notations. If the worldsheet has no boundary, say $\Sigma=\mathbb{R}^2$, then a $(2,2)$ supersymmetry is generated by the four supercharges $Q_\pm$ and $\overline{Q}_\pm$ with the following supercurrents

$$\begin{aligned} G^0_\pm &= g_{i\bar{j}}(\partial_0\pm\partial_1)\overline{\phi^{j}}\psi^i_\pm \mp \frac{i}{2}\overline{\psi^{i}_{\mp}}\partial_{\bar{i}}\overline{W}, \\ G^1_\pm &= \mp g_{i\bar{j}}(\partial_0\pm\partial_1)\overline{\phi^{j}}\psi^i_\pm - \frac{i}{2}\overline{\psi^{i}_{\mp}}\partial_{\bar{i}}\overline{W}, \end{aligned} \tag{39.150}$$

$$\begin{aligned} \overline{G}^0_\pm &= g_{i\bar{j}}\overline{\psi^{j}_{\pm}}(\partial_0\pm\partial_1)\phi^i \pm \frac{i}{2}\psi^i_\mp\partial_iW, \\ \overline{G}^1_\pm &= \mp g_{i\bar{j}}\overline{\psi^{j}_{\pm}}(\partial_0\pm\partial_1)\phi^i + \frac{i}{2}\psi^i_\mp\partial_iW. \end{aligned} \tag{39.151}$$

Let us now formulate the theory on the left half-plane $\Sigma=(-\infty,0]\times\mathbb{R}$. We will consider the boundary condition on $\partial\Sigma=\{0\}\times\mathbb{R}$ corresponding to

a D-brane wrapped on a submanifold $\gamma\subset X$. Thus, the boundary of Σ has to be mapped into γ:

$$\phi:\partial\Sigma\longrightarrow\gamma. \tag{39.152}$$

In what follows we will determine more precise boundary conditions on the fields ϕ^i and $\psi^i_\pm,\overline{\psi^{i}_{\pm}}$ so that $\mathcal{N}=1$ supersymmetry is preserved. Since we will only consider $\mathcal{N}=1$ supersymmetry here, we will use the real notation: The fields corresponding to coordinates x^I are denoted by ϕ^I and $\psi^I_\pm$.

The first constraint comes from the equation of motion for the fields $X=(\phi^I,\psi^I_\pm)$. If we vary the fields, the variation of the action consists of a bulk term plus a boundary term;

$$\delta S = \int_\Sigma d^2x\,\delta X(\text{bulk equation of motion}) + \int_{\partial\Sigma} dt\,\delta X(\text{boundary term}), \tag{39.153}$$

where the bulk equation of motion is the same as the one in the case without boundary. We require both the bulk equation of motion and "δX (boundary term)" to vanish. This imposes the constraint on the boundary

$$\left.\begin{aligned} &g_{IJ}\delta\phi^I\partial_1\phi^J=0, \\ &g_{IJ}(\psi^I_-\delta\psi^J_- - \psi^I_+\delta\psi^J_+)=0. \end{aligned}\right\}\quad \text{on } \partial\Sigma \tag{39.154}$$

Let us look at the first equation. Since ϕ maps the boundary into γ we know that $\delta\phi^I$ is tangent to γ. Then, the first equation holds if and only if $\partial_1\phi^I$ is normal to γ. The second equation will be automatically satisfied if the fields obey the constraint from $\mathcal{N}=1$ supersymmetry which we discuss next.

The $\mathcal{N}=1$ supersymmetry transformation is given by

$$\delta\phi^I = i\epsilon(\psi^I_+ + \psi^I_-), \tag{39.155}$$

$$\delta\psi^I_\pm = -\epsilon(\partial_0\pm\partial_1)\phi^I \mp \epsilon g^{IJ}\partial_J\mathrm{Im}(W) \pm i\epsilon\Gamma^I_{JK}\psi^J_+\psi^K_-. \tag{39.156}$$

Under this transformation, the action varies as

$$\begin{aligned} \delta S = \frac{i\epsilon}{2}\int_{\partial\Sigma} dx^0\Big\{ & -g_{IJ}\partial_0\phi^I(\psi^J_- - \psi^J_+) \\ & -g_{IJ}\partial_1\phi^I(\psi^J_- + \psi^J_+) - \frac{i}{2}(\psi^I_- + \psi^I_+)\partial_I(W-\overline{W})\Big\}. \end{aligned} \tag{39.157}$$

$\psi_+^I + \psi_-^I$ at the boundary is tangent to γ since it is the (supersymmetry) variation of ϕ^I by Eq. (39.155). Then the second term of δS vanishes automatically (using $\partial_1\phi^I$ = normal to γ). The first and the third terms must vanish independently since $\partial_0\phi^I$ and $\psi_+^I + \psi_-^I$ can be independent tangent vectors of γ. This requires $\psi_-^I - \psi_+^I$ to be normal to γ and $\mathrm{Im}(W)$ to be annihilated by any tangent vector of γ. The last condition means that $\mathrm{Im}(W)$ has to be locally constant on γ. It is easy to see that the second equation of Eq. (39.154) can be satisfied by using the boundary conditions obtained so far.

To summarize, we found the following boundary conditions

(39.158) $$T_b := \partial_0\phi^I \text{ is tangent to } \gamma,$$

(39.159) $$N_b := \partial_1\phi^I \text{ is normal to } \gamma,$$

and

(39.160) $$T_f := \psi_-^I + \psi_+^I \text{ is tangent to } \gamma,$$

(39.161) $$N_f := \psi_-^I - \psi_+^I \text{ is normal to } \gamma.$$

We also found the condition on the submanifold γ itself: $\mathrm{Im}(W)$ has to be a constant on each connected component of γ.

We have shown the $\mathcal{N}=1$ supersymmetry of the action under the above boundary condition. However, for the system to be really $\mathcal{N}=1$ supersymmetric, the boundary condition itself must also be invariant under the $\mathcal{N}=1$ supersymmetry. The condition that the worldsheet boundary is mapped to γ by ϕ is certainly preserved by the supersymmetry, given Eq. (39.160). To see that the conditions Eqs. (39.160)–(39.161) themselves are also invariant, we note that

$$\delta T_f = -2\epsilon T_b, \quad \delta N_f = 2\epsilon N_b + 2\epsilon g^{IJ}\partial_J \mathrm{Im}(W),$$

in the background $\psi_\pm^I = 0$. We see that the conditions are indeed invariant provided $g^{IJ}\partial_J\mathrm{Im}(W) = 0$, which is the case if $\mathrm{Im}(W)$ is locally constant on the D-brane.

39.2.2.1. *Inclusion of B-field.* Recall that for a worldsheet without boundary one can deform the theory by a B-field which is a closed two-form on X. We can also deform the boundary theory by the B-field term

(39.162) $$\frac{1}{2}\int_\Sigma B_{IJ}d\phi^I \wedge d\phi^J.$$

The variation of the integrand is a total derivative, if the closedness $dB=0$ is used, and it is a boundary term $\int_{\partial\Sigma} dt\delta\phi^I\partial_t\phi^J B_{IJ}$. It therefore changes the first equation of Eq. (39.154) to

(39.163) $$g_{IJ}\delta\phi^I(\partial_1\phi^I + g^{IJ}B_{JK}\partial_0\phi^K) = 0,$$

although the second equation of Eq. (39.154) remains the same. This modifies the above story, and the new boundary condition is

(39.164) $$T_b := \partial_0\phi^I \text{ is tangent to } \gamma,$$

(39.165) $$N_b := \partial_1\phi^I + g^{IJ}B_{JK}\partial_0\phi^K \text{ is normal to } \gamma,$$

and

(39.166) $$T_f := \psi_-^I + \psi_+^I \text{ is tangent to } \gamma,$$

(39.167) $$N_f := (\psi_-^I - \psi_+^I) - g^{IJ}B_{JK}(\psi_-^K + \psi_+^K) \text{ is normal to } \gamma.$$

The condition that $\mathrm{Im}(W)$ is locally constant on γ remains the same.

39.2.2.2. *Gauge Field on The Brane.* Let $A_M dx^M$ be a $U(1)$ gauge field on γ, where we shall use $M, N, \ldots$ for coordinate indices on γ. One can couple the worldsheet boundary to this gauge field by the term

(39.168) $$\int_{\partial\Sigma} A_M d\phi^M = \int_{\partial\Sigma} dt\,\partial_t\phi^M A_M.$$

Its variation is given by $\int_{\partial\Sigma} dt\delta\phi^M\partial_t\phi^N F_{MN}$ where $F_{MN} = \partial_M A_N - \partial_N A_M$ is the curvature of A. The effect is identical to having the B-field $B_{MN} = F_{MN}$. Thus, the boundary condition on the fields is given by

(39.169) $$T_b := \partial_0\phi^I \text{ is tangent to } \gamma,$$

(39.170) $$N_b := \partial_1\phi^I + g_\gamma^{IM}F_{MN}\partial_0\phi^N \text{ is normal to } \gamma,$$

and

(39.171) $$T_f := \psi_-^I + \psi_+^I \text{ is tangent to } \gamma,$$

(39.172) $$N_f := (\psi_-^I - \psi_+^I) - g_\gamma^{IM}F_{MN}(\psi_-^N + \psi_+^N) \text{ is normal to } \gamma,$$

where the inclusion $T\gamma \hookrightarrow TX$ is implicit, and $g_{\gamma MN}$ is the metric induced on γ from g_{IJ}. The condition that $\mathrm{Im}(W)$ is locally constant on γ remains the same.

There is actually an alternative formulation of the boundary coupling to the gauge field on the brane. Let us consider the boundary action

$$\int_{\partial\Sigma}\mathcal{A}=\int_{\partial\Sigma}dt\,\left(\partial_0\phi^M A_M - iF_{MN}\psi^M\psi^N\right), \tag{39.173}$$

where $\psi^M=(\psi_+^M+\psi_-^M)/2$. It is easy to see that this is invariant under $\mathcal{N}=1$ supersymmetry. Thus, under the ordinary boundary condition Eqs. (39.158)–(39.161), the sum of S and Eq. (39.173) is invariant, since each of them is. Of course this boundary condition is different from the above, Eqs. (39.169)–(39.172). However, it is known that the two formulations lead to the same space-time physics. This formulation has an advantage of being easy to generalize to non-abelian gauge fields. Let $A_M dx^M$ be a non-abelian gauge field on γ with gauge group $U(r)$. Then one can consider the weight

$$P\exp\left[\,i\int_{\partial\Sigma}dt\,\left(\partial_0\phi^M A_M - iF_{MN}\psi^M\psi^N\right)\right], \tag{39.174}$$

where P stands for the path-ordered product. This is invariant under the $\mathcal{N}=1$ supersymmetry, and can be used as the supersymmetric path-integral weight.

39.2.3. A-Branes and B-Branes. We have determined the $\mathcal{N}=1$ supersymmetric boundary condition on the fields ϕ^I, $\psi_\pm^I$ corresponding to a D-brane wrapped on a submanifold γ of X (in both the absence and presence of B-field and/or gauge field on the brane). We would now like to find a condition for $\mathcal{N}=2$ supersymmetry: A-type or B-type supersymmetry. Since the boundary condition for a given brane is already fixed, the condition for $\mathcal{N}=2$ supersymmetry is imposed on the brane itself. We will call a D-brane preserving A-type (resp. B-type) supersymmetry an *A-brane* (resp. a *B-brane*).

39.2.3.1. *A-Branes.* We start with A-type supersymmetry. We first set the phase to be trivial, $e^{i\alpha}=1$. The condition is that the space component of the supercurrent vanishes at the boundary, $\overline{G}_+^1+G_-^1=0$. Using the expressions Eqs. (39.150)–(39.151), we find that

$$\begin{aligned}\overline{G}_+^1+G_-^1 &= -g_{i\bar{j}}\overline{\psi}_+^{\bar{j}}(\partial_0+\partial_1)\phi^i\\ &\quad +g_{i\bar{j}}(\partial_0-\partial_1)\overline{\phi}^{\bar{j}}\psi_-^i+\frac{i}{2}\psi_-^i\partial_i W-\frac{i}{2}\overline{\psi}_+^{\bar{j}}\partial_{\bar{j}}\overline{W}\\ &= \frac{1}{2}g(\partial_0\phi,\psi_--\psi_+)+\frac{i}{2}\omega(\partial_0\phi,\psi_-+\psi_+)\\ &\quad -\frac{1}{2}g(\partial_1\phi,\psi_-+\psi_+)-\frac{i}{2}\omega(\partial_1\phi,\psi_--\psi_+)\\ &\quad -\frac{1}{2}(\psi_-+\psi_+)^I\partial_I\mathrm{Im}(W)+\frac{i}{2}(\psi_--\psi_+)^I\partial_I\mathrm{Re}(W),\end{aligned} \tag{39.175}$$

where $g(v,w)=g_{i\bar{j}}(v^iw^{\bar{j}}+v^{\bar{j}}w^i)$ and $\omega(v,w)=ig_{i\bar{j}}(v^iw^{\bar{j}}-v^{\bar{j}}w^i)$ are the metric and the Kähler form on X. Suppose the B-field and the gauge field on the brane are set equal to zero. Then, as we have seen in Eqs. (39.158)–(39.161), at the boundary $\partial\Sigma$ the vectors $T_b=\partial_0\phi$ and $T_f=\psi_-+\psi_+$ are tangent to γ while $N_b=\partial_1\phi$ and $N_f=\psi_--\psi_+$ are normal to γ. Also, $\mathcal{N}=1$ supersymmetry requires $\mathrm{Im}(W)$ to be annihilated by a tangent vector of γ. Then, the above expression simplifies to

$$\overline{G}_+^1+G_-^1=\frac{i}{2}\omega(T_b,T_f)-\frac{i}{2}\omega(N_b,N_f)+\frac{i}{2}N_f^I\partial_I\mathrm{Re}(W). \tag{39.176}$$

The three terms must vanish individually since T_b, T_f, N_b and N_f are arbitrary tangent and normal vectors. Vanishing of $\omega(T_b,T_f)$ means that γ is isotropic with respect to ω (i.e., $\omega|_{T\gamma}=0$) while vanishing of $\omega(N_b,N_f)$ means that γ is co-isotropic (i.e., $\omega^{-1}|_{(T\gamma)^\perp}=0$, where $(T\gamma)^\perp\subset T^*X$ is the annihilator of $T\gamma$). Namely, γ is a middle-dimensional Lagrangian submanifold of X. Vanishing of the last term means that the gradient of $\mathrm{Re}(W)$ is tangent to γ. This condition holds if the gradient of $\mathrm{Im}(W)$ is normal to γ (as required by $\mathcal{N}=1$ supersymmetry) and if γ is Lagrangian, since $\mathrm{grad}[\mathrm{Re}(W)]=-J\mathrm{grad}[\mathrm{Im}(W)]$ where J is the complex structure of X.

We need to show that the boundary condition itself is invariant under the A-type supersymmetry with variation parameter $\epsilon_+=\overline{\epsilon}_-$ and $\epsilon_-=\overline{\epsilon}_+$. We first look at the boundary variation of bosonic fields:

$$\delta\phi^i=\epsilon_+\psi_-^i-\overline{\epsilon}_+\psi_+^i=\epsilon_1(\psi_-^i-\psi_+^i)+i\epsilon_2(\psi_-^i+\psi_+^i), \tag{39.177}$$

where ϵ_1 and ϵ_2 are the real and imaginary parts of ϵ_+. This shows that, for a real parameter ϵ, $\epsilon(\psi_-^i-\psi_+^i)$ and $i\epsilon(\psi_-^i+\psi_+^i)$ are the holomorphic components of tangent vectors of γ. On the other hand, $\mathcal{N}=1$ supersymmetry requires $i\epsilon(\psi_-^i-\psi_+^i)$ and $i\epsilon(\psi_-^i+\psi_+^i)$ to be the holomorphic components of normal and tangent vectors of γ respectively. Thus, $\mathcal{N}=2$ invariance of the boundary condition requires that multiplication by $i=\sqrt{-1}$ on the holomorphic components sends normal vectors to tangent vectors. This holds provided γ is a Lagrangian submanifold. Next we note

$$\delta[i\epsilon(\psi_-^i+\psi_+^i)] = 2i\epsilon\epsilon_2(\partial_0\phi^i) + 2i\epsilon\epsilon_1(i\partial_1\phi^i - \frac{1}{2}g^{i\bar{j}}\partial_{\bar{j}}\overline{W}), \tag{39.178}$$

$$\delta[\epsilon(\psi_-^i-\psi_+^i)] = 2i\epsilon\epsilon_1(\partial_0\phi^i) + 2i\epsilon\epsilon_2(i\partial_1\phi^i - \frac{1}{2}g^{i\bar{j}}\partial_{\bar{j}}\overline{W}), \tag{39.179}$$

in the background with $\psi_\pm^i=0$. For $\mathcal{N}=2$ invariance, these both must be tangent to γ. This too holds provided $g^{i\bar{j}}\partial_{\bar{j}}\overline{W}$ is the holomorphic component of a tangent vector to γ.

To summarize, when the B-field and the gauge field on the brane are zero:

A D-brane wrapped on γ is an A-brane if and only if γ is a middle dimensional Lagrangian submanifold of (X,ω) whose W-image is a straight line parallel to the real axis, which is invariant under the gradient flow of $\mathrm{Re}(W)$.

Wave-Front Trajectories. The basic example of a cycle satisfying this condition is the wave-front trajectory emanating from a critical point of the superpotential. Let $p_*\in X$ be a non-degenerate critical point of W. Let us consider the gradient vector field of the function $\mathrm{Re}(W)$ on X, which generates a one-parameter family of diffeomorphisms f_t. The maps f_t have a fixed point p_* at which the gradient is zero. We consider the collection of all gradient flow lines originating from p_*. The collection sweeps out a subset

$$\gamma_{p_*} = \left\{p\in X\,\middle|\, \lim_{t\to-\infty} f_t(p) = p_*\right\}. \tag{39.180}$$

The gradient vector of $\mathrm{Re}(W)$ annihilates $\mathrm{Im}(W)$ and therefore every point of γ_{p_*} has the same value of $\mathrm{Im}(W)$ as p_*. Thus, the image of γ_{p_*} in the W-plane is a straight line emanating from the critical value $W(p_*)$. Below, we show that γ_{p_*} is a middle-dimensional Lagrangian submanifold of (X,ω).

In a neighborhood of p_*, one can choose a complex coordinate system (z^i) such that W can be expanded as

$$W = W(p_*) + \sum_{i=1}^{n}(z_i)^2 + \cdots \tag{39.181}$$

where $+\cdots$ are cubic or higher-order terms. If the metric is the standard one $ds^2=\sum_{i=1}^n|dz_i|^2$, then the flow lines are $z_i(t)=c_i\,e^t+\cdots$ as $t\to-\infty$, in which c_i are real parameters. Thus, in a small neighborhood of p_*, the subset γ_{p_*} is a real section in the coordinate system (z_i), which is a submanifold of dimension n. (If the metric is deformed, the real section is deformed to another submanifold of dimension n.) Now, by definition, the family of maps f_t acts on γ_{p_*} with a fixed point p_*. This shows that γ_{p_*} is a submanifold of X of dimension n. To see that γ_{p_*} is a Lagrangian submanifold of X with respect to the Kähler form $\omega = ig_{i\bar{j}}dz^i\wedge d\bar{z}^{\bar{j}}$, it is crucial to note that, for $v=\mathrm{grad}[\mathrm{Re}(W)]$,

$$i_v\omega = i(g_{i\bar{j}}v^i d\bar{z}^{\bar{j}} - g_{i\bar{j}}dz^i\bar{v}^{\bar{j}}) = id(\overline{W}-W), \tag{39.182}$$

and hence

$$\mathcal{L}_v\omega = di_v\omega + i_v d\omega = 0. \tag{39.183}$$

Thus, ω is invariant under the diffeomorphisms f_t. Let V_1 and V_2 be tangent vectors of γ_{p_*} at any point. Since the Kähler form is f_t-invariant, $\omega(f_tV_1,f_tV_2) = (f_t^*\omega)(V_1,V_2)$ is independent of t. However, in the limit $t\to-\infty$, the vectors f_tV_i become the zero vector at p_*. Thus, we have shown $\omega(V_1,V_2)=0$. Namely, γ_{p_*} is Lagrangian.

It is easy to recover the phase $e^{i\alpha}$. A convenient way is to use the $U(1)_V$ R-symmetry which rotates the fermions as $\psi_\pm\to e^{i\alpha/2}\psi_\pm^i$ and the superpotential as $W\to e^{-i\alpha}W$. Identification of $\psi_\pm^i$, $\overline{\psi}_\pm^i$ and $\psi_\pm^I$ is rotated accordingly, and so is the condition on the cycle: γ has to be invariant under the gradient flow of $\mathrm{Re}(e^{-i\alpha}W)$ and its image in the W plane is a straight line in the $e^{i\alpha}$ direction. For instance γ_{p_*} is deformed so that its W-image is rotated by the angle α around the critical value $W(p_*)$.

Inclusion of B-field. Let us now turn on the B-field. This changes the boundary condition to Eqs. (39.164)–(39.167). In particular, it is $N_b=\partial_1\phi+g^{-1}B\partial_0\phi$ and $N_f=(\psi_--\psi_+)-g^{-1}B(\psi_-+\psi_+)$ that are normal to γ. It is a straghtforward computation to show

$$\begin{aligned}\overline{G}_+^1+G_-^1 &= \frac{i}{2}(\omega+B\omega^{-1}B)(T_b,T_f) - \frac{i}{2}\omega(N_b,N_f)\\ &\quad+\frac{i}{2}\omega^{-1}(gN_b,BT_f)+\frac{i}{2}\omega^{-1}(gN_f,BT_b)\\ &\quad+\frac{i}{2}N_f^I\partial_I\mathrm{Re}(W)+\frac{i}{2}(g^{-1}BT_f)^I\partial_I\mathrm{Re}(W).\end{aligned} \tag{39.184}$$

For A-type supersymmetry, each of these six terms must vanish. Vanishing of $\omega(N,N)$ means that γ is co-isotropic. In other words, the subspace $(T\gamma)^\circ$ orthogonal to $T\gamma$ with respect to ω belongs to $T\gamma$. (This can be shown by noting that $(T\gamma)^\circ = \omega^{-1}(T\gamma)^\perp$.) Vanishing of $\omega^{-1}(gN, BT)$ then means that $B=0$ on $(T\gamma)^\circ \times T\gamma$. Vanishing of $(\omega + B\omega^{-1}B)(T,T)$ implies that B is non-degenerate on $T\gamma/(T\gamma)^\circ$. Vanishing of $N\mathrm{Re}(W)$ means that the gradient of $\mathrm{Re}(W)$ is tangent to γ while vanishing of $(g^{-1}BT)\mathrm{Re}(W)$ implies (together with other conditions obtained so far) that the gradient of $\mathrm{Re}(W)$ belongs to $(T\gamma)^\circ$. However, this condition is automatically satisfied if $\mathrm{Im}(W)$ is locally constant on γ (as required by $\mathcal{N}=1$ supersymmetry), since $\mathrm{grad}[\mathrm{Re}(W)] = -J\mathrm{grad}[\mathrm{Im}(W)]$.

To summarize, the condition of $\mathcal{N}=2$ supersymmetry is

$$\begin{aligned} &(T\gamma)^\circ \subset T\gamma, \\ &B = 0 \text{ on } (T\gamma)^\circ \times T\gamma, \\ &\omega + B\omega^{-1}B = 0 \text{ on } T\gamma, \end{aligned} \tag{39.185}$$

in addition to the condition that $\mathrm{Im}(W)$ *is locally constant on* γ. It is easy to see that the D-brane boundary condition Eqs. (39.164)–(39.167) itself is invariant under the A-type $\mathcal{N}=2$ supersymmetry. In the case where γ is a middle-dimensional Lagrangian submanifold of X (where $(T\gamma)^\circ = T\gamma$), the B-field has to vanish when restricted to γ.

Coupling to the Gauge Field on the Brane. The condition of A-type supersymmetry for a D-brane supporting a $U(1)$ gauge field is obtained essentially by replacing the B-field in the above consideration by the curvature F. A point one has to be careful about is that F is now defined only on γ. The subtlety resides in the generalization of $B\omega^{-1}B$ in the third line of Eq. (39.185). However, the second line, $F=0$ on $(T\gamma)^\circ \times T\gamma$, assures that F defines a form on $T\gamma/(T\gamma)^\circ$. By definition, ω also defines a form on $T\gamma/(T\gamma)^\circ$ and is invertible there. Thus, $F\omega^{-1}F$ makes sense as a form on $T\gamma/(T\gamma)^\circ$. To summarize, the condition is stated as

$$\begin{aligned} &(T\gamma)^\circ \subset T\gamma, \\ &F = 0 \text{ on } (T\gamma)^\circ \times T\gamma, \\ &\omega + F\omega^{-1}F = 0 \text{ on } T\gamma/(T\gamma)^\circ, \end{aligned} \tag{39.186}$$

and $\mathrm{Im}(W)$ *is locally constant on* γ. In particular, if γ is a Lagrangian submanifold of X, then the gauge field on the brane has vanishing curvature $F=0$. Namely, it is flat.

39.2.3.2. *B-Branes.* We next consider B-type supersymmetry (first with the trivial phase $e^{i\beta}=1$). The condition is $\overline{G}^1_+ + \overline{G}^1_- = 0$. Using Eqs. (39.150)–(39.151), we find

$$\begin{aligned} \overline{G}^1_+ + \overline{G}^1_- &= g_{i\bar{j}}(\overline{\psi}^{\bar{j}}_- - \overline{\psi}^{\bar{j}}_+)\partial_0\phi^i - g_{i\bar{j}}(\overline{\psi}^{\bar{j}}_- + \overline{\psi}^{\bar{j}}_+)\partial_1\phi^i + \frac{i}{2}(\psi^i_- + \psi^i_+)\partial_i W \\ &= \frac{1}{2}g(\partial_0\phi, \psi_- - \psi_+) - \frac{i}{2}\omega(\partial_0\phi, \psi_- - \psi_+) \\ &\quad -\frac{1}{2}g(\partial_1\phi, \psi_- + \psi_+) + \frac{i}{2}\omega(\partial_1\phi, \psi_- + \psi_+) \\ &\quad +\frac{i}{2}(\psi_- + \psi_+)^I\partial_I(i\mathrm{Im}(W) + \mathrm{Re}(W)). \end{aligned} \tag{39.187}$$

Suppose the B-field and the gauge field on the brane are zero. Then, at the boundary $\partial\Sigma$, the vectors $T_b = \partial_0\phi$ and $T_f = \psi_- + \psi_+$ are tangent to γ while $N_b = \partial_1\phi$ and $N_f = \psi_- - \psi_+$ are normal to γ. Also, $\mathrm{Im}(W)$ is annihilated by a tangent vector to γ. Then the above expression simplifies to

$$\overline{G}^1_+ + \overline{G}^1_- = -\frac{i}{2}\omega(T_b, N_f) + \frac{i}{2}\omega(N_b, T_f) + \frac{i}{2}T_f^I\partial_I\mathrm{Re}(W). \tag{39.188}$$

For B-type supersymmetry to be conserved, all three terms must vanish. Vanishing of the first two terms means the same thing: $\omega(T,N)=0$ for any tangent and normal vectors T and N to γ. We now recall that the complex structure J is defined by $\omega g^{-1} = {}^tJ$. Therefore we have $0 = \omega(T,N) = \langle JT, gN\rangle$. Since $gN \in (T\gamma)^\perp$, this means that J sends $T\gamma$ onto itself. Namely, $T\gamma$ is a complex subspace of (TX, J). Vanishing of $T^I\partial_I\mathrm{Re}(W)$ simply means that $\mathrm{Re}(W)$ is locally constant on γ. Together with the condition from $\mathcal{N}=1$, we find that W itself is locally constant on γ.

Let us see whether the boundary condition itself is invariant under B-type $\mathcal{N}=2$ supersymmetry with the variation parameter $\epsilon_- = -\epsilon_+$ and $\bar{\epsilon}_- = -\bar{\epsilon}_+$. The bosonic fields ϕ^i transform as

$$\delta\phi^i = \epsilon_+(\psi^i_- + \psi^i_+). \tag{39.189}$$

This is indeed tangent to γ for any complex ϵ_+ if γ is a complex submanifold. The supersymmetry transformation of the tangent vector $\psi^i_- + \psi^i_+$ is $\delta(\psi^i_- + \psi^i_+) = -2i\bar{\epsilon}_+\partial_0\phi^i$ which is indeed tangent to γ. On the other hand, the normal vector $\psi_-^i - \psi_+^i$ transforms as

$$\delta(\psi_-^i - \psi_+^i) = 2i\bar{\epsilon}_+\partial_1\phi^i + \epsilon_+ g^{i\bar{j}}\partial_{\bar{j}}\overline{W}, \tag{39.190}$$

at $\psi_\pm^i = 0$. This is indeed normal to γ provided $\partial_1\phi^i$ is normal and W is locally constant on γ.

To summarize, when the B-field and the gauge field on the brane are zero:

A D-brane wrapped on γ is a B-brane if and only if γ is a complex submanifold of (X, J) and W is locally constant on γ.

B-field and Gauge Field on the Brane. If we introduce the B-field, the boundary condition is changed, and accordingly, Eq. (39.188) changes to

$$\begin{aligned}\overline{G}_+^1 + \overline{G}_-^1 = & -\frac{i}{2}\omega(T_b, N_f) + \frac{i}{2}\omega(N_b, T_f) \\ & + \frac{i}{2}(BJ + {}^tJB)(T_b, T_f) + \frac{i}{2}T_f^I\partial_I \mathrm{Re}(W).\end{aligned} \tag{39.191}$$

Thus, we simply obtain the extra condition that $BJ + {}^tJB = 0$ on $T\gamma$, or $(B|_{T\gamma})^{(2,0)} = 0$. Namely, *the restriction of B to γ has only a $(1,1)$-form component.*

The same is true for the field strength of the gauge field on the brane. Thus, B-type supersymmetry requires (in addition to the requirement that γ is a complex submanifold of X and W is constant on γ) that the curvature F has only a $(1,1)$-form component. This is equivalent to saying that the operator $\overline{\partial}_A = \overline{\partial} + A^{(0,1)}$ is nilpotent,

$$\overline{\partial}_A^2 = 0, \tag{39.192}$$

so that it defines a holomorphic structure on the associated complex line bundle.

One can examine the condition of B-type supersymmetry also by using the second formulation to couple the boundary to the gauge fields on the brane. For simplicity we present it for a $U(1)$ gauge group, but the generalization to higher ranks is straightforward. B-type supersymmetry is generated by $\overline{Q}_+ + \overline{Q}_-$ and $Q_+ + Q_-$, and we have already checked the $\mathcal{N} = 1$ supersymmetry generated by their sum. So we only have to check the invariance of the boundary coupling as shown by Eq. (39.173) by the

difference $(\overline{Q}_+ + \overline{Q}_-) - (Q_+ + Q_-)$. The corresponding variation is

$$\begin{aligned}\delta\phi^i &= 2\epsilon\psi^i, \\ \delta\psi^i &= -i\epsilon\partial_0\phi,\end{aligned}$$

where ϵ is a real anti-commuting parameter. One important point here is that B-type supersymmetry preserves the subset of the boundary fields generated by $\phi^I|_{\partial\Sigma}$ and $\psi^I = \frac{1}{2}(\psi_- + \psi_+)^I|_{\partial\Sigma}$. It is simple algebra to show that

$$\delta\int_{\partial\Sigma}\mathcal{A} = 4\epsilon\int_{\partial\Sigma}\left[F_{ij}\psi^i\partial_0\phi^j - F_{\bar{i}\bar{j}}\overline{\psi^i}\partial_0\overline{\phi^j} + i\partial_{\bar{i}}F_{jk}\overline{\psi^i}\psi^j\psi^k - i\partial_i F_{\bar{i}\bar{k}}\psi^i\overline{\psi^j}\overline{\psi^k}\right]. \tag{39.193}$$

Thus, it is invariant if and only if $F_{ij} = F_{\bar{i}\bar{j}} = 0$. Namely, the system has B-type supersymmetry if and only if the curvature F has only a $(1,1)$-form component.

39.2.4. Appendix. Here we point out a subtlety in using the supercurrent when determining the supersymmetric boundary condition. Since this subtlety exists (and the resolution holds) for an arbitrary symmetry, we describe it in general terms. Suppose a system of variable X with Lagrangian $\mathcal{L}(X, \partial_\mu X)$, formulated on a space without a boundary, has a symmetry under $X \to X + \epsilon f(X)$. This means that the action $S = \int d^2x\mathcal{L}$ is invariant, that is, for a position dependent $\epsilon(x)$, we have

$$\delta\mathcal{L} = \partial_\mu\epsilon\, a^\mu + \epsilon\,\partial_\mu b^\mu. \tag{39.194}$$

The conserved current of the system is

$$j^\mu = b^\mu - a^\mu. \tag{39.195}$$

Now let us formulate the system on a strip that has a timelike boundary $\partial\Sigma$. Naïvely, the conservation of the charge $Q = \int dx^1\, j^0$ holds under the condition

$$j^1 = 0 \ \text{ at } \partial\Sigma. \tag{39.196}$$

However, for a constant ϵ, the action varies as

$$\delta S = \int_\Sigma d^2x\,\epsilon\partial_\mu b^\mu = \int_{\partial\Sigma} dx^0\,\epsilon b^1. \tag{39.197}$$

Thus, the condition for the symmetry of the system is

$$b^1 = 0 \ \text{ at } \partial\Sigma. \tag{39.198}$$

Thus, the two conditions differ by $b^1 - j^1 = a^1$. For instance, let us consider the system $\mathcal{L} = -\frac{1}{2}\partial_\mu X \partial^\mu X$ which has a symmetry under $X \to X + \epsilon$. For this symmetry, we find $a^\mu = -\partial^\mu X$ and $b^\mu = 0$, and thus $j^\mu = \partial^\mu X$ (leading to the "target space momentum" $p = \int dx^1\, \partial^0 X$). If it is formulated on a strip, the charge conservation Eq. (39.196) is non-trivial, $\partial_1 X = 0$ at $\partial\Sigma$, but the symmetry condition given by Eq. (39.198) is trivial, $0 = 0$. What does this mean? For example, for Dirichlet boundary condition where $\partial_1 X$ is not necessarily zero, the charge is not conserved but the symmetry is preserved? Of course not! The point is that the Dirichlet boundary condition itself, $X =$ constant at $\partial\Sigma$, is not invariant under the translation $X \to X + \epsilon$. This example reminds us that a symmetry should preserve *both* the allowed field configuration space *and* the action. Thus, the symmetry is preserved in the boundary theory if

(i) It preserves the boundary condition.
(ii) $j^1 = 0$ at $\partial\Sigma$.
(iii) $a^1 = 0$ at $\partial\Sigma$.

Are they all independent? We claim that this is not the case, when the boundary condition is chosen so that it is compatible with the bulk equation of motion. Let us consider the variation $\delta X = \epsilon f(X)$ where $\epsilon = \epsilon(x)$ is position-dependent. Then, the action varies as

$$\delta S = \int_\Sigma d^2x\, \epsilon\, \partial_\mu j^\mu + \int_{\partial\Sigma} dx^0\, \epsilon\, a^1. \tag{39.199}$$

Under the (compatible) boundary condition that is preserved by the transformation $\delta X = \epsilon f(X)$, this has to vanish if the bulk equation of motion is satisfied. Since $\partial_\mu j^\mu = 0$ if the bulk equation of motion holds, this means a^1 has to automatically vanish on the boundary. This argument shows that it is enough to require (i) and (ii).

For completeness, we record here the expressions of ϵj^1 and ϵa^1 for the $(2,2)$ supersymmetry on the system considered in this section. The supersymmetry transformation of the fields $\delta X = \epsilon f(X)$ is

$$\delta\phi^i = \epsilon_+\psi_-^i - \epsilon_-\psi_+^i, \tag{39.200}$$

$$\delta\psi_+^i = i\overline{\epsilon}_-(\partial_0 + \partial_1)\phi^i + \epsilon_+ F^i, \tag{39.201}$$

$$\delta\psi_-^i = -i\overline{\epsilon}_+(\partial_0 - \partial_1)\phi^i + \epsilon_- F^i, \tag{39.202}$$

where

$$F^i = -\frac{1}{2}g^{i\overline{j}}\partial_{\overline{j}}\overline{W} + \Gamma^i_{jk}\psi_+^j\psi_-^k. \tag{39.203}$$

The action varies as

$$\begin{aligned}\delta S = \frac{1}{2}\int_{\partial\Sigma} dx^0 \Big\{ &\epsilon_+\left(-g_{i\overline{j}}(\partial_0+\partial_1)\overline{\phi}^{\overline{j}}\psi_-^i + \frac{i}{2}\overline{\psi}_+^{\overline{i}}\partial_{\overline{i}}\overline{W}\right) \\ &+\epsilon_-\left(-g_{i\overline{j}}(\partial_0-\partial_1)\overline{\phi}^{\overline{j}}\psi_+^i - \frac{i}{2}\overline{\psi}_-^{\overline{i}}\partial_{\overline{i}}\overline{W}\right) \\ &+\overline{\epsilon}_+\left(g_{i\overline{j}}\overline{\psi}_-^{\overline{j}}(\partial_0+\partial_1)\phi^i + \frac{i}{2}\psi_+^i\partial_i W\right) \\ &+\overline{\epsilon}_-\left(g_{i\overline{j}}\overline{\psi}_+^{\overline{j}}(\partial_0-\partial_1)\phi^i - \frac{i}{2}\psi_-^i\partial_i W\right)\Big\}.\end{aligned} \tag{39.204}$$

This shows what ϵb^1 is. This decomposes into $\epsilon(j^1 + a^1)$ where

$$\epsilon j^1 = \epsilon_- G_+^1 - \epsilon_+ G_-^1 + \overline{\epsilon}_+\overline{G}_-^1 - \overline{\epsilon}_-\overline{G}_+^1, \tag{39.205}$$

and

$$\begin{aligned}\epsilon a^1 = \;&\epsilon_-\left[\frac{1}{2}g_{i\overline{j}}(\partial_0+3\partial_1)\overline{\phi}^{\overline{j}}\psi_+^i + \frac{i}{4}\overline{\psi}_-^{\overline{j}}\partial_{\overline{j}}\overline{W}\right] \\ &+\epsilon_+\left[\frac{1}{2}g_{i\overline{j}}(\partial_0-3\partial_1)\overline{\phi}^{\overline{j}}\psi_-^i - \frac{i}{4}\overline{\psi}_+^{\overline{j}}\partial_{\overline{j}}\overline{W}\right] \\ &-\overline{\epsilon}_+\left[\frac{1}{2}g_{i\overline{j}}\overline{\psi}_-^{\overline{j}}(\partial_0-3\partial_1)\phi^i + \frac{i}{4}\psi_+^i\partial_i W\right] \\ &-\overline{\epsilon}_-\left[\frac{1}{2}g_{i\overline{j}}\overline{\psi}_+^{\overline{j}}(\partial_0+3\partial_1)\phi^i - \frac{i}{4}\psi_-^i\partial_i W\right].\end{aligned} \tag{39.206}$$

A-type (resp. B-type) supersymmetry generates the transformation with $\epsilon_+ = \overline{\epsilon}_-$ (resp. $\epsilon_+ = -\epsilon_-$). In the main text, we have found the condition for A-branes (resp. B-branes) so that $\epsilon j^1 = 0$ holds, and we have also checked the invariance of the boundary condition itself. Thus, ϵa^1 will also vanish under the same boundary condition. It is also straightforward to check this explicitly.

39.3. R-Anomaly

In Ch. 13, we saw that the axial $U(1)$ R-symmetry of the (bulk) nonlinear sigma model on X can be anomalous depending on the first Chern class $c_1(X)$. It is anomaly-free if $c_1(X) = 0$, that is, if X is a Calabi–Yau manifold. Otherwise it is broken to a discrete subgroup including $\mathbb{Z}_2$; if the greatest common divisor of $\int_\beta c_1(X)$, $\beta \in H^2(X,\mathbb{Z})$, is N, it is broken down to $\mathbb{Z}_{2N}$. We would like to ask whether the R-symmetry can be anomalous for the boundary theory. Of course, the bulk anomaly persists and the boundary R-symmetry cannot be larger than than that of the bulk. The question is whether the bulk R-symmetry can be further broken by the boundary conditions or boundary interactions. At the classical level, the B-type boundary condition breaks the axial $U(1)$ R-symmetry to $\mathbb{Z}_2$, while the A-type boundary condition preserves it. Thus, this question is non-trivial for A-branes. In this section, we show that the boundary can break the bulk R-symmetry to a smaller subgroup. We find that a topological invariant called the *Maslov index* plays a role analogous to the first Chern class in the bulk theory. We will also determine the condition for the full $U(1)$ axial R-symmetry to be unbroken. An example is the D-brane wrapped on a special Lagrangian submanifold of a Calabi–Yau manifold.

39.3.1. Fermionic Zero Modes. As in the bulk theory, it is useful to count the number of fermionic zero modes in a given bosonic background. Here we consider the problem in a toy model which is motivated by the theory of an open string ending on an A-brane.

Let Σ be a Riemann surface with boundary to which we give a metric as described in Sec. 39.1.4. At the boundary, we have a unitary isomorphism between the canonical bundle and its dual,

$$\tau : K|_{\partial\Sigma} \to K^*|_{\partial\Sigma}, \tag{39.207}$$

that identifies the parallel frames of the Levi–Civita connections (which is possible since the holonomy is trivial along $\partial\Sigma$ under the above choice of metric on Σ). Let E be a rank n complex vector bundle on Σ with a Hermitian metric h and a Hermitian connection A. Suppose we have a unitary isomorphism between E and its dual at the boundary

$$\tau : E|_{\partial\Sigma} \to E^*|_{\partial\Sigma}. \tag{39.208}$$

We consider a free fermion system

$$\psi_- \in \Gamma(\Sigma, E), \quad \overline{\psi}_- \in \Gamma(\Sigma, E^*\otimes K), \tag{39.209}$$

$$\overline{\psi}_+ \in \Gamma(\Sigma, E^*), \quad \psi_+ \in \Gamma(\Sigma, E\otimes K^*), \tag{39.210}$$

with the action

$$S = \frac{i}{2\pi}\int_{\Sigma} \left(\overline{\psi}_-\overline{D}\psi_- + \psi_+ D\overline{\psi}_+\right) \tag{39.211}$$

where D and $\overline{D}$ are the $(1,0)$ and $(0,1)$ components of the covariant derivative (of E^* and E respectively). We impose the boundary condition

$$\psi_- = \overline{\psi}_+, \quad \overline{\psi}_- = \psi_+, \quad \text{on } \partial\Sigma. \tag{39.212}$$

Implicit here are the identifications $E|_{\partial\Sigma} \cong E^*|_{\partial\Sigma}$ and $(E^*\otimes K)|_{\partial\Sigma} \cong (E\otimes K^*)|_{\partial\Sigma}$ that are induced from Eqs. (39.207)–(39.208). We would like to compute the index of the twisted Dirac operator $\mathcal{D}$;

$$\mathrm{Index}\,\mathcal{D} = \#[(\psi_-, \overline{\psi}_+)\text{ zero modes}] - \#[(\overline{\psi}_-, \psi_+)\text{ zero modes}]. \tag{39.213}$$

The essential technique in computing the index is the doubling of the Riemann surface Σ together with the Hermitian vector bundle with connection (E, A). Since the metric of Σ around each component of $\partial\Sigma$ is that of a flat cylinder, without any effort one can glue Σ and its orientation reversal Σ^* along $\partial\Sigma$. We denote by $\Sigma\#\Sigma^*$ the resulting closed Riemann surface (with metric). By using the isomorphism $K|_{\partial\Sigma} \cong K^*|_{\partial\Sigma}$ shown in Eq. (39.207), one can glue K over Σ and K^* over Σ^*, along with the Levi–Civita connection. It is easy to show that this leads to the canonical bundle $K_{\Sigma\#\Sigma^*}$ with the Levi–Civita connection of the double. Since we have a unitary isomorphism $E|_{\partial\Sigma} \cong E^*|_{\partial\Sigma}$ as shown by Eq. (39.208), one can also glue (E, h) over Σ and (E^*, h^*) over Σ^* topologically. In general, however, this does not define a smooth Hermitian vector bundle with metric on $\Sigma\#\Sigma^*$, and the connections do not even glue continuously (e.g., the holonomies along $\partial\Sigma$ do not match). At this stage, we take advantage of the deformation invariance of the index: As far as the computation of the index is concerned, one can deform the metric h and the connection A. Suppose one can change them so that (E, h, A) over Σ and (E^*, h^*, A^*) over Σ^* smoothly glue along $\partial\Sigma$. (We will shortly exhibit how this can be done for a class of bundles with the isomorphism given by Eq. (39.208).) We denote the resulting Hermitian bundle with a connection by $(E\#E^*, h\#h^*, A\#A^*)$, or simply by $E\#E^*$.

We will consider ψ_- and $\overline{\psi}_-$ as fields on Σ and ψ_+ and $\overline{\psi}_+$ as fields on Σ^*. By the boundary condition shown in Eq. (39.212), ψ_- on Σ and $\overline{\psi}_+$ on Σ^* continuously glue along $\partial\Sigma$. Thus they define a continuous section of $E\#E^*$, which we denote by $\psi_-\#\overline{\psi}_+$. On the other hand, $\overline{\psi}_-$ and ψ_+ define a continuous section $\overline{\psi}_-\#\psi_+$ of $(E\#E^*)^*\otimes K_{\Sigma\#\Sigma^*}$. If $\overline{D}\psi_- = 0$, $\psi_-\#\overline{\psi}_+$ is obviously holomorphic on $\Sigma\subset\Sigma\#\Sigma^*$. If $D\overline{\psi}_+ = 0$, $\overline{\psi}_+$ is holomorphic on Σ^* because of the flip of the orientation, and therefore $\psi_-\#\overline{\psi}_+$ is holomorphic on $\Sigma^*\subset\Sigma\#\Sigma^*$. Thus, if $\overline{D}\psi_- = D\overline{\psi}_+ = 0$, $\psi_-\#\overline{\psi}_+$ is entirely holomorphic on $\Sigma\#\Sigma^*$ (in particular ψ_- and $\overline{\psi}_+$ holomorphically glue along $\partial\Sigma$). Likewise, if $\overline{D\psi}_- = D\psi_+ = 0$, $\overline{\psi}_-\#\psi_+$ is entirely holomorphic on $\Sigma\#\Sigma^*$.

This shows that the index of $\mathcal{D}$ is the index of the Dolbeault operator of $E\#E^*$;

$$\begin{aligned} \operatorname{Index}\mathcal{D} &= \dim H^0(\Sigma\#\Sigma^*, E\#E^*) - \dim H^0(\Sigma\#\Sigma^*, (E\#E^*)^*\otimes K) \\ &= \operatorname{Ind}\left(\overline{\partial}_{A\#A^*} : \Omega^{0,0}(\Sigma\#\Sigma^*, E\#E^*) \to \Omega^{0,1}(\Sigma\#\Sigma^*, E\#E^*)\right) \\ &= c_1(E\#E^*) + \operatorname{rank} E(1 - g_{\Sigma\#\Sigma^*}). \end{aligned} \tag{39.214}$$

Here $g_{\Sigma\#\Sigma^*}$ is the genus of the double $\Sigma\#\Sigma^*$, which is $2g+h-1$ if Σ has genus g and h boundary circles.

Now, we focus on more specific models of bundles with the boundary isomorphisms given by Eq. (39.208). We assume that the restriction of E to the boundary has an orthogonal decomposition (over $\mathbb{R}$)

$$E|_{\partial\Sigma} \cong E_{\mathbb{R}} \oplus iE_{\mathbb{R}} \tag{39.215}$$

where $E_{\mathbb{R}}$ is a real subbundle of $E|_{\partial\Sigma}$. We denote by $E^*_{\mathbb{R}}$ the image of $E_{\mathbb{R}}$ under the complex conjugation $\sigma : E \to E^*$ with respect to the Hermitian metric (i.e., $\langle\sigma(v), w\rangle := h(v,w)$). Then $\tau(v) := \sigma(v)$ and $\tau(iv) := -\sigma(iv) = i\sigma(v)$ for $v \in E_{\mathbb{R}}$ defines a unitary isomorphism shown in Eq. (39.208) at the boundary. Now, what is $c_1(E\#E^*)$ for such a τ? We note here that one can trivialize E over Σ as long as the boundary is non-empty. We choose and fix a unitary trivialization $E \cong \Sigma\times\mathbb{C}^n$. A subbundle $E_{\mathbb{R}}$ giving a decomposition shown in Eq. (39.215) is provided by a map $t \in \partial\Sigma \mapsto [u_t] \in U(n)/O(n)$ (where u_t is a representative in $U(n)$ (possibly multivalued)); the fibre of $E_{\mathbb{R}}$ at t is given by $u_t(\mathbb{R}^n) \subset \mathbb{C}^n$. Assuming the dual trivialization $E^* \cong \Sigma\times\mathbb{C}^n$, the Hermitian conjugation $\sigma : E \to E^*$ is given by the simple complex conjugation $^- : \mathbb{C}^n \to \mathbb{C}^n$. Thus, the subbundle $E^*_{\mathbb{R}}$ is associated with the map $t \in \partial\Sigma \mapsto [\overline{u_t}] \in U(n)/O(n)$. In the construction of $E\#E^*$, we identify $u_t(v)$ with $\overline{u_t}(v)$ for $v \in \mathbb{R}^n$. Therefore the transition function of $E\#E^*$ at $\partial\Sigma$ is given by $t \mapsto g_t$ where

$$g_t := u_t\overline{u_t}^{-1} = u_t u_t^T, \tag{39.216}$$

which is single-valued since $u_t o(u_t o)^T = u_t o o^T u_t^T = u_t u_t^T$ for $o \in O(n)$. Thus, the first Chern class $c_1(E\#E^*)$ is given by the winding number of $t \mapsto \det g_t = \det^2 u_t$. The latter is a topological invariant called the *Maslov index* of the pair $(E, E_{\mathbb{R}})$, denoted by $\mu(E, E_{\mathbb{R}})$. We therefore find

$$c_1(E\#E^*) = \mu(E, E_{\mathbb{R}}). \tag{39.217}$$

Now as promised, we construct a connection A of E so that (E, A) on Σ and (E^*, A^*) on Σ^* glues smoothly. We extend $t \to u_t$ to a (multi-valued) function u defined on a neighborhood U of $\partial\Sigma = \Sigma\cap\Sigma^*$ in $\Sigma\#\Sigma^*$. Then the transition function $g = uu^T : U \to U(n)$ provides a smooth structure on $E\#E^*$. Then, we set

$$A = udu^{-1} \quad \text{on } U, \tag{39.218}$$

and extend it to the interior using a bump function that is identically 1 on U but decays quickly to 0. This yields $A^* = \overline{u}d\overline{u}^{-1}$ on U. It is easy to check that $A^* = g^{-1}Ag + g^{-1}dg$ on U, and therefore the two connections glue smoothly.

Thus, for this class of models, we find that the index is given by Eq. (39.214), or

$$\operatorname{Index}\mathcal{D} = \mu(E, E_{\mathbb{R}}) + (\operatorname{rank} E)(2 - 2g - h), \tag{39.219}$$

where $\mu(E, E_{\mathbb{R}})$ is the Maslov index of the pair $(E, E_{\mathbb{R}})$, g is the genus of Σ and h is the number of boundary components of Σ.

39.3.2. Axial Anomaly for A-Branes. Let us now consider the actual objects of interest — A-branes. To be specific we consider a D-brane wrapped on a Lagrangian submanifold L of a Kähler manifold X. We fix a bosonic background of the open string worldsheet Σ: a map ϕ into X sending the boundary to L;

$$\phi : (\Sigma, \partial\Sigma) \to (X, L). \tag{39.220}$$

The worldsheet fermions are spinors that take values in the pulled-back tangent bundle; $\Psi_\pm \in \Gamma(\Sigma, \phi^*TX \otimes S_\pm)$. Here $S_\pm$ are spin bundles which could also be denoted as $S_- = \sqrt{K}$ and $S_+ = \sqrt{K}^*$. The boundary condition for $\Psi_\pm$ is given in Eqs. (39.160)–(39.161) for the flat Minkowski space: $\Psi_- + \Psi_+$ is tangent to L while $\Psi_- - \Psi_+$ is normal to L. ($\psi^I_\pm$ there is denoted by $\Psi^I_\pm$ here.) This can be restated on the Euclidean worldsheet Σ as follows. First, on the boundary $\partial\Sigma$, we identify the spin bundles of opposite chirality

$$\tau : \sqrt{K}|_{\partial\Sigma} \to \sqrt{K}^*|_{\partial\Sigma}, \tag{39.221}$$

by identifying parallel sections with respect to the Levi–Civita connection. (Again, this is possible since the holonomies of the two bundles are the same, $(\pm 1) = (\pm 1)^*$, for our choice of metric. For more details, see Sec. 39.1.4.) Next we introduce a map $\tau : TM|_L \to TM|_L$ which is the identity on the tangent bundle of L and is $(-1)\times$ identity on the normal bundle of L. Then the boundary condition is simply written as

$$\tau(\Psi_-) = \Psi_+. \tag{39.222}$$

Now, we decompose $\Psi_\pm$ into $\psi_\pm$ and $\overline{\psi}_\pm$ which take values in $\phi^*T^{(1,0)}X$ and $\phi^*T^{(0,1)}X$ respectively. We note here that τ acts linearly on the complexification $TX_L \otimes \mathbb{C} = T^{(1,0)}X|_L \oplus T^{(0,1)}X|_L$, exchanging the $(1,0)$ and $(0,1)$ components. This is because if $t \in TL$ then $Jt \in NL$, and thus $\tau : (t - iJt) \in T^{(1,0)}X|_L \mapsto (t + iJt) \in T^{(0,1)}X|_L$. Thus, we have

$$\tau : \phi^*T^{(1,0)}X|_{\partial\Sigma} \to \phi^*T^{(0,1)}X|_{\partial\Sigma}. \tag{39.223}$$

The above boundary condition can then be restated as

$$\tau(\psi_-) = \overline{\psi}_+, \quad \tau(\overline{\psi}_-) = \psi_+. \tag{39.224}$$

This is exactly the boundary condition of the type shown in Eq. (39.212) where E is given by $\phi^*T^{(1,0)}X \otimes \sqrt{K}$ and $\tau : E|_{\partial\Sigma} \to E^*|_{\partial\Sigma}$ is induced from Eqs. (39.221)–(39.223). Furthermore, the linear map τ is the one associated with the orthogonal decomposition $E|_{\partial\Sigma} = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ resulting from the decompositions

$$T^{(1,0)}X|_L = [TL]^{(1,0)} \oplus i[TL]^{(1,0)}, \tag{39.225}$$

$$\sqrt{K}\Big|_{\partial\Sigma} = \sqrt{K}_{\mathbb{R}} \oplus i\sqrt{K}_{\mathbb{R}}. \tag{39.226}$$

The first one is obtained from $TX|_L = TL \oplus NL$ by the map $(1 - iJ) : TX \to T^{(1,0)}X$, while $\sqrt{K}_{\mathbb{R}}$ is spanned (over $\mathbb{R}$) by a parallel frame of $\sqrt{K}$ on $\partial\Sigma$ (which is single-valued up to ± 1 multiplication). To be explicit, $E_{\mathbb{R}} = \phi^*[TL]^{(1,0)} \otimes_{\mathbb{R}} \sqrt{K}_{\mathbb{R}}$.

Thus, we can apply the formula Eq. (39.219) or Eq. (39.214) to compute the index of the Dirac operator $\mathcal{D}$. It is easy to see that the resulting

$\sqrt{K}\#\sqrt{K}^*$ is a spinor bundle $\sqrt{K_{\Sigma\#\Sigma^*}}$ of the double. Now, we have

$$\begin{aligned} c_1(E\#E^*) &= c_1\left((\phi^*T^{(1,0)}X\#\phi^*T^{(0,1)}X) \otimes \sqrt{K_{\Sigma\#\Sigma^*}}\right) \\ &= c_1\left(\phi^*T^{(1,0)}X\#\phi^*T^{(0,1)}X\right) \\ &\qquad + \operatorname{rank}\left(\phi^*T^{(1,0)}X\right)\, c_1\left(\sqrt{K_{\Sigma\#\Sigma^*}}\right) \\ &= \mu\left(\phi^*T^{(1,0)}X, \phi^*[TL]^{(1,0)}\right) + \operatorname{rank} E\,(g_{\Sigma\#\Sigma^*} - 1). \end{aligned} \tag{39.227}$$

Thus we find

$$\operatorname{Index}\mathcal{D} = \mu\left(\phi^*T^{(1,0)}X, \phi^*[TL]^{(1,0)}\right). \tag{39.228}$$

Suppose Σ has a disc topology and let us fix a trivialization $\phi^*TX \cong \Sigma \times \mathbb{C}^n$. Then the sub-bundle ϕ^*TL on $\partial\Sigma = S^1$ defines a loop in the Lagrangian Grassmannian $\Lambda(\mathbb{C}^n)$, the space of Lagrangian subspaces of $\mathbb{C}^n$. The Maslov index $\mu(\phi^*T^{(1,0)}X, \phi^*[TL]^{(1,0)})$ is clearly the Maslov index of the loop. (See Appendix of this section for the definition and topology of $\Lambda(\mathbb{C}^n)$ and the Maslov index of loops therein.) More generally if Σ has h boundary components, we have h loops in $\Lambda(\mathbb{C}^n)$. The index $\mu(\phi^*T^{(1,0)}X, \phi^*[TL]^{(1,0)})$ is just the sum of the Maslov indices of the loops.

If this Maslov index is nonzero for some map $\phi : (\Sigma, \partial\Sigma) \to (X, L)$ the axial $U(1)$ R-symmetry is anomalously broken. If X is not Calabi–Yau, the axial $U(1)$ R-symmetry of the bulk theory is anomalous. This is also true of the boundary theory which can be explicitly seen as follows. Let us choose a map ϕ_1 from a closed Riemann surface Σ_1 to X such that $\int_{\Sigma_1} \phi_1^* c_1(X) \neq 0$. Then, "composition" of this map with any map $\phi : (\Sigma, \partial\Sigma) \to (X, L)$ increases the index by $2\int_{\Sigma_1} \phi_1^* c_1(X)$, as can be seen by the formula Eq. (39.214). Thus, if the R-symmetry is broken to $\mathbb{Z}_{2N}$ in the bulk theory, the axial R-symmetry of the boundary theory is not larger than $\mathbb{Z}_{2N}$.

One thing we would like to have in the boundary theory is the distinction between bosonic and fermionic states. This is possible only if at least the $\mathbb{Z}_2$ subgroup of the R-symmetry group is unbroken. This is assured if the Dirac index is always an even integer. This condition is satisfied if the Lagrangian submanifold is oriented, so that the Maslov index is twice the oriented Maslov index and is necessarily even. In what follows, we will only consider oriented Lagrangians.

39.3.2.1. *Anomaly-Free Condition.* If X is a Calabi–Yau manifold, the bulk axial $U(1)$ R-symmetry is not broken. Under what condition is the boundary R-symmetry unbroken as well? On a Calabi–Yau n-fold there is a nowhere vanishing holomorphic n-form Ω. Let us put $\Omega|_L = c\cdot \mathrm{vol}(L)$ where $\mathrm{vol}(L)$ is the volume form on L. It is easy to see that c is nowhere vanishing if L is Lagrangian. Thus, we have a map

$$c : L \to \mathbb{C}^{\times}. \tag{39.229}$$

The holomorphic n-form defines an isomorphism $\wedge^n T_x^{(1,0)} X \to \mathbb{C}$ at each point x of X and determines a trivialization of the determinant bundle $\det(T^{(1,0)}X)$. If we have a map $\phi : (\Sigma, \partial\Sigma) \to (X, L)$, $\phi^*\Omega$ trivializes $\det(\phi^* T^{(1,0)} X)$, and the Maslov index of the real subbundle $\phi^*[TL]^{(1,0)}$ is measured by the function $\phi^* c^2$. Namely

$$\mu(\phi^* T^{(1,0)} X, \phi^*[TL]^{(1,0)}) = 2 \times \text{winding number of } (\phi^* c : \partial\Sigma \to \mathbb{C}^{\times}). \tag{39.230}$$

Thus, if the map c can be written as $c = \mathrm{e}^f$ for some single-valued function $f : L \to \mathbb{C}$, then the Maslov index is always zero, and the axial $U(1)$ R-symmetry is unbroken. An example of such a Lagrangian submanifold is a special Lagrangian submanifold, for which c is a constant. Fig. 4 depicts two

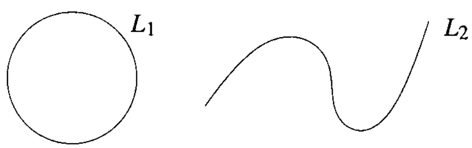


FIGURE 4. Two Lagrangians in $\mathbb{R}^2$

Lagrangian subspaces in $\mathbb{R}^2 = \mathbb{C}$ with $\Omega = dz$. The first one (L_1) is a circle parametrized by a periodic coordinate $t \equiv t + 2\pi$. We have $c(t) = i\,\mathrm{e}^{it}$ and thus the Maslov index is 2. A disc can be mapped to $\mathbb{C}$ with the boundary mapped isomorphically onto the circle. The Dirac index for such a map is 2, and thus the R-symmetry is indeed broken down to $\mathbb{Z}_2$. For the second one (L_2), the function c can be written as $c = \mathrm{e}^f$ for a single-valued function f on L_2. Thus, the R-symmetry is anomaly-free for this submanifold.

39.3.3. Some Applications. Let us A-twist the above theory on the open string worldsheet. After twisting, the quantum number of the fermions changes so that ψ_- takes values in $\phi^* T^{(1,0)} X$. One can see that the argument of the toy model is applicable to this case as well, in which we have

$E = \phi^* T^{(1,0)} X$ and $\tau : E \to E^*$ is associated with $E_{\mathbb{R}} = \phi^*[TL]^{(1,0)}$. We would like to compute the index of the Dirac operator $\widetilde{\mathcal{D}}$ of the twisted system. The computation is almost the same as in the untwisted model except that we do not have the $\sqrt{K}$ factors now. Thus we find

$$\mathrm{Index}\, \widetilde{\mathcal{D}} = \mu\left(\phi^* T^{(1,0)} X, \phi^*[TL]^{(1,0)}\right) + (\dim X)(2 - 2g - h). \tag{39.231}$$

This is the (real) dimension of the moduli space of holomorphic maps from $(\Sigma, \partial\Sigma)$ to (X, L).

We next consider the open version of topological gravity. Its fermionic sector can be identified as the one with $E = K^{\otimes 2}$, and the boundary condition is the one corresponding to the isomorphism as shown by Eq. (39.207). Since $K^{\otimes 2} \# (K^*)^{\otimes 2} = K^{\otimes 2}_{\Sigma\#\Sigma^*}$, which has $c_1 = 4g_{\Sigma\#\Sigma^*} - 4 = 8g - 8 + 4h$, the index of the Dirac operator $\mathcal{D}_{\mathrm{TG}}$ is $\mathrm{Index}\, \mathcal{D}_{\mathrm{TG}} = (8g-8+4h) + (2-2g+h) = 6g - 6 + 3h$. This can be regarded as the dimension of the moduli space $\mathcal{M}_{g,h}$ of Riemann surfaces of genus g and h boundary components:

$$\dim_{\mathbb{R}} \mathcal{M}_{g,h} = 6g - 6 + 3h. \tag{39.232}$$

We can combine the above two systems, namely, an open topological sigma model coupled to open topological gravity. The index of the fermionic operator can be regarded as the dimension of the moduli space $\mathcal{M}_{g,h}(X, L; \phi)$ of open holomorphic curves in (X, L) in the same homotopy class as ϕ. It is simply the sum of the indices, Eqs. (39.231)–(39.232):

$$\dim_{\mathbb{R}} \mathcal{M}_{g,h}(X, L; \phi) = \mu\left(\phi^* T^{(1,0)} X, \phi^*[TL]^{(1,0)}\right) + (3 - \dim X)\,(2g - 2 + h). \tag{39.233}$$

In particular, if X is a Calabi–Yau threefold and L is a special Lagrangian submanifold, the dimension is zero for all values of g and h and for all classes of the maps ϕ.

39.3.4. Appendix: Maslov Index. Let us consider the Euclidean space $\mathbb{R}^{2n}$ equipped with a symplectic form $\omega = \sum_{i=1}^n dx^i \wedge dy^i$, where $(x^1, y^1, \ldots, x^n, y^n)$ are orthonormal coordinates of $\mathbb{R}^{2n}$. A Lagrangian subspace L of $\mathbb{R}^{2n}$ is a middle-dimensional linear subspace such that the restriction of ω to L is vanishes. We will study the space of Lagrangian subspaces in $\mathbb{R}^{2n}$;

$$\Lambda(\mathbb{R}^{2n}, \omega) = \left\{ L \subset \mathbb{R}^{2n} \ \text{ Lagrangian subspace} \right\}. \tag{39.234}$$

We often abbreviate it as $\Lambda(\mathbb{R}^{2n})$ when there is no room for confusion. We call this space the *Lagrangian Grassmannian* of $(\mathbb{R}^{2n},\omega)$. One can also consider the space of oriented Lagrangian subspaces, the *oriented Lagrangian Grassmannian* $\widetilde{\Lambda}(\mathbb{R}^{2n},\omega)$.

The space $\mathbb{R}^{2n}$ can be considered as a complex plane $\mathbb{C}^n$ with coordinates $z^i = x^i + iy^i$, and the Euclidean metric g of $\mathbb{R}^{2n}$ is related to the symplectic form by $g(v,w)=\omega(v,Jw)$, where J is the multiplication by i. The two-form $h := g + i\omega$ defines a Hermitian inner product on $\mathbb{C}^n$; $h(v,Jw) = ih(v,w) = -h(Jv,w)$. An example of a Lagrangian subspace is the subspace $\mathbb{R}^n \subset \mathbb{C}^n$ spanned by $\partial/\partial x^i$.[2] Let L be a Lagrangian subspace. If we choose an orthonormal basis $e_1,\dots,e_n$ of L with respect to the metric induced from g of $\mathbb{R}^{2n}$, it defines, over $\mathbb{C}$, an orthonormal basis of $\mathbb{C}^n$ with respect to the Hermitian metric h,

$$h(e_i,e_j) = g(e_i,e_j) + i\omega(e_i,e_j) = \delta_{i,j} + i\times 0. \tag{39.235}$$

Conversely, given an orthonormal basis $f_1,\dots,f_n$ of $(\mathbb{C}^n,h)$, its $\mathbb{R}$-linear span is a Lagrangian subspace because $h(f_i,f_j) = \delta_{i,j} \in \mathbb{R}$. Thus, any Lagrangian subspace L can be written as

$$L = u(\mathbb{R}^n), \quad u \in U(n), \tag{39.236}$$

and any subspace of this form is a Lagrangian subspace. Since $u(\mathbb{R}^n) = u'(\mathbb{R}^n)$ if and only if $u' = u\times o$ for some $o \in O(n)$, we have found

$$\Lambda(\mathbb{R}^{2n}) \cong U(n)/O(n). \tag{39.237}$$

If we consider oriented Lagrangian subspaces, we obviously have

$$\widetilde{\Lambda}(\mathbb{R}^{2n}) \cong U(n)/SO(n). \tag{39.238}$$

The fundamantal groups of the oriented and ordinary Lagrangian Grassmannians are both $\mathbb{Z}$;

$$\mu : \pi_1\Big(\Lambda(\mathbb{R}^{2n})\Big) \xrightarrow{\cong} \mathbb{Z}, \tag{39.239}$$

$$\widetilde{\mu} : \pi_1\Big(\widetilde{\Lambda}(\mathbb{R}^{2n})\Big) \xrightarrow{\cong} \mathbb{Z}. \tag{39.240}$$

The isomorphisms are explicitly given as follows. A loop in $\Lambda(\mathbb{R}^{2n})$ is given by a map $t\in[0,1] \mapsto u_t(\mathbb{R}^n)$ where $u_t \in U(n)$ and $u_1 = u_0 \times g$ for some $g \in O(n)$. Then the map $t \in [0,1] \mapsto \det^2 u_t \in U(1)$ determines a loop

[2]We identify the tangent space at the origin $T_0\mathbb{R}^{2n}$ with $\mathbb{R}^{2n}$ itself.

in $U(1)$ since $\det^2 g = 1$. The winding number of this loop is called the *Maslov index* of the loop $\{u_t(\mathbb{R}^n)\}$. The Maslov index is additive under the composition of loops, and defines the isomorphism μ. For a loop in the oriented Grassmannian $\widetilde{\Lambda}(\mathbb{R}^{2n})$, the above g belongs to $SO(n)$, $\det g = 1$, and therefore $\det u_t$ defines a loop in $U(1)$. Its winding number, *oriented Maslov index*, defines the isomorphism $\widetilde{\mu}$. Note that there is a projection $p : \widetilde{\Lambda}(\mathbb{R}^{2n}) \to \Lambda(\mathbb{R}^{2n})$ corresponding to forgetting the orientation, and the indices are related by

$$\mu \circ p = 2\widetilde{\mu}. \tag{39.241}$$

Let us compute the Maslov index for a loop in the (oriented) Lagrangian Grassmannian in $\mathbb{R}^2$ ($n=1$ case). Let us consider the unit circle $S^1 = \{\, e^{it}\}$ in $\mathbb{R}^2 = \mathbb{C}$, and regard the tangent space at each point e^{it} as a subspace L_t of $\mathbb{R}^2$, which is of course Lagrangian. (See Fig. 5.) Thus, we have a loop of Lagrangian subspaces $\{L_t\}_{0\le t\le 2\pi}$. A tangent vector at t is given by

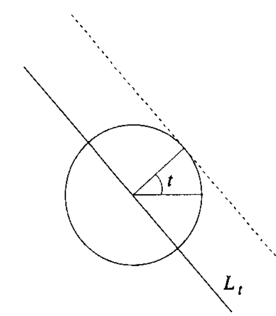


FIGURE 5. The Lagrangian subspace L_t

$u_t = i\,e^{it}$. The Maslov index is the winding number of $u_t^2 = -\,e^{2it}$, which is 2. Actually, the subspaces L_t can be given an orientation in such a way that u_t is a positive vector. The oriented Maslov index is the winding number of $u_t = i\,e^{it}$, which is 1.

39.4. Supersymmetric Ground States

In this section, we study the quantum mechanics of an open string stretched between two D-branes. In particular, we consider the case where the two D-branes are both A-branes with the same phase $e^{i\alpha}$ or both B-branes with the same phase $e^{i\beta}$. Then the quantum mechanics has a supersymmetry with the generator $Q = Q_A$ and $Q^\dagger = Q_A^\dagger$ or $Q = Q_B$ and $Q^\dagger = Q_B^\dagger$. We will study the supersymmetric ground states of this quantum mechanics. Naïvely, we expect that the supercharges obey the anti-commutation relations

$$Q^2 = (Q^\dagger)^2 = 0, \qquad \{Q, Q^\dagger\} = 2H. \tag{39.242}$$

If this is indeed the case, the supersymmetric ground states are in one-to-one correspondence with the Q-cohomology classes. However, one cannot always quantize the system so that Eq. (39.242) holds. In general, even if Q and $Q^\dagger$ are conserved, their anti-commutation relations could be modified as

$$Q^2 = \mathcal{Z}, \ (Q^\dagger)^2 = \mathcal{Z}^\dagger, \qquad \{Q, Q^\dagger\} = 2(H + \Delta). \tag{39.243}$$

In fact, this form comes from the $(2,2)$ supersymmetry algebra with non-trivial central charges: $\mathcal{Z} = \widetilde{Z}$ and $\Delta = \mathrm{Re}(Z)$ for $Q = Q_A$ while $\mathcal{Z} = Z$ and $\Delta = \mathrm{Re}(\widetilde{Z})$ for $Q = Q_B$. The correction Δ can be absorbed by redefinition of the Hamiltonian. However, if $\mathcal{Z}$ is nonzero we lose the cohomological structure of the theory. Of course there would be no supersymmetric ground states in such a case. Using the deformation invariance property, we will compute the Witten index $\mathrm{Tr}(-1)^F e^{-\beta(H+\Delta)}$ by going to a convenient region in the parameter space. We will see in many examples that the index has an interesting geometrical meaning.

We start with the intersecting D1-branes in $\mathbb{C}$ (A-branes) where one can explicitly quantize the open string theory. We next consider more general A-branes in non-linear sigma models and Landau–Ginzburg models. We are led to quantize the space of paths using a Morse function, in the same way as in the analysis of the supersymmetric ground states for the closed string theory in Ch. 13. The space of supersymmetric ground states we will find is the Floer homology groups for a pair of A-branes. We will also study B-branes: space-filling branes supporting holomorphic vector bundles, and point-like branes in Landau–Ginzburg models. In the former case, the space

of supersymmetric ground states is the Ext groups of the pair of holomorphic bundles. In the latter case, the space of ground states is the exterior algebra on $\mathbb{C}^n$, where n is the number of LG fields.

39.4.1. Intersecting Straight D1-Branes in $\mathbb{R}^2$. The first example we consider is the intersecting D1-branes in $\mathbb{R}^2 = \mathbb{C}$. We have already studied the open string quantum mechanics for the bosonic theory. We have to add the Dirac fermion system, but we have already studied the corresponding open string theory too. Thus, essentially, we only have to combine the two results. Since a line in $\mathbb{C}$ is a Lagrangian submanifold, the D1-branes we consider are A-branes. Throughout this subsection, we set the phase to be trivial, $e^{i\alpha} = 1$.

We are considering the sum of a massless complex scalar field ϕ and a Dirac fermion $\psi_\pm, \overline{\psi}_\pm$. We consider an open string stretched from one D1-brane at $\mathrm{Im}(\phi) = 0$ to the other at $\mathrm{Im}(e^{-i\theta}\phi) = 0$. The boundary conditions on the fields are

$$\left.\begin{array}{l} \partial_s \mathrm{Re}(\phi) = 0, \quad \phi = \overline{\phi}, \\ \psi_- = \overline{\psi}_+, \end{array}\right\} \text{ at } s = 0, \tag{39.244}$$

$$\left.\begin{array}{l} \partial_s \mathrm{Re}(e^{-i\theta}\phi) = 0, \quad e^{-i\theta}\phi = \overline{e^{i\theta}\phi}, \\ e^{-i\theta}\psi_- = \overline{e^{-i\theta}\psi_+}, \end{array}\right\} \text{ at } s = \pi. \tag{39.245}$$

The bosonic part is identical to the system considered in Sec. 39.1.2.4 (z there and ϕ here are related by $\phi = \sqrt{2}^{-1} z$) while the fermionic part is the A_0-$A_{-2\theta}$ system studied in Sec. 39.1.2.6. The bosonic and fermionic parts of the Hamiltonian are

$$H_B = \sum_{r \in \mathbb{Z} + \frac{\theta}{\pi}} |r| a_r^\dagger a_r + \frac{1}{24} - \frac{1}{2}\left(\frac{\theta}{\pi} - \left[\frac{\theta}{\pi}\right] - \frac{1}{2}\right)^2, \tag{39.246}$$

$$H_F = \sum_{r \in \mathbb{Z} + \frac{\theta}{\pi}} r {:}\overline{\psi}_{-r}\psi_r{:} - \frac{1}{24} + \frac{1}{2}\left(\frac{\theta}{\pi} - \left[\frac{\theta}{\pi}\right] - \frac{1}{2}\right)^2. \tag{39.247}$$

Thus, the total is simply

$$H = \sum_{r \in \mathbb{Z} + \frac{\theta}{\pi}} \left(|r| a_r^\dagger a_r + r {:}\overline{\psi}_{-r}\psi_r{:} \right). \tag{39.248}$$

We note that there is a unique ground state for each $\theta \notin \pi\mathbb{Z}$: it is

$$|0\rangle_B \otimes |0,-2\theta\rangle_F, \tag{39.249}$$

where $|0\rangle_B$ is the ground state of the bosonic sector annihilated by a_r $(\forall r)$ and $|0,-2\theta\rangle_F$ is the ground state of the fermionic sector introduced in Sec. 39.1.2.6. The fields have the following mode expansions:

$$\phi = \frac{i}{\sqrt{2}} \sum_{r\in\mathbb{Z}+\frac{\theta}{\pi}} \frac{1}{\sqrt{|r|}} \left(a_r\, \mathrm{e}^{i(rs-|r|t)} - a_r^\dagger\, \mathrm{e}^{i(rs+|r|t)} \right), \tag{39.250}$$

$$\psi_- = \sum_{r\in\mathbb{Z}+\frac{\theta}{\pi}} \psi_r\, \mathrm{e}^{ir(s-t)}, \quad \overline{\psi}_- = \sum_{r'\in\mathbb{Z}-\frac{\theta}{\pi}} \overline{\psi}_{r'}\, \mathrm{e}^{ir'(s-t)}, \tag{39.251}$$

$$\psi_+ = \sum_{r'\in\mathbb{Z}-\frac{\theta}{\pi}} \overline{\psi}_{r'}\, \mathrm{e}^{-ir'(s+t)}, \quad \overline{\psi}_+ = \sum_{r\in\mathbb{Z}+\frac{\theta}{\pi}} \psi_r\, \mathrm{e}^{-ir(s+t)}. \tag{39.252}$$

The system has A-type supersymmetry generated by

$$\begin{aligned} Q = \overline{Q}_+ + Q_- &= \frac{1}{2\pi}\int_0^\pi ds\, \{\overline{\psi}_+(\partial_t+\partial_s)\phi + (\partial_t-\partial_s)\overline{\phi}\psi_-\} \\ &= \sqrt{2}\left(\sum_{r>0} \sqrt{|r|}\psi_r a_r^\dagger + \sum_{r<0} \sqrt{|r|}\psi_r a_r \right), \end{aligned} \tag{39.253}$$

and

$$Q^\dagger = \sqrt{2}\left(\sum_{r>0} \sqrt{|r|}\overline{\psi}_{-r} a_r + \sum_{r<0} \sqrt{|r|}\overline{\psi}_{-r} a_r^\dagger \right). \tag{39.254}$$

It is straightforward to show that the relations given by Eq. (39.242) hold. The system also has an R-symmetry generated by F_A in Eq. (39.62). We shift the R-charges as

$$F^\theta := F_A + \frac{\theta}{\pi} - \frac{1}{2}, \tag{39.255}$$

so that it has integral eigenvalues.

The weighted partition function is essentially the product of the one of Eq. (39.49) for the bosonic system and the one of Eq. (39.63) for the fermionic system with $\alpha = -2\theta$:

(39.256)
$$\begin{aligned} \mathrm{Tr}\, \mathrm{e}^{2\pi i(a-\frac{1}{2})F^\theta} q_o^{H_o} &= \mathrm{e}^{2\pi i(a-\frac{1}{2})(\frac{\theta}{\pi}-\frac{1}{2})} \\ &\quad \times \mathrm{e}^{\pi i(\frac{\theta}{\pi}-[\frac{\theta}{\pi}]-\frac{1}{2})} \frac{\eta(iT)}{\vartheta\left[\begin{smallmatrix} \frac{\theta}{\pi}-\frac{1}{2} \\ \frac{1}{2} \end{smallmatrix}\right](0,iT)} \cdot \frac{\vartheta\left[\begin{smallmatrix} -\frac{\theta}{\pi}-\frac{1}{2} \\ a-\frac{1}{2} \end{smallmatrix}\right](0,iT)}{\eta(iT)} \\ &= (-1)^{[\frac{\theta}{\pi}]} \frac{\mathrm{e}^{2\pi i a(\frac{\theta}{\pi}-\frac{1}{2})} \vartheta\left[\begin{smallmatrix} \frac{\theta}{\pi}-\frac{1}{2} \\ \frac{1}{2}-a \end{smallmatrix}\right](0,iT)}{\vartheta\left[\begin{smallmatrix} \frac{\theta}{\pi}-\frac{1}{2} \\ \frac{1}{2} \end{smallmatrix}\right](0,iT)}. \end{aligned}$$

In particular, the Witten index, which is the value at $a \in \mathbb{Z}$, is

$$\mathrm{Tr}(-1)^{F^\theta} \mathrm{e}^{-\beta H} = (-1)^{[\frac{\theta}{\pi}]}. \tag{39.257}$$

One important lesson of this example is that there is no canonical way to make the R-charges to be integers, and there is an ambiguity in the choice of the "ground". For instance, suppose at $\theta = \pi/2$ we set $F = F_A$ and let us increase θ with the R-charge defined by Eq. (39.255). After increasing θ by πn we come back to the same configuration but the integral R-charge changes to $F = F_A + n$. However, if we give an orientation to each D1-brane, the configuration comes back to itself with period 2π in θ. Then the mod 2 R-charge can be well defined. We note that the Witten index from Eq. (39.257) can be interpreted as the intersection number of the two oriented lines in $\mathbb{R}^2$. There is actually a prescription to make the integral R-charge well defined: It is simply to keep track of how many times the D1-brane is rotated. In such a prescription, we need to assign an integer to each D1-brane.

39.4.2. Intersecting Lagrangians. We next consider more general A-branes for which the open string theory is not necessarily free. Let L_0 and L_1 be Lagrangian submanifolds of X. We consider the theory of an open string stretched between them. The bosonic field $\phi : [0,\pi]\times\mathbb{R} \to X$ obeys the boundary conditions

(39.258)
$$\left.\begin{array}{l} \phi \in L_0, \\ \partial_1\phi \in \phi^* NL_0, \end{array}\right\} \text{ at } x^1 = 0 \quad\text{and}\quad \left.\begin{array}{l} \phi \in L_1, \\ \partial_1\phi \in \phi^* NL_1, \end{array}\right\} \text{ at } x^1 = \pi,$$

while the fermionic fields $\psi_\pm \in \phi^*T^{(1,0)}X \otimes S_\pm$ and $\overline{\psi}_\pm \in \phi^*T^{(0,1)}X \otimes S_\pm$ obey

$$\left.\begin{array}{l}\tau_0\psi_- = \overline{\psi}_+,\\ \tau_0\overline{\psi}_- = \psi_+,\end{array}\right\}\text{ at } x^1=0 \quad\text{and}\quad \left.\begin{array}{l}\tau_1\psi_- = \overline{\psi}_+,\\ \tau_1\overline{\psi}_- = \psi_+,\end{array}\right\}\text{ at } x^1=\pi. \tag{39.259}$$

Here τ_0 is a linear map from $\phi^*T^{(1,0)}X \otimes S_\pm|_{x^1=0}$ to $\phi^*T^{(0,1)}X \otimes S_\mp|_{x^1=0}$ (and back) associated with the decomposition $TX|_{L_0} = TL_0 \oplus NL_0$, acting as 1 on TL_0 but as -1 on NL_0. Similarly for τ_1.

The space of bosonic configurations (at a given time) is thus the space of paths $\phi : [0,\pi] \to X$ obeying the boundary conditions in Eq. (39.258);

$$\Omega(L_0, L_1) = \left\{ \phi : [0,\pi] \to X \;\middle|\; \text{obeying Eq. (39.258)} \right\}. \tag{39.260}$$

The Hilbert space of states is formally considered as the space of differential forms on this space. The fermionic fields, obeying the anti-commutation relations $\{\psi^i_\pm(x^1), \overline{\psi}^{\overline{\jmath}}_\pm(y^1)\} = g^{i\overline{\jmath}}\delta(x^1-y^1)$, are identified as the operators

$$\begin{aligned}\psi^i_- &= \delta\phi^i\wedge, \quad \overline{\psi}^{\overline{\jmath}}_- = g^{i\overline{\jmath}} i_{\left(\frac{\delta}{\delta\phi^i}\right)},\\ \overline{\psi}^{\overline{\jmath}}_+ &= \delta\overline{\phi}^{\overline{\jmath}}\wedge, \quad \psi^i_+ = g^{i\overline{\jmath}} i_{\left(\frac{\delta}{\delta\overline{\phi}^{\overline{\jmath}}}\right)}.\end{aligned} \tag{39.261}$$

In general, the path space $\Omega(L_0, L_1)$ consists of connected components, and the Hilbert space decomposes into the sum of the corresponding subspaces. We will focus on one component, $\Omega_{\phi_0}(L_0, L_1)$, consisting of the paths homotopic to some base path ϕ_0.

The supercharge $Q = Q_A$ is expressed as

$$iQ = \int_{[0,\pi]} \left\{ ig_{i\overline{\jmath}}\overline{\psi}^{\overline{\jmath}}_+\partial_0\phi^i + ig_{i\overline{\jmath}}\psi^i_-\partial_0\overline{\phi}^{\overline{\jmath}} - ig_{i\overline{\jmath}}\psi^i_-\partial_1\overline{\phi}^{\overline{\jmath}} + ig_{i\overline{\jmath}}\overline{\psi}^{\overline{\jmath}}_+\partial_1\phi^i \right\} dx^1. \tag{39.262}$$

It can be written in the familiar form

$$iQ = \delta + \delta h\wedge \tag{39.263}$$

if there is a function h of $\phi(x^1)$ such that

$$\frac{\delta h}{\delta\phi^i} = -ig_{i\overline{\jmath}}\partial_1\overline{\phi}^{\overline{\jmath}}, \quad \frac{\delta h}{\delta\overline{\phi}^{\overline{\jmath}}} = ig_{i\overline{\jmath}}\partial_1\phi^i. \tag{39.264}$$

We encountered a similar problem in Sec. 13.3. As in the discussion there we put

$$h = \int_{[0,\pi]\times[0,1]} \widehat{\phi}^*\omega, \tag{39.265}$$

where ω is the Kähler form and $\widehat{\phi}$ is a homotopy in $\Omega_{\phi_0}(L_0, L_1)$ connecting the base path ϕ_0 to ϕ. Namely, a map $(x^1,\tau) \in [0,\pi]\times[0,1] \mapsto \widehat{\phi}(x^1,\tau) \in X$ such that $\widehat{\phi}|_{\tau=0} = \phi_0$, $\widehat{\phi}|_{\tau=1} = \phi$, $\widehat{\phi}|_{x^1=0} \in L_0$ and $\widehat{\phi}|_{x^1=\pi} \in L_1$. Under a deformation of ϕ and $\widehat{\phi}$, it varies as

$$\delta h = \int_{\partial([0,\pi]\times[0,1])} \left\{ -ig_{i\overline{\jmath}}\delta\widehat{\phi}^i d\widehat{\overline{\phi}}^{\overline{\jmath}} + ig_{i\overline{\jmath}}d\widehat{\phi}^i\delta\widehat{\overline{\phi}}^{\overline{\jmath}} \right\}. \tag{39.266}$$

Since $\widehat{\phi}$ is fixed to be ϕ_0 at the $\tau=0$ boundary, we have $\delta\phi|_{\tau=0} = 0$. At the string end boundary $x^1=0$ the integrand vanishes since both $\delta\widehat{\phi}$ and $\partial_\tau\widehat{\phi}$ are tangent to L_0 which is Lagrangian. The same is true of the other end $x^1=\pi$. Thus, we only have the contribution from the $\tau=1$ boundary

$$\delta h = \int_{[0,\pi]} \left\{ -ig_{i\overline{\jmath}}\delta\phi^i d\overline{\phi}^{\overline{\jmath}} + ig_{i\overline{\jmath}}d\phi^i\delta\overline{\phi}^{\overline{\jmath}} \right\}. \tag{39.267}$$

Since $\delta h = 0$ as long as $\delta\phi = 0$, h does not change under small deformations of the homotopy $\widehat{\phi}$ fixing ϕ. Therefore h is well defined at least locally on $\Omega_{\phi_0}(L_0, L_1)$. Furthermore the result Eq. (39.267) shows that Eq. (39.265) solves the problem shown in Eq. (39.264). However, h can change for a "large" change of homotopy $\widehat{\phi}$ — under the change of the homotopy class of the homotopy $\widehat{\phi}$. To make it well defined globally we consider the universal cover of $\Omega_{\phi_0}(L_0, L_1)$ and require the wave-functions to be invariant under the covering transformations. A model for the universal cover is indeed the space of pairs $(\phi, \widehat{\phi})$ up to homotopy:

$$\widetilde{\Omega}_{\phi_0}(L_0, L_1) = \left\{ (\phi, \widehat{\phi}) \;\middle|\; \widehat{\phi}|_{\tau=0} = \phi_0,\; \widehat{\phi}|_{\tau=1} = \phi \right\} \Big/ \simeq, \tag{39.268}$$

where $(\phi, \widehat{\phi}) \simeq (\phi', \widehat{\phi}')$ if and only if $\phi = \phi'$ and $\widehat{\phi}$ can be continuously deformed to $\widehat{\phi}'$. The function h is globally well defined on this space, as we have seen explicitly.

We now apply the method developed in Sec. 10.5 to determine the space of supersymmetric ground states of the theory. The first step is to find the critical points of the (multi-valued) function h. By Eq. (39.267) or Eq.

(39.264), the critical points in $\Omega_{\phi_0}(L_0, L_1)$ are the paths with $\partial_1\phi = 0$, i.e., the constant maps $\phi(x^1) \equiv p$. Since the left and the right ends of ϕ must lie in L_0 and L_1, this is possible if and only if $p \in L_0 \cap L_1$. Thus, the critical points of h in $\Omega_{\phi_0}(L_0, L_1)$ are the constant maps to $L_0 \cap L_1$. We denote its lift to the cover $\widetilde{\Omega}_{\phi_0}(L_0, L_1)$ by $\mathrm{Crit}_{\phi_0}(L_0, L_1)$. At each point of $\mathrm{Crit}_{\phi_0}(L_0, L_1)$, there is an approximate supersymmetric ground state of the quantum mechanics on $\widetilde{\Omega}_{\phi_0}(L_0, L_1)$. The state shown in Eq. (39.249) is a model of such a perturbative ground state. If we take the average so that it is invariant under the covering transformations, it defines a state of the quantum mechanics on $\Omega_{\phi_0}(L_0, L_1)$ which is peaked at a constant map to $L_0 \cap L_1$.

39.4.2.1. *R-charge.* We next compute the axial R-charges of these (approximate) ground states. To this end, we first determine the Morse index of h at the critical points, $\mathrm{Crit}_{\phi_0}(L_0, L_1)$. The Hessian $\delta^2 h/\delta\phi^I(x^1)\delta\phi^J(y^1)$ is essentially the space derivative ∂_1 and has infinitely many positive as well as negative eigenvalues. Thus we have to regularize the Morse index. One may use zeta-function regularization, as in the definition of the R-charge for intersecting straight D1-branes. However, it is not enough to do it point-by-point, but it has to be compatible with the relative Morse index between different points. As observed in Sec. 10.5, the relative Morse index between two points is equal to the index of the fermion Dirac operator in a bosonic background that connects them. We now express the Dirac index in terms of a topological invariant of the bosonic background.

Computation of Relative Morse Index. Let $(p, \widehat{\phi}_p)$ and $(q, \widehat{\phi}_q)$ be critical points of h, where $p, q \in L_0 \cap L_1$. We suppose that there is a homotopy $\tau \in [0,1] \mapsto \widehat{\phi}(-,\tau) \in \Omega_{\phi_0}(L_0, L_1)$ connecting the constant maps q and p such that the composition $\widehat{\phi}\#\widehat{\phi}_q$ is continuously connected to $\widehat{\phi}_p$. We fix a trivialization $\widehat{\phi}^*TX \cong ([0,\pi]\times[0,1]) \times \mathbb{C}^n$ such that

- it is constant along $\tau = 0$ and $\tau = 1$,
- the subspaces $\widehat{\phi}^*TL_0$ at $(s,\tau) = (0,0)$ and $(0,1)$ are mapped to $\mathbb{R}^n$,
- the subspaces $\widehat{\phi}^*TL_1$ at $(s,\tau) = (\pi,0)$ and $(\pi,1)$ are mapped to $i\mathbb{R}^n$.

Let $\lambda_{\widehat{\phi}} : \partial([0,\pi]\times[0,1]) \to \Lambda(\mathbb{C}^n)$ be a loop defined by

$$\lambda_{\widehat{\phi}}(s,0) = e^{is/2}\mathbb{R}^n, \quad \lambda_{\widehat{\phi}}(s,1) = e^{is/2}\mathbb{R}^n, \tag{39.269}$$

$$\lambda_{\widehat{\phi}}(0,\tau) = T_{\widehat{\phi}(0,\tau)}L_0, \quad \lambda_{\widehat{\phi}}(\pi,\tau) = T_{\widehat{\phi}(\pi,\tau)}L_1. \tag{39.270}$$

We claim that the relative Morse index is equal to the Maslov index of this loop:

$$\mu(p, \widehat{\phi}_p) - \mu(q, \widehat{\phi}_q) = \mu(\lambda_{\widehat{\phi}}). \tag{39.271}$$

Below, we give a derivation of this claim.

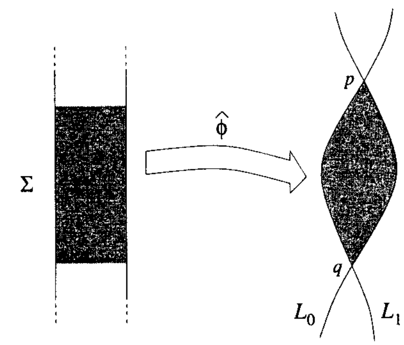


FIGURE 6. The map $\widehat{\phi} : \Sigma \to X$

We extend $\widehat{\phi}$ to a map from an infinitely long strip $\Sigma = [0,\pi]\times\mathbb{R}$ to X (see Fig. 6), by first attaching the regions $\tau \le 0$ and $\tau \ge 1$ that are mapped constantly to q and p respectively, and then smoothing. It follows from the general story in Sec. 10.5 that the relative Morse index is equal to the index of the Dirac operator acting on normalizable fermionic fields on Σ.

$$\Delta\mu = \mathrm{Index}\,\mathcal{D}.$$

We assume for simplicity that L_0 and L_1 orthogonally intersect at p and q. (This assumption is not essential and the following computation applies for more general cases as well.) Then the fermionic sector at $\tau < 0$ and $\tau > 1$ is identical to the tensor product of n copies of the fermionic sector that appears in the system of intersecting D1-branes with $\theta = \pm\pi/2$ in Sec. 39.4.1, or A_0-$A_{\mp\pi}$ system in Sec. 39.1.2.6. The zero modes can be expressed as Eqs. (39.251)–(39.252), with the Wick rotation $t \to -i\tau$ understood. For the normalizability at $\tau \to \mp\infty$, only the $r \le -\frac{1}{2}$ and the $r' \le -\frac{1}{2}$ components can be nonzero at $\tau < 0$ while we need only $r \ge \frac{1}{2}$ and $r' \ge \frac{1}{2}$

components at $\tau > 1$. In particular, we find

$$\psi_- = \begin{cases} \sum_{n=0}^{\infty} \psi_{n+\frac{1}{2}}\, \mathrm{e}^{i(n+\frac{1}{2})(s+i\tau)}, & \tau > 1, \\ \sum_{n=0}^{\infty} \psi_{-n-\frac{1}{2}}\, \mathrm{e}^{-i(n+\frac{1}{2})(s+i\tau)}, & \tau < 0, \end{cases} \tag{39.272}$$

$$\overline{\psi}_+ = \begin{cases} \sum_{n=0}^{\infty} \psi_{n+\frac{1}{2}}\, \mathrm{e}^{-i(n+\frac{1}{2})(s-i\tau)}, & \tau > 1, \\ \sum_{n=0}^{\infty} \psi_{-n-\frac{1}{2}}\, \mathrm{e}^{i(n+\frac{1}{2})(s-i\tau)}, & \tau < 0. \end{cases} \tag{39.273}$$

Let us introduce coordinates $z = \mathrm{e}^{\zeta}$ and $w = \mathrm{e}^{-\zeta}$ where $\zeta = -i(s+i\tau)$. The strip Σ can be compactified to a hemisphere $\overline{\Sigma}$ by adding the points $z = 0$ and $w = 0$ (See Fig. 7). Eqs. (39.272)-(39.273) shows that

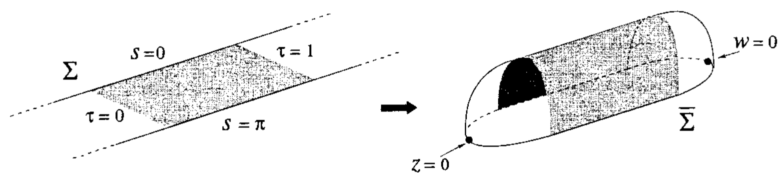


FIGURE 7. Compactification of the strip

$\psi_- z^{-\frac{1}{2}}$ and $\overline{\psi}_+ \overline{z}^{-\frac{1}{2}}$ extend regularly to $z = 0$ while $\psi_- w^{-\frac{1}{2}}$ and $\overline{\psi}_+ \overline{w}^{-\frac{1}{2}}$ extend regularly to $w = 0$. Since $(d\zeta)^{\frac{1}{2}} = z^{-\frac{1}{2}}(dz)^{\frac{1}{2}} = (-w)^{-\frac{1}{2}}(dw)^{\frac{1}{2}}$, this means that $\psi_-(d\zeta)^{\frac{1}{2}}$ and $\overline{\psi}_+(d\overline{\zeta})^{\frac{1}{2}}$ extend over $\overline{\Sigma}$ as sections of the bundle $E = \widehat{\phi}^* T^{(1,0)}X \otimes \sqrt{K_{\overline{\Sigma}}}$ and its dual E^*. The boundary condition for the fermion was the one associated with the decomposition $E|_{\partial\Sigma} = E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ where $E_{\mathbb{R}}$ is spanned by $\widehat{\phi}^*[TL_0]^{(1,0)} \otimes (d\zeta)^{\frac{1}{2}}$ on $s = 0$ and by $\widehat{\phi}^*[TL_1]^{(1,0)} \otimes (d\zeta)^{\frac{1}{2}}$ on $s = \pi$. The sub-bundles $E_{\mathbb{R}}|_{s=0}$ and $E_{\mathbb{R}}|_{s=\pi}$ glue smoothly at $z = 0$ and $w = 0$. To see this we first note that, in a region close to $z = 0$, $\widehat{\phi}^*[TL_0]^{(1,0)} = \mathbb{R}^n$ along $s = 0$ and $\widehat{\phi}^*[TL_1]^{(1,0)} = i\mathbb{R}^n$ along $s = \pi$. Around $z = 0$, we have $\mathbb{R}(d\zeta)^{\frac{1}{2}} = \mathbb{R}(dz)^{\frac{1}{2}}$ along $s = 0$, while $\mathbb{R}(d\zeta)^{\frac{1}{2}} = i\mathbb{R}(dz)^{\frac{1}{2}}$ along $s = \pi$. Thus we can put $E_{\mathbb{R}} = \mathbb{R}^n \otimes (dz)^{\frac{1}{2}}$ in a neighborhood of $z = 0$. Similarly we can put $E_{\mathbb{R}} = i\mathbb{R}^n \otimes (dw)^{\frac{1}{2}}$ in a neighborhood of $w = 0$. Thus the decomposition $E_{\mathbb{R}} \oplus iE_{\mathbb{R}}$ extends over $\overline{\Sigma}$. Applying the formula from Eq. (39.219) we find that the index of the Dirac operator is

$$\mathrm{Index}\, \mathcal{D} = \mu(E, E_{\mathbb{R}}) + n. \tag{39.274}$$

Let us fix a trivialization of E over the whole $\overline{\Sigma}$, extending the previous one on $[0,\pi] \times [0,1]$. Along the arc in $\partial\overline{\Sigma}$ that starts from $(s,\tau) = (0,0)$, passes through $z = 0$ and ends on $(s,\tau) = (\pi, 0)$, the sub-bundle $E_{\mathbb{R}} = \mathbb{R}^n \otimes (dz)^{\frac{1}{2}}$ rotates in phase by $-\pi/2$. Along another arc starting from $(s,\tau) = (0,1)$, passing through $w = 0$ and ending on $(s,\tau) = (\pi, 1)$, the sub-bundle $E_{\mathbb{R}} = i\mathbb{R}^n \otimes (dw)^{\frac{1}{2}}$ rotates in phase by $\pi/2$. The loop in $\Lambda(\mathbb{C}^n)$ determined by $E_{\mathbb{R}}$ is thus homotopically equivalent to the loop $\lambda_E : [0,\pi] \times [0,1] \to \Lambda(\mathbb{C}^n)$ given by

$$\lambda_E(s,0) = \mathrm{e}^{-is/2}\mathbb{R}^n, \quad \lambda_E(s,1) = \mathrm{e}^{is/2}\mathbb{R}^n, \tag{39.275}$$

$$\lambda_E(0,\tau) = T_{\widehat{\phi}(0,\tau)}L_0, \quad \lambda_E(\pi,\tau) = T_{\widehat{\phi}(\pi,\tau)}L_1. \tag{39.276}$$

This differs from $\lambda_{\widehat{\phi}}$ in Eqs. (39.269)–(39.270) simply along the $\tau = 0$ segment. The difference is simply the loop $\mathrm{e}^{-is/2}\mathbb{R}^n$: $0 \leq s \leq 2\pi$. Thus, we find

$$\mu(E, E_{\mathbb{R}}) = \mu(\lambda_E) = \mu(\lambda_{\widehat{\phi}}) + \mu(\{\mathrm{e}^{-is/2}\mathbb{R}^n\}_{0\leq s\leq 2\pi}) = \mu(\lambda_{\widehat{\phi}}) - n. \tag{39.277}$$

Inserting this into (39.274), we find Eq. (39.271).

A Consistent Assignment of the Morse Index μ. We now assign the (regularized) Morse index at each point of $\mathrm{Crit}_{\phi_0}(L_0, L_1)$, so that the consistency condition Eq. (39.271) is satisfied. We choose and fix a Lagrangian sub-bundle E_0 of the bundle $\phi_0^* TX$ over $[0,\pi]$ such that $E_0 = \phi_0^* TL_0$ at $s = 0$ and $E_0 = \phi_0^* TL_1$ at $s = \pi$. Now, let $(p, \widehat{\phi}_p)$ be a point of $\mathrm{Crit}_{\phi_0}(L_0, L_1)$. Choose a trivialization $\widehat{\phi}_p^* TX \cong ([0,\pi] \times [0,1]) \times \mathbb{C}^n$ such that

- it is constant along $\tau = 1$, and $E_0 = \mathbb{R}^n$ along $\tau = 0$,
- $T_p L_0 = \mathbb{R}^n$ and $T_p L_1 = i\mathbb{R}^n$.

Let us define a loop $\lambda_{\widehat{\phi}_p} : \partial([0,\pi] \times [0,1]) \to \Lambda(\mathbb{C}^n)$ by

$$\lambda_{\widehat{\phi}_p}(s,0) = E_0|_s, \quad \lambda_{\widehat{\phi}_p}(s,1) = \mathrm{e}^{is/2}\mathbb{R}^n, \tag{39.278}$$

$$\lambda_{\widehat{\phi}_p}(0,\tau) = T_{\widehat{\phi}_p(0,\tau)}L_0, \quad \lambda_{\widehat{\phi}_p}(\pi,\tau) = T_{\widehat{\phi}_p(\pi,\tau)}L_1. \tag{39.279}$$

We define the Morse index at $(p, \widehat{\phi}_p)$ as the Maslov index of this loop:

$$\mu(p, \widehat{\phi}_p) := \mu(\lambda_{\widehat{\phi}_p}). \tag{39.280}$$

It is obvious that the consistency condition Eq. (39.271) is satisfied for this definition.

R-charge from Morse Index. The regularized Morse index given by Eq. (39.280) can be considered as the regularized R-charge of the approximate ground state peaked at the critical point $(p, \widehat{\phi}_p)$ in $\widetilde{\Omega}_{\phi_0}(L_0, L_1)$. However, we are considering the covering space $\widetilde{\Omega}_{\phi_0}(L_0, L_1)$ just for a technical reason: to make h single-valued. What we are really interested in is the quantum mechanics on the path space $\Omega_{\phi_0}(L_0, L_1)$ itself. Thus, we would like to determine the R-charge of the averages of wave-functions that are invariant under the covering transformation group. However, we encounter the possibility that the regularized Morse index shown in Eq. (39.280) may not be invariant under the covering transformations. This is the case when the relative Morse index from Eq. (39.271) is nonzero for some $\widehat{\phi}$ connecting a point p to itself.

This does not happen when X is Calabi–Yau and the anomaly-free condition in Sec. 39.3.2.1 is satisfied for both L_0 and L_1, namely, if the functions $c_0 : L_0 \to \mathbb{C}^\times$ and $c_1 : L_1 \to \mathbb{C}^\times$ can be represented as $c_0 = e^{f_0}$ and $c_1 = e^{f_1}$ for single-valued functions f_0 and f_1. In such a case, for any homotopy $\widehat{\phi}$ connecting a constant map $p \in L_0 \cap L_1$ to itself, the loop $\lambda_{\widehat{\phi}}$ is topologically trivial; along the $s = 0$ and $s = \pi$ segments, it comes back to the same points without any winding. Thus, if the anomaly-free condition is satisfied on both L_0 and L_1, the integer-valued R-charge is well defined for the theory of an open string stretched between L_0 and L_1.

Otherwise, integer-valued R-charge is not well defined. If X is not Calabi–Yau, this is always the case. However, if the relative Morse index for $p = q$ is always divisible by M, the R-charge is well-defined modulo M, or $\mathbb{Z}_M$ grading can be put on the ground states. If the axial R-symmetry is broken into $\mathbb{Z}_{M_0}$ and $\mathbb{Z}_{M_1}$ on L_0 and L_1 respectively, M is not larger than the greatest common divisor of M_0 and M_1. If L_0 and L_1 are oriented, the loop $\lambda_{\widehat{\phi}}$ for $\widehat{\phi}$ with $p = q$ can lift to a loop in the oriented Grassmannian $\widetilde{\Lambda}(\mathbb{C}^n)$. Since the Maslov index for such a loop is always even, as shown by Eq. (39.241), the relative Morse index for $p = q$ is always even. Thus, the mod 2 R-charge is always well defined for oriented Lagrangians.

There is also a subtlety in the choice of the "ground" of the R-charge, as in the example of the intersecting D1-branes in $\mathbb{R}^2$. The regularized Morse index from Eq. (39.280) may shift by an integer under a change of E_0. However, for *oriented* Lagrangians L_0, L_1, mod 2 R-charge can be well defined by taking as E_0 an oriented Lagrangian sub-bundle such that the

isomorphisms $E_0 = \phi_0^* TL_0$ at $s = 0$ and $E_0 = \phi_0^* TL_1$ at $s = \pi$ hold as oriented vector spaces: any change in such an E_0 results in the shift of the Maslov index by an even integer.

Witten Index. We have discussed the R-charge of the approximate ground states that are peaked at the constant maps to the intersection points of L_0 and L_1. In particular, we found that the mod 2 R-charge is well-defined if L_0 and L_1 are oriented. Then the Witten index of the theory is well defined and can be computed by summing $(-1)^{\mu(p)}$ over $p \in L_0 \cap L_1$. We claim that $\mu(p)$ is even (resp. odd) if and only if L_0 and L_1 intersect positively (resp. negatively) at p; namely, the wedge product of the volume forms of L_0 and L_1 is proportional to that of X with a positive (resp. negative) proportionality constant. To see this it is enough to note that the loop $\lambda_{\widehat{\phi}_p}$ lifts to a loop in the oriented Grassmannian $\widetilde{\Lambda}(\mathbb{C}^n)$ if and only if L_0 and L_1 intersect positively at p. This shows that

$$\mathrm{Tr}\,(-1)^F = \#(L_0 \cap L_1). \tag{39.281}$$

This generalizes the observation made for intersecting D1-branes in $\mathbb{R}^2$.

39.4.2.2. *Floer Homology Group.* In order to find the true ground states of the theory, we need to compute the Q-cohomology of the Morse–Witten complex

$$\cdots \xrightarrow{Q} C^{i-1} \xrightarrow{Q} C^i \xrightarrow{Q} C^{i+1} \xrightarrow{Q} \cdots, \tag{39.282}$$

which is constructed on the set of critical points $\mathrm{Crit}_{\phi_0}(L_0, L_1)$,

$$C^i = \sum_{\mu(p,\widehat{\phi}_p)=i} \mathbb{C}[(p, \widehat{\phi}_p)]. \tag{39.283}$$

The operator Q is defined in terms of the number of gradient flow lines from $(q, \widehat{\phi}_q)$ to $(p, \widehat{\phi}_p)$ with $\mu(p, \widehat{\phi}_p) - \mu(q, \widehat{\phi}_q) = 1$. The gradient flow equation reads as

$$\frac{d}{d\tau}\phi^i = g^{i\bar{\jmath}} \frac{\delta h}{\delta \overline{\phi^{\jmath}}} = i\partial_1 \phi^i, \tag{39.284}$$

which is nothing but the Cauchy-Riemann equation $\partial_{\bar{z}}\phi^i = 0$ on the worldsheet with complex coordinate $z = x^1 + i\tau$. Thus, the coboundary operator Q is determined by counting the number of holomorphic maps

$\phi : [0,\pi]\times\mathbb{R}\to X$ such that

$$\phi|_{x^1=0}\in L_0,\quad \phi|_{x^1=\pi}\in L_1,$$
$$\lim_{\tau\to-\infty}\phi = q,\quad \lim_{\tau\to+\infty}\phi = p, \tag{39.285}$$
$$\phi\#\widehat{\phi}_q \simeq \widehat{\phi}_p.$$

The cohomology group is called the *Floer cohomology group for the pair of Lagrangian submanifolds* (L_0, L_1), and is denoted by $HF(L_0,L_1)$. Clearly, this group is the space of supersymmetric ground states of the open string

$$\mathcal{H}_{\rm SUSY}\cong HF^\bullet(L_0,L_1). \tag{39.286}$$

For example, let us consider two Lagrangian submanifolds L_0 and L_1 in $\mathbb{R}^2$ as depicted in Fig. 8, where a base path ϕ_0 is also chosen. The

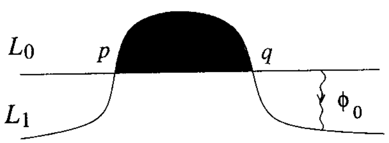


FIGURE 8

two Lagrangians intersect at two points, p and q in Fig. 8. There is an obvious and unique homotopy connecting ϕ_0 to q (or to p). Thus, the critical point set is $\mathrm{Crit}_{\phi_0}(L_0,L_1)=\{p,q\}$ where we suppressed the information on homotopy (since it is unique). It is straightforward to show that (under an obvious choice of the bundle E_0 over ϕ_0) the Morse index is given by $\mu(p)=0$ and $\mu(q)=-1$. Thus, we have

$$C^i=\begin{cases}\mathbb{C}[p] & i=0,\\ \mathbb{C}[q] & i=-1.\end{cases} \tag{39.287}$$

The coboundary operator acts by zero on $[p]$. The action on the state $[q]$ is determined by counting the number of holomorphic maps $\phi:[0,\pi]\times\mathbb{R}\to\mathbb{C}$ obeying Eq. (39.285). There is only a single map, where the image is depicted in the shaded region in the figure. Thus, we find $Q[q]=[p]$. This shows that

$$HF^\bullet(L_0,L_1)=0. \tag{39.288}$$

There is no supersymmetric ground state, which is consistent with $\#(L_0\cap L_1)=0$.

As another example, consider the Lagrangians in Fig. 9. They intersect

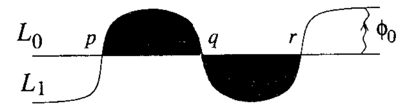


FIGURE 9

at three points, p, q, and r, whose Morse indices are 0, -1 and 0 respectively. Thus, we have

$$C^i=\begin{cases}\mathbb{C}[p]\oplus\mathbb{C}[r] & i=0,\\ \mathbb{C}[q] & i=-1.\end{cases} \tag{39.289}$$

The coboundary operator is given by $Q[q]=[p]\pm[r]$ while $Q[p]=Q[r]=0$. Thus, the Floer cohomology group is given by

$$HF^i(L_0,L_1)=\delta_{i,0}\mathbb{C}. \tag{39.290}$$

There is a non-trivial supersymmetric ground state. This is consistent with the fact that $\#(L_0\cap L_1)=1$.

Anomaly of $Q^2=0$. We have seen in the above examples that $Q^2=0$. However, we should note that the relation $Q^2=0$ is not guaranteed. In the finite dimensional case, we have seen that $Q^2=0$ can be proved directly (See Sec. 10.5.4). The essential point was the boundary relation shown in Eq. (10.355): for a given pair of critical points of relative Morse index 2, the boundary of a family of gradient flow lines between them consists of broken gradient flow lines, where breaking occurs at other critical points of the mid-value of the Morse index. We should therefore check if a boundary or a limit of a family of holomorphic maps consists only of "broken holomorphic maps". There can be two other possibilities:

(I) Bubbling-off of holomorphic spheres at some points of the strip;

(II) Bubbling-off of holomorphic disks on the boundaries of the strip.

One may notice that the possibility (I) could have endangered the relation $Q^2=0$ even in the closed string case. However, such a phenomenon occurs in codimension 2, and cannot affect the boundary relation. The possibility (II) is more serious; it occurs in codimension 1, and may indeed destroy the relation $Q^2 = 0$. In the above two examples (Figs 8–9) there are no holomorphic disks ending on L_0 or L_1, and therefore phenomenon (II) did not happen. However, in general, L_0 or L_1 admits holomorphic disks and one should worry about phenomenon (II). To be more precise, $\langle [(p,\widehat{\phi}_p)], Q^2[(q,\widehat{\phi}_q)]\rangle$ receives a non-trivial contribution if there is a holomorphic disk $w : (D^2, \partial D^2) \to (X, L_0)$ or (X, L_1), and a holomorphic map $\phi' : [0,\pi] \times \mathbb{R} \to X$ obeying Eq. (39.285) with the last one replaced by $w\#\phi'\#\widehat{\phi}_q = \widehat{\phi}_p$, such that they intersect on their boundaries. If the sum of the contributions from all such maps is nonzero, the relation $Q^2 = 0$ is obstructed. In other words, there is an anomaly in the supersymmetry algebra due to an instanton effect.

Suppose $Q^2 = 0$ fails in this way. Is this the end of the story and will we never have a cochain complex nor supersymmetric ground states? One may actually try to modify the operator Q so that $Q^2 = 0$ is satisfied. This is indeed possible if a certain condition is met. This corresponds to modifying the worldsheet theory, say by deforming the location of the D-brane. We will not develop the story further, and refer the reader to the work by Fukaya, Oh, Ohta and Ono [**97**].

39.4.3. Parallel Lagrangians in Landau–Ginzburg Models. Let us next consider A-branes in Landau–Ginzburg models. We assume that the Landau–Ginzburg variables are coordinates of a (non-compact) Calabi–Yau manifold X and the superpotential W grows at the infinity of X. We also assume that W has only non-degenerate critical points so that the bulk theory has a mass-gap. We consider D-branes wrapped on the wave-front trajectories γ_a and γ_b emanating from critical points p_a and p_b of W. We recall that $\operatorname{Im} W$ is constant on γ_a and γ_b, and we assume that the constants are separated, $\operatorname{Im} W(p_a) \neq \operatorname{Im} W(p_b)$.

The supercharge can be written as

$$iQ = \int_{[0,\pi]} \Big\{ ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}\partial_0\phi^i + ig_{i\bar{j}}\psi_-^i\partial_0\overline{\phi}^{\bar{j}} - ig_{i\bar{j}}\psi_-^i\partial_1\overline{\phi}^{\bar{j}} + ig_{i\bar{j}}\overline{\psi}_+^{\bar{j}}\partial_1\phi^i - \frac{1}{2}\psi_-^i\partial_i W - \frac{1}{2}\overline{\psi}_+^{\bar{i}}\overline{\partial}_{\bar{i}}\overline{W} \Big\} dx^1. \tag{39.291}$$

From this we see that Q and its Hermitian conjugate obey the anti-commutation relation

$$\{Q, Q^\dagger\} = 2(H + \Delta\operatorname{Im} W), \tag{39.292}$$

where $\Delta\operatorname{Im} W = \operatorname{Im} W(p_b) - \operatorname{Im} W(p_a)$. The nilpotency relation $Q^2 = 0$ is not modified (at least to the leading order). One can make the anti-commutation relation into the standard form given by Eq. (39.242) by redefining the Hamiltonian as $\widetilde{H} = H + \Delta\operatorname{Im} W$, or by the shift of the action as

$$\widetilde{S} = S - \int_{x^1=\pi} dx^0 \operatorname{Im} W + \int_{x^1=0} dx^0 \operatorname{Im} W. \tag{39.293}$$

The space of states can be considered as the space of differential forms on the path space $\Omega(\gamma_a,\gamma_b)$. As before, the supercharge can be written as $iQ = \delta + \delta h \wedge$, where h is a function

$$h = \int_{[0,\pi]\times[0,1]} \widehat{\phi}^*\omega - \int_{[0,\pi]} \operatorname{Re} W dx^1, \tag{39.294}$$

which is well-defined on a cover of $\Omega(\gamma_a,\gamma_b)$. The critical points of h are the paths obeying the equation $\delta h = 0$ which is given by

$$\partial_1\phi^i = -\frac{i}{2} g^{i\bar{j}}\partial_{\bar{j}}\overline{W}. \tag{39.295}$$

This is the equation for the gradient flow of the function $-\operatorname{Im} W$. The image in the W-plane is thus a straight line from $W(\gamma_a)$ down to $W(\gamma_b)$ in the negative imaginary direction. In particular, there is no solution if $\operatorname{Im} W(p_a) < \operatorname{Im} W(p_b)$. There is no supersymmetric ground state in such a case.

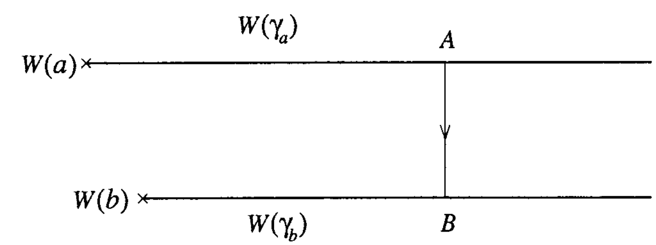


FIGURE 10. The image of D-branes in the W-plane and a path between them

If $\operatorname{Im} W(p_a) > \operatorname{Im} W(p_b)$, critical points may exist. Let us count how many of them are there. We are considering the situation as depicted in Fig. 10. We consider the wave-front at the point B along the straight line from $W(p_b)$ and another wave-front at B along the broken segment starting from $W(p_a)$ and cornering at the point A. From the general theory of singularities, the two wave-fronts have intersection number $\Delta_a \circ \Delta_b$, the same as the soliton number between a and b if there are no critical values between $W(\gamma_a)$ and $W(\gamma_b)$. We assume for simplicity that all the intersection points are positive. (We will comment on the case where there are also negative intersections.) Then there are $\Delta_a \circ \Delta_b$ gradient flow lines from γ_a to γ_b that map to the straight segment $\overrightarrow{AB}$ in the W-plane. Since this holds for any starting point A, there are $\Delta_a \circ \Delta_b$ families of such paths parametrized by $w_1 := \operatorname{Re} A = \operatorname{Re} B$. The length of x^1 for each path P is given by

$$\Delta x^1 = \left| 2 \int_P \frac{d\operatorname{Im} W}{|\partial W|^2} \right|. \tag{39.296}$$

Only the paths with $\Delta x^1 = \pi$ define the solutions to Eq. (39.295) with the right boundary condition. If the starting point A or the end point B is the critical value $W(p_a)$ or $W(p_b)$, the required length is infinity, $\Delta x^1 = +\infty$. By the assumption on the superpotential, Δx^1 approaches zero when $w_1 = \operatorname{Re} A$ goes to infinity. Thus, for each of the $\Delta_a \circ \Delta_b$ families, Δx^1 is roughly a decreasing function as a function of w_1. If it is a monotonic function, the function shown in Eq. (39.296) cuts through $\Delta x^1 = \pi$ exactly once. However, it is also possible that Δx^1 is not monotonic and cuts through $\Delta x^1 = \pi$ more than once, as depicted in Fig. 11.

We next define the R-charge of the approximate ground states that are peaked at these critical points. This can be done, as before, by first computing the relative Morse index. By the assumption on X and the fact that γ_a and γ_b are contractible, we see that the axial $U(1)$ R-symmetry is anomaly-free. Thus, we expect to have a well-defined integral R-charge. In particular, the mod 2 R-charge and Witten index are well defined. We can compute the Witten index by using the deformation invariance property. Let us deform the theory by rescaling the superpotential as $W \to e^t W$. Let us focus on one of the $\Delta_a \circ \Delta_b$ families of paths. The rescaling of W changes the function given by Eq. (39.296) as $\Delta x^1 \to e^{-t}\Delta x^1$. For an appropriate choice of e^t one can make $e^{-t}\Delta x^1$ cut through π exactly once. Then the

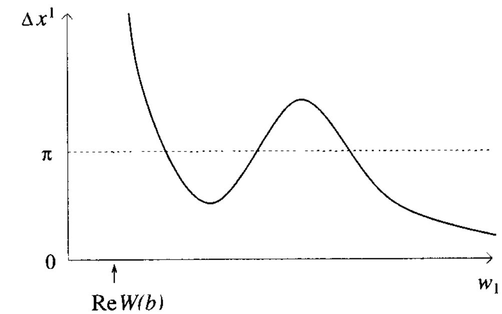


FIGURE 11. A graph of Δx^1 as a function of $w_1 = \operatorname{Re} A = \operatorname{Re} B$. Corresponding to the situation in Fig. 10, the left end is set at $w_1 = \operatorname{Re} W(b)$.

contribution to the Witten index is ± 1. To determine the sign, we need to set the ground of the R-charge. If two families are in a common connected component of $\Omega(\gamma_a, \gamma_b)$, the relative Morse index of the two paths is even. Thus, one can set the sign to be all $+1$ and the total index is $\Delta_a \circ \Delta_b$. (If there were paths corresponding to negative intersections of the wave-fronts, the Morse indices would be odd compared to those for positive intersections. Thus the total Witten index is still $\Delta_a \circ \Delta_b$.)

Finally, we consider the case where $\gamma_a = \gamma_b$. In this case, the critical point of h is the constant map to $p_a = p_b$. Thus, there is a unique supersymmetric ground state. In particular, the Witten index can be set equal to 1.

To summarize, the Witten index is given by

$$\operatorname{Tr}(-1)^F = \begin{cases} \Delta_a \circ \Delta_b & \text{if } \operatorname{Im} W(p_a) > \operatorname{Im} W(p_b), \\ 1 & \text{if } \gamma_a = \gamma_b, \\ 0 & \text{if } \operatorname{Im} W(p_a) < \operatorname{Im} W(p_b). \end{cases} \tag{39.297}$$

Here we have chosen the ground of the R-charges so that the states corresponding to positive intersection have even R-charge; we have shown that one can do so. However, in general the ground states are in various connected components of $\Omega(\gamma_a, \gamma_b)$, and there is no canonical way to set the ground of the R-charge for all of them. Another choice is also possible and may alter the total Witten index. What is invariant is the Witten index in each connected component. One can represent this refined information as a weighted Witten index where the weight distinguishes the homotopy class of the paths.

Let us determine the space of supersymmetric ground states. The problem is non-trivial only for the case $\mathrm{Im}\,W(p_a) > \mathrm{Im}\,W(p_b)$. (As we have seen above, there is none if $\mathrm{Im}\,W(p_a) < \mathrm{Im}\,W(p_b)$, and there is one if $\gamma_a = \gamma_b$.) We keep the assumption that there is no negative intersection between the two wave-fronts at B, for any value of w_1. Then we claim that there are as many supersymmetric ground states as the Witten index $\Delta_a \circ \Delta_b$, all with even R-charges. If the $\Delta_a \circ \Delta_b$ families of paths are all in different components of $\Omega(\gamma_a, \gamma_b)$, one can take the R-charges to be all 0. Let us first suppose that the function Δx^1 cuts through π exactly once in each of the $\Delta_a \circ \Delta_b$ families. Then there are $\Delta_a \circ \Delta_b$ critical points of h with all even R-charges. Since no tunneling is possible between such critical points, the claim follows. Let us now consider the general case where the function Δx^1 cuts through π several times (as in Fig. 11) in some of the $\Delta_a \circ \Delta_b$ families. One way to proceed is to construct the Morse–Witten complex and compute the cohomology. This will lead to the characterization of the space of supersymmetric ground states as the Landau–Ginzburg analogue of the Floer cohomology group

$$\mathcal{H}_{\mathrm{SUSY}} \cong HF^{\bullet}_{W}(\gamma_a, \gamma_b). \tag{39.298}$$

However, there is actually an easy way to find the cohomology. Note that the rescaling $W \to e^t W$, which we used in computing the Witten index, is in fact implemented by a similarity transformation of the supercharge $Q \to e^{-\Delta h} Q\, e^{\Delta h}$, where $\Delta h = (1-e^t)\int_0^\pi dx^1 \mathrm{Re}\,W$. Then the computation reduces to the previous case, and hence the claim follows.

39.4.4. Holomorphic Bundles. We now turn to considering open strings stretched between B-branes. We first take the non-linear sigma model on X and consider D-branes wrapped totally on X and supporting holomorphic vector bundles.

Let E_a and E_b be complex vector bundles on X with Hermitian connections $A^{(a)}$ and $A^{(b)}$. We consider an open string with the boundary interaction as shown by Eq. (39.174), with $A = -A^{(a)t}$ at $s=0$ and $A = A^{(b)}$ at $s=\pi$. We require the curvatures of $A^{(a)}$ and $A^{(b)}$ to have only $(1,1)$-form components so that holomorphic structures are defined on E_a and E_b. Then the open string theory is invariant under B-type supersymmetry generated by $Q = \overline{Q}_+ + \overline{Q}_-$ and $Q^\dagger = Q_+ + Q_-$. Since the boundary interaction in Eq. (39.174) includes the time derivatives of the fields, the Noether charges are modified. The modified supercharge Q is expressed as

$$\begin{aligned} Q = &\int_0^\pi ds \left\{ g_{i\bar{j}}(\overline{\psi}_+^{\bar{j}} + \overline{\psi}_-^{\bar{j}})\partial_0\phi^i - g_{i\bar{j}}(\overline{\psi}_+^{\bar{j}} - \overline{\psi}_-^{\bar{j}})\partial_1\phi^i \right\} \\ &+ (\overline{\psi}_+^{\bar{j}} + \overline{\psi}_-^{\bar{j}})A_{\bar{j}}^{(b)}\Big|_{s=\pi} - (\overline{\psi}_+^{\bar{j}} + \overline{\psi}_-^{\bar{j}})A_{\bar{j}}^{(a)t}\Big|_{s=0}. \end{aligned} \tag{39.299}$$

The supercharges Q and $Q^\dagger$ obey the anti-commutation relations from Eq. (39.242) without any correction, i.e., $\Delta = \mathcal{Z} = 0$ in Eq. (39.243). The system has an exact $U(1)$ R-symmetry F, coming from the bulk vector R-symmetry, with the commutation relation $[F, \psi_\pm] = -\psi_\pm$ and $[F, \overline{\psi}_\pm] = \overline{\psi}_\pm$.

We first quantize the system in the zero mode approximation where we assume all the fields have no s-dependence. From the boundary condition, the left and the right fermionic zero modes are related as $\psi_{-0}^i = \psi_{+0}^i$ and $\overline{\psi}_{-0}^{\bar{i}} = \overline{\psi}_{+0}^{\bar{i}}$. We can identify the quantum mechanical Hilbert space as the space of $(0,p)$-forms with values in $E_a^* \otimes E_b$;

$$\mathcal{H}^{\text{zero mode}} = \bigoplus_{p=0}^{n} \Omega^{0,p}(X, E_a^* \otimes E_b), \tag{39.300}$$

on which the fermionic zero modes act as

$$(\overline{\psi}_{+0}^{\bar{i}} + \overline{\psi}_{-0}^{\bar{i}}) \longleftrightarrow d\overline{z}^{\bar{i}}\wedge, \tag{39.301}$$

$$g_{i\bar{j}}(\psi_{+0}^i + \psi_{-0}^i) \longleftrightarrow i_{\partial/\partial\overline{z}^{\bar{j}}}. \tag{39.302}$$

Note that we only have the combinations $\psi_{-0} + \psi_{+0}$ and $\overline{\psi}_{-0} + \overline{\psi}_{+0}$ because $\overline{\psi}_{+0}^{\bar{j}} - \overline{\psi}_{-0}^{\bar{j}} = 0$ and $\psi_{+0}^i - \psi_{-0}^i = 0$ from the boundary condition. The R-charge counts the form-degree; $F = p$ on $(0,p)$-forms. The supercharge Q corresponds to the Dolbeault operator on the bundle $E_a^* \otimes E_b$:

$$Q \leftrightarrow \overline{\partial}_A = d\overline{z}^{\bar{i}}\left(\partial_{\bar{i}} + A_{\bar{i}}^{(b)} - A_{\bar{i}}^{(a)}\right). \tag{39.303}$$

Thus, in the zero mode approximation, the space of supersymmetric ground states is identified as the Dolbeault cohomology group

$$\mathcal{H}_{\mathrm{SUSY}}^{\text{zero mode}} = \bigoplus_{p=0}^{n} H_{\overline{\partial}_A}^{0,p}(X, E_a^* \otimes E_b). \tag{39.304}$$

In particular, the Witten index is equal to the Dolbeault index

$$\begin{aligned}\mathrm{Tr}(-1)^F &= \sum_{p=0}^{n}(-1)^p \dim H^{0,p}(X, E_a^*\otimes E_b) =: \chi(E_a, E_b)\\ &= \int_X \mathrm{ch}(E_a^*\otimes E_b)\mathrm{Td}(X),\end{aligned} \tag{39.305}$$

where $\mathrm{Td}(X)$ is the total Todd class of the tangent bundle of X (See Sec. 3.5.2). As we will show below, this holds also for the full theory.

Let us come back to the full theory. The space of bosonic field configurations is the space of paths obeying the Neumann boundary condition at both ends:

$$\Omega_N(X) = \left\{ \phi : [0,\pi] \to X \;\middle|\; \partial_s\phi = 0 \text{ at } s = 0, \pi \right\}. \tag{39.306}$$

This space inherits a complex structure from X, where the holomorphic tangent space at $\phi \in \Omega_N(X)$ is isomorphic to the space of sections $v^i(s)$ of ϕ^*T_X that obey the Neumann boundary condition $\partial_s v^i(s) = 0$ at $s = 0, \pi$. The vector $\delta/\delta\phi^j(s')$ corresponds to the section $v^i(s) = \delta^i_j\delta(s-s')$ and belongs to a tangent vector only if $s' \neq 0, \pi$. (For $s' = 0, \pi$ the Neumann boundary condition would not be satisfied.) There are two holomorphic maps, p_0 and p_π, from $\Omega_N(X)$ to X that send a path ϕ to the two end points $p_s[\phi] = \phi(s)$, $s = 0, \pi$. The Hilbert space of states can be considered as the space of anti-holomorphic forms on $\Omega_N(X)$ with values in the bundle $p_0^*E_a^*\otimes p_\pi^*E_b\otimes\bigwedge^\bullet T_{\Omega_N(X)}$:

$$\mathcal{H} = \Omega^{0,\bullet}(\Omega_N(X), p_0^*E_a^*\otimes p_\pi^*E_b\otimes \textstyle\bigwedge^\bullet T_{\Omega_N(X)}), \tag{39.307}$$

where $\bigwedge^\bullet T_{\Omega_N(X)}$ is the exterior powers of the holomorphic tangent bundle $T_{\Omega_N(X)}$. The fermionic fields are identified as the operators

$$(\overline{\psi}^{\,i}_+ + \overline{\psi}^{\,i}_-) \longleftrightarrow \delta\overline{\phi}^{\,i}\wedge, \qquad (\psi^i_+ + \psi^i_-) \longleftrightarrow g^{i\bar{j}}i_{(\frac{\delta}{\delta\overline{\phi}^{\,j}})}, \tag{39.308}$$

$$(\overline{\psi}^{\,i}_+ - \overline{\psi}^{\,i}_-) \longleftrightarrow g^{\bar{j}i}\frac{\delta}{\delta\phi^j}\wedge, \qquad (\psi^i_+ - \psi^i_-) \longleftrightarrow i_{\delta\phi^i}. \tag{39.309}$$

The boundary condition $\psi_- - \psi_+ = \overline{\psi}_- - \overline{\psi}_+ = 0$ at $s = 0$ and π follows from the absence of $\delta/\delta\phi^i(0)$ and $\delta/\delta\phi^i(\pi)$ in the tangent space. Then the supercharge can be identified as the following operator acting on the Hilbert space shown in Eq. (39.307),

$$Q = \overline{\delta}_A + \mathcal{V}\wedge. \tag{39.310}$$

Here, $\overline{\delta}_A$ is the Dolbeault operator on $p_0^*E_a^*\otimes p_\pi^*E_b\otimes\bigwedge^\bullet T_{\Omega_N(X)}$ with respect to the connection induced from that of $E_a^*\otimes E_b$, while $\mathcal{V}$ is the holomorphic vector field on $\Omega_N(X)$ given by

$$\mathcal{V} = \int_{(0,\pi)} ds\,\partial_s\phi^i(s)\frac{\delta}{\delta\phi^i(s)}. \tag{39.311}$$

Thus, the system is an example of the class of supersymmetric quantum mechanics studied in Sec. 10.4.5. As in the discussion there, the Q-cohomology is invariant under the rescaling $\mathcal{V} \to \mathrm{e}^t\mathcal{V}$, and the ground state spectrum is independent of e^t as long as it is finite. A zero energy state remains as a zero energy state in the limit $\mathrm{e}^t \to +\infty$, and therefore the number of ground states is bounded from above by the number of ground states of the limiting theory. But the limiting theory is the theory on the $\mathcal{V} = 0$ locus, namely, the zero mode theory we have studied above. Thus, the space of ground states is bounded from above by the space $\mathcal{H}^{\text{zero mode}}_{\text{SUSY}}$ in Eq. (39.304). It is possible that some of the states in $\mathcal{H}^{\text{zero mode}}_{\text{SUSY}}$ acquire nonzero energies for finite e^t, but that happens always in pairs — together with superpartners. This explains why the Witten index of the full theory remains the same as Eq. (39.305). The pair-lifting of states in $\mathcal{H}^{\text{zero mode}}_{\text{SUSY}}$ cannot occur, however, if the Dolbeault cohomology $H^{0,p}(X, E_a^*\otimes E_b)$ is non-trivial only for even p (or only for odd p). In such a case, the space of supersymmetric ground states of the full theory is indeed the same as the space shown in Eq. (39.304),

$$\mathcal{H}_{\text{SUSY}} \cong \mathcal{H}^{\text{zero mode}}_{\text{SUSY}}. \tag{39.312}$$

The equivalence from Eq. (39.312) holds also when X is a Calabi–Yau manifold. To understand this it is best to consider topological twisting. This will be explained in the next subsection in a similar model.

In general, the index shown in Eq. (39.305) is neither symmetric nor anti-symmetric under the exchange of a and b. This is related by mirror symmetry, as we will discuss later, to the fact noted earlier, that the supersymmetric index for Lagrangian D-branes in LG models is neither symmetric nor anti-symmetric. However, since odd Todd classes are divisible by the first Chern class of X, for a Calabi–Yau manifold $\mathrm{Td}(X)$ is a sum of $4k$-forms. Under the exchange $E_a^*\otimes E_b \to E_b^*\otimes E_a$ the Chern character changes by sign flip in the $(4k+2)$-form components. Thus, for a Calabi–Yau manifold of dimension n, the index is symmetric for even n and anti-symmetric for odd n under the exchange of E_a and E_b.

39.4.5. Holomorphic Branes in Landau–Ginzburg Models. Finally, we consider B-branes in Landau–Ginzburg models. Let X and W be the target space and the superpotential. As we have seen in Sec. 39.2.3, a D-brane wrapped on a complex submanifold of X on which W is locally constant is a B-brane. We study the supersymmetric ground states of the theory of an open string stretched between two such submanifolds Z_a and Z_b. Throughout this discussion, we assume that X is a non-compact Calabi–Yau manifold so that one can consider B-twisting.

The system has B-type supersymmetry with the supercharge expressed as

$$Q = \frac{1}{2\pi}\int dx^1 \Big\{ g_{i\bar{j}}(\overline{\psi}^{\bar{j}}_- + \overline{\psi}^{\bar{j}}_+)\partial_0\phi^i - g_{i\bar{j}}(\overline{\psi}^{\bar{j}}_- - \overline{\psi}^{\bar{j}}_+)\partial_1\phi^i + \frac{i}{2}(\psi^i_- - \psi^i_+)\partial_i W \Big\}. \tag{39.313}$$

It is straightforward to see, by using the canonical (anti-)commutation relations, that the supercharge squares to

$$Q^2 = -\frac{i}{2\pi}\left(W_b - W_a\right) \tag{39.314}$$

where W_a and W_b are the constant values of the superpotential on Z_a and Z_b respectively, $W_a := W|_{Z_a}$, $W_b := W|_{Z_b}$. Thus, the theory loses its cohomological structure unless the values of the superpotential on Z_a and Z_b coincide with each other. In particular, there are no supersymmetric ground states if $W_b \neq W_a$. In what follows, we consider the cases where Z_a and Z_b have the same W-value.

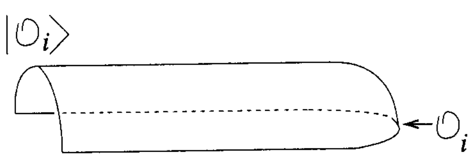


FIGURE 12. Field-State Correspondence

Let us consider an open string ending on the same D-brane, $Z_a = Z_b =: Z$. The supersymmetric ground states are in one-to-one correspondence with the Q-cohomology classes of states. In the present case, the latter are in turn in one-to-one correspondence with the Q-cohomology classes of boundary operators. To see this, let us consider the tongue-shaped worldsheet

(See Fig. 12) where the D-brane boundary condition is imposed on the sides through the tip of the tongue, and the theory is B-twisted in the curved region. We insert an operator $\mathcal{O}_i$ at the tip. Then the state corresponding to $\mathcal{O}_i$ is the one that appears at the back of the tongue. The claimed correspondence follows from the fact that the state $|\mathcal{O}_i\rangle$ is Q-closed (resp. Q-exact) if and only if $\mathcal{O}_i$ is Q-closed (resp. Q-exact). This argument is applicable to the open string in Calabi–Yau with holomorphic bundles, and explains why the equivalence from Eq. (39.312) holds in such a case.

Thus, it is enough to determine the Q-cohomology classes of local boundary operators in the twisted theory. As in the bulk theory, we rename the fields as

$$\eta^{\bar{i}} = -(\overline{\psi}^{\bar{i}}_- + \overline{\psi}^{\bar{i}}_+), \quad \theta_i = g_{i\bar{j}}(\overline{\psi}^{\bar{j}}_- - \overline{\psi}^{\bar{j}}_+), \quad \rho^i_z = \psi^i_-, \quad \rho^i_{\bar{z}} = \psi^i_+. \tag{39.315}$$

We recall that the Q-variation of the fields is given by

$$\begin{aligned} &\delta\phi^i = 0, && \delta\overline{\phi}^{\bar{i}} = \bar{\epsilon}\eta^{\bar{i}}, \\ &\delta\theta_i = \bar{\epsilon}\partial_i W, && \delta\eta^{\bar{i}} = 0. \\ &\delta\rho_\mu = -2\bar{\epsilon}J^\nu_\mu\partial_\nu\phi^i, && \end{aligned} \tag{39.316}$$

The boundary condition is given by

$$\left.\begin{array}{c} \phi \in Z, \\ \rho^i_n : \text{tangent to } Z, \ \ \rho^i_t : \text{normal to } Z, \\ \eta^{\bar{i}} : \text{tangent to } Z, \ \ \theta_i : \text{normal to } Z, \end{array}\right\} \text{ on } \partial\Sigma, \tag{39.317}$$

where ρ^i_n and ρ^i_t are the normal and tangent components of ρ^i_μ on the worldsheet. Since η is tangent to Z, $\eta^{\bar{i}}$ can be identified as $d\overline{z}^{\bar{i}}$ where z^i are the coordinates of Z (not X). The condition on θ can also be stated as $t^i\theta_i = 0$ if $t^i \in T_Z$. This shows that the algebra generated by θ_i can be identified as the exterior algebra of the normal bundle to Z:

$$N_{Z/X} = T_X/T_Z.$$

Now, one can identify the Q-cohomology of the local boundary operators as the cohomology of the operator

$$\overline{\partial} + \partial W. \tag{39.318}$$

acting on

$$\bigoplus_{p,q} \Omega^{0,p}(Z, \textstyle\bigwedge^q N_{Z/X}). \tag{39.319}$$

Here $\partial W\cdot$ stands for the contraction by $dW \in T_X^*$, which makes sense on $N_{Z/X}$ since $dW = 0$ on T_Z (W is a constant on Z). Thus, we find that the space of supersymmetric ground states is given by the cohomology group

$$\mathcal{H}_{\mathrm{SUSY}} \cong H(\Omega^{0,\bullet}(Z, \textstyle\bigwedge^\bullet N_{Z/X}), \overline{\partial} + \partial W\cdot). \tag{39.320}$$

Let us determine the action of the R-symmetry group. Suppose a subgroup Γ of $U(1)$ acts holomorphically on X, $\lambda \in \Gamma : \phi \in X \mapsto R_\lambda(\phi) \in X$, in such a way that

$$W(R_\lambda(\phi)) = \lambda^2 W(\phi), \tag{39.321}$$

and preserving the submanifold Z,

$$R_\lambda(Z) = Z. \tag{39.322}$$

Then there is a (vector) R-symmetry Γ in the theory of an open string ending on Z. The action on the fields is such that $\eta^{\bar{\imath}} \to \lambda R_{\lambda *}{}^{\bar{\imath}}{}_{\bar{\jmath}} \eta^{\bar{\jmath}}$ and $\theta_i \to \lambda R_{\lambda^{-1} *}{}^j{}_i \theta_j$. Thus, the R-symmetry group Γ acts on the space shown in Eq. (39.319) as

$$\omega \mapsto \lambda^{p+q} R_\lambda^* \omega \quad \text{on} \quad \Omega^{0,p}(Z, \textstyle\bigwedge^q N_{Z/X}), \tag{39.323}$$

where R_λ^* acts on $N_{Z/X}$ as $R_{\lambda^{-1}*}$. The supercharge $Q = \overline{\partial} + \partial W\cdot$ increases the R-charge by 1, and hence the Q-complex is graded by the R-charge. This determines the action of the R-symmetry group on the supersymmetric ground states. Depending on the cases, the R-symmetry group Γ can range from the full $U(1)$ group to the $\mathbb{Z}_2$ subgroup (acting trivially on X).

39.4.5.1. *Examples.*

Z does not include any critical point. If Z does not pass through any of the critical points of W, then there are no supersymmetric ground states

$$\mathcal{H}_{\mathrm{SUSY}} = 0. \tag{39.324}$$

This is simply because supersymmetry requires $\partial_i W = 0$ on the boundary, which is impossible to satisfy if Z is away from all the critical points. It is an interesting problem to directly show that the $(\overline{\partial} + \partial W\cdot)$-cohomology vanishes in this case.

Z = a critical point. Let us consider a D0-brane located at a critical point p_* of W, namely, the case with $Z = \{p_*\}$. Then the normal bundle is just the complex vector space $N_{\{p_*\}/X} = \mathbb{C}^n$ and the supercharge is trivial, $Q = 0$. Thus, the cohomology group is simply the exterior algebra

$$\mathcal{H}_{\mathrm{SUSY}} = \textstyle\bigwedge^\bullet \mathbb{C}^n \cong \mathbb{C}^{2^n}. \tag{39.325}$$

With respect to the $\mathbb{Z}_2$ R-symmetry group acting trivially on X, we find that even powers of $\mathbb{C}^n$ are bosonic and odd powers are fermionic,

$$\mathcal{H}^B_{\mathrm{SUSY}} = \textstyle\bigwedge^{even} \mathbb{C}^n, \quad \mathcal{H}^F_{\mathrm{SUSY}} = \textstyle\bigwedge^{odd} \mathbb{C}^n. \tag{39.326}$$

In particular, the Witten index is zero,

$$I(p_*, p_*) = 0. \tag{39.327}$$

Taking the example of the free theory where $X = \mathbb{C}^n$ and $W = \sum_{i=1}^n \Phi_i^2$, one can explicitly quantize the open string and check these results. We leave this as an excercise for the readers.

Z = a line through a critical point. Let us consider the case $X = \mathbb{C}^2 = \{(u,v)\}$ and $W = UV$. The system has a non-degenerate critical point at $u = v = 0$ with critical value $W = 0$, and we consider a submanifold in $W = 0$. There are three smooth submanifolds: $Z_0 = \{(0,0)\}$, $Z_2 = \{v = 0\}$, and $Z_2' = \{u = 0\}$. The first one, Z_0, was already studied above. Thus, we will focus on the second one, $Z = Z_2$ (the third should be similar). It is isomorphic to a complex line $\mathbb{C} = \{u\}$ and the normal bundle $N_{\mathbb{C}/\mathbb{C}^2}$ is spanned by $\partial/\partial v$. The general element of $\Omega^{0,\bullet}(\mathbb{C}, \bigwedge^\bullet N_{\mathbb{C}/\mathbb{C}^2})$ is expressed as

$$f = f_0 + f_{\overline{u}} d\overline{u} + f^v \frac{\partial}{\partial v} + f^v_{\overline{u}} d\overline{u} \otimes \frac{\partial}{\partial v}, \tag{39.328}$$

and the supercharge is given by

$$Q = \overline{\partial} + u dv\cdot\,. \tag{39.329}$$

Here $f_0, f_{\overline{u}}, f^v, f^v_{\overline{u}}$ are normalizable functions of $(u, \overline{u})$ and $\overline{\partial} = d\overline{u}\partial_{\overline{u}}$. We see that

$$Qf = uf^v + (\partial_{\overline{u}} f_0 - uf^v_{\overline{u}}) d\overline{u} + \partial_{\overline{u}} f^v d\overline{u} \otimes \frac{\partial}{\partial v}, \tag{39.330}$$

$$Q^\dagger f = -\partial_u f_{\overline{u}} + (\overline{u} f_0 - \partial_u f^v_{\overline{u}}) \frac{\partial}{\partial u} - \overline{u} f_{\overline{u}} d\overline{u} \otimes \frac{\partial}{\partial v}. \tag{39.331}$$

Finding Q-cohomology classes is the same as solving $Q = Q^\dagger = 0$. We find there is a unique solution to the latter,

$$f = \mathrm{e}^{-|u|^2}\left(1 - d\overline{u} \otimes \frac{\partial}{\partial v}\right). \tag{39.332}$$

With respect to the $\mathbb{Z}_2$ R-symmetry acting trivially on $X = \mathbb{C}^2$, this state is invariant. Thus, the Witten index is

$$I(Z_2, Z_2) = 1. \tag{39.333}$$

The R-charge of the ground state is also trivial with respect to the $U(1)$ R-symmetry $U \to \lambda^{q_U} U$, $V \to \lambda^{q_V} V$ such that $q_U + q_V = 2$.

Since this is a free theory, one can explicitly quantize the open string system and reproduce the above results, including Eq. (39.327) which says $I(Z_0, Z_0) = 0$, and also Eq. (39.333). Moreover, one can also quantize the Z_0-Z_2 and Z_2-Z_2' open strings, and show the following: For the Z_0-Z_2 string, there are two supersymmetric ground states, one bosonic and one fermionic. For the Z_2-Z_2' string, there is a unique supersymmetric ground state. In particular, we have

$$I(Z_0, Z_2) = 0, \tag{39.334}$$
$$I(Z_2, Z_2') = 1. \tag{39.335}$$

We leave the derivation as an exercise for the readers.

This story extends to cases with a larger number of variables. Let us consider $X = \mathbb{C}^{2m} = \{(u_i, v_i)\}$ with the superpotential $W = \sum_{i=1}^m U_i V_i$ that has a non-degenerate critical point at $u_i = v_i = 0$. The submanifold $Z = \{v_1 = \cdots = v_m = 0\}$ has a constant superpotential $W = 0$ and passes through the critical point. The system is the tensor product of m-copies of the previous system of two variables. In particular, there is a unique supersymmetric ground state with trivial R-charge and therefore $I(Z, Z) = 1$.

More General Cases. One can use the above results to determine open string ground states (and Witten index) in more general models. Let us consider the following situation. Z is a middle-dimensional (complex) submanifold of X. For each critical point of W that belongs to Z, Z is given by $Z = \{v_1 = \cdots = v_m = 0\}$ for a coordinate system (u_i, v_i) in which W is expanded as $W = \text{const} + \sum_{i=1}^m U_i V_i + \cdots$. Then near each critical point belonging to Z, we find a unique supersymmetric ground state. Thus, we find as many approximate supersymmetric ground states as the number of such critical points. With respect to the $\mathbb{Z}_2$ R-symmetry acting trivially on the bosonic variables, they all have trivial R-charge. Thus, there is no room for non-perturbative lifting of the ground states. We thus conclude that there are as many supersymmetric ground states as the index

$$I(Z, Z) = \#\Big(\text{critical points in } Z\Big). \tag{39.336}$$

An example of such a case occurs in the LG model with $2m$ variables $X_1, \ldots, X_m$, $Y_1, \ldots, Y_m$ with the superpotential $W = f(X_1, \ldots, X_m) - f(Y_1, \ldots, Y_m)$. Then the "diagonal" $Z = \{X_i = Y_i\}$ satisfies the above condition. Indeed, it passes through the diagonal critical points of f in the desired way. In particular, the number of supersymmetric ground states (and Witten index) equals the number of critical points of f.

39.5. Boundary States and Overlap with RR Ground States

In Sec. 39.1.4, we introduced the notion of boundary states. It was useful in computing open string partition functions, but we also saw that it carries important information about the D-brane. In this section, we study boundary states for D-branes preserving $\mathcal{N}=2$ supersymmetry. We will see that one can learn important topological information about the D-branes even by looking at the coefficients of the RR ground states.

39.5.1. General Aspects. Let a be an A- or B-brane boundary condition in a $(2,2)$ supersymmetric field theory. We study the properties of the corresponding boundary states, $|a\rangle$ for an incoming boundary and $\langle a|$ for an outgoing boundary. First of all, they obey some conditions corresponding to the fact that the boundary conditions preserve A- or B-type supersymmetry. Let $w = x^1 + ix^2$ be the flat coordinate near the boundary, which is located at $x^1 = 0$ where the worldsheet is in the region $x^1 \le 0$ for an incoming boundary, while it is in $x^1 \ge 0$ for an outgoing boundary. This is the coordinate system for the open string. In this coordinate system, the condition of $\mathcal{N}=2$ supersymmetry is that the supercurrents obey $\overline{G}^1_+ + G^1_- = G^1_+ + \overline{G}^1_- = 0$ for an A-brane and $\overline{G}^1_+ + \overline{G}^1_- = G^1_+ + G^1_- = 0$ for a B-brane (for the trivial phases $\mathrm{e}^{i\alpha} = \mathrm{e}^{i\beta} = 1$). In terms of the coordinates for the closed string, $(x^{1'}, x^{2'}) = (x^2, -x^1)$, the conditions are expressed as $\mathrm{e}^{\frac{\pi i}{4}}\overline{G}^{2'}_{+'} + \mathrm{e}^{-\frac{\pi i}{4}}G^{2'}_{-'} = \mathrm{e}^{\frac{\pi i}{4}}G^{2'}_{+'} + \mathrm{e}^{-\frac{\pi i}{4}}\overline{G}^{2'}_{-'} = 0$ for an A-brane and $\mathrm{e}^{\frac{\pi i}{4}}\overline{G}^{2'}_{+'} + \mathrm{e}^{-\frac{\pi i}{4}}\overline{G}^{2'}_{-'} = \mathrm{e}^{\frac{\pi i}{4}}G^{2'}_{+'} + \mathrm{e}^{-\frac{\pi i}{4}}G^{2'}_{-'} = 0$ for a B-brane. Here the phases $e^{\pm\frac{\pi i}{4}}$ come from the spin of the supercurrent, as in the case of free fermions which we have studied in Sec. 39.1.4. This means that the boundary states must obey the condition

$$
\left.\begin{aligned}
\left(\overline{G}^{2'}_{+'} - iG^{2'}_{-'}\right)|a\rangle &= \left(G^{2'}_{+'} - i\overline{G}^{2'}_{-'}\right)|a\rangle = 0,\\
\langle a|\left(\overline{G}^{2'}_{+'} - iG^{2'}_{-'}\right) &= \langle a|\left(G^{2'}_{+'} - i\overline{G}^{2'}_{-'}\right) = 0,
\end{aligned}\right\}\quad\text{for A-brane,} \tag{39.337}
$$

$$
\left.\begin{aligned}
\left(\overline{G}^{2'}_{+'} - i\overline{G}^{2'}_{-'}\right)|a\rangle &= \left(G^{2'}_{+'} - iG^{2'}_{-'}\right)|a\rangle = 0,\\
\langle a|\left(\overline{G}^{2'}_{+'} - i\overline{G}^{2'}_{-'}\right) &= \langle a|\left(G^{2'}_{+'} - iG^{2'}_{-'}\right) = 0,
\end{aligned}\right\}\quad\text{for B-brane.} \tag{39.338}
$$

If the axial or vector $U(1)$ R-symmetry is not broken in the bulk theory, and if the boundary condition preserves that symmetry, the corresponding boundary states must obey additional conditions:

$$J^{2'}_A|a\rangle = 0,\quad \langle a|J^{2'}_A = 0,\quad \text{for A-brane,} \tag{39.339}$$

$$J^{2'}_V|a\rangle = 0,\quad \langle a|J^{2'}_V = 0,\quad \text{for B-brane.} \tag{39.340}$$

In the quantization of the closed strings, the Hermiticity condition is imposed so that $(G^{\mu'}_{\pm'})^\dagger = \overline{G}^{\mu'}_{\pm'}$ (whereas the quantization of open strings would lead to $(G^{\mu}_{\pm})^\dagger = \overline{G}^{\mu}_{\pm}$). Thus, the above conditions on the boundary states are not invariant under Hermitian conjugation. If a is a D-brane preserving A- or B-type supersymmetry with the phase $e^{i\alpha}$ or $e^{i\beta}$, the Hermitian conjugates $\langle\overline{b}|$ and $|\overline{a}\rangle$ correspond to the boundary conditions preserving A- or B-type supersymmetry with the phase $-e^{i\alpha}$ or $-e^{i\beta}$. If the sign flip $(-1)^{F_L}$ of the left-moving worldsheet fermions is a symmetry of the theory, the states $\langle\overline{b}|(-1)^{F_L}$ and $(-1)^{F_L}|\overline{a}\rangle$ correspond to the boundary conditions preserving the A- or B-type supersymmetry with the phase $e^{i\alpha}$ or $e^{i\beta}$, which is the same as the original supersymmetry.

Let a and b be D-brane boundary conditions that preserve the same combinations of the supercharges (A-type or B-type). We can use the boundary states to represent the supersymmetric index for the open string stretched between a and b as

$$I(a,b) = {}_{\rm RR}\langle a|\, e^{-TH(\beta)}|b\rangle_{\rm RR}, \tag{39.341}$$

where ${}_{\rm RR}\langle a|$ and $|b\rangle_{\rm RR}$ are the boundary states in the RR sector. By the basic property of the index, it is independent of the various parameters,

such as β and T. It is an integer and therefore must be invariant under the complex conjugation that induces the replacement $(a,b) \to (\overline{b},\overline{a})$. We note, however, that the latter preserves a different combination of the supercharges compared to the original one.

39.5.2. Overlap with RR ground states. The boundary states are in general the sum of infinitely many eigenstates of the Hamiltonian. Important information on the boundary states can be obtained by looking at the contribution by the supersymmetric ground states, which can be measured by taking the overlap of the boundary state and the ground states. This may be considered as the $\mathcal{N} = 2$ analogue of the boundary entropy discussed in Sec. 39.1.5. In supersymmetric field theory, there are several supersymmetric ground states $|i\rangle$. Thus, we consider

$$
\begin{aligned}
\Pi^a_i &= {}_{\rm RR}\langle a|i\rangle,\\
\widetilde{\Pi}^a_i &= \langle i|a\rangle_{\rm RR}.
\end{aligned} \tag{39.342}
$$

These overlaps encode information of D-brane charge, as we will explain later in more detail.

The open string Witten index can be expressed in terms of these overlaps. Let us look at the formula Eq. (39.341). The index is an integer and in particular is independent of the parameters T and β. Let us expand the time evolution operator in Eq. (39.341) in terms of a complete basis of the closed string states. If we take the limit $T \to \infty$ holding fixed β, only the ground states contribute. Thus, we obtain

$$I(a,b) = \sum_{i\overline{j}} \Pi^a_i g^{i\overline{j}}\, \widetilde{\Pi}^b_{\overline{j}}. \tag{39.343}$$

Here, $|i\rangle$ and $\langle\overline{j}|$ are basis sets and $g^{i\overline{j}}$ is the inverse of the ground state metric $g_{\overline{j}i} = \langle\overline{j}|i\rangle$. The two basis sets are not necessarily related to each other. Suppose there is a subset of D-branes $\{a_1, \dots, a_N\}$, with $N =$ the number of ground states, such that $I(a_i, a_j)$ has an inverse I^{a_j,a_i} and such that $\Pi^{a_j}_i$ is non-degenerate. Then from Eq. (39.343) one can deduce

$$\widetilde{\Pi}^{a_k}_{\overline{j}} I^{a_k a_l} \Pi^{a_l}_i = g_{\overline{j}i}. \tag{39.344}$$

Roughly speaking, one may say $\sum_{k,l} |a_k\rangle I^{a_k a_l}\langle a_l| = 1$ on the space of ground states.

If the axial R-symmetry is unbroken, we see from Eq. (39.337) that the boundary state for an A-brane has zero axial charge. Thus for the overlaps given by Eq. (39.342) to be non-vanishing, the ground state $|i\rangle$ must also have zero axial R-charge. Likewise, if vector R-symmetry exists, the overlaps given by Eq. (39.342) for B-type boundary state are non-vanishing only for the ground state $|i\rangle$ with zero vector R-charge. If the theory has a mass gap, this selection rule is vacuous since all ground states have zero R-charge.

If the vector (resp. axial) R-charge is conserved and integral, there is a one-to-one correspondence between the supersymmetric ground states and the elements of the *ac* ring (resp. *cc* ring), as discussed in Ch. 16. The state $|\phi_i\rangle$ corresponding to a chiral ring element ϕ_i is the one that appears at the boundary S^1 of the semi-infinite cigar Σ with the insertion of ϕ_i at the tip, where the theory is twisted to a topological field theory in the curved region. Thus, for those states, the overlaps ${}_{\rm RR}\langle a|\phi_i\rangle$ can be identified as the path-integral on the semi-infinite cigar where the boundary condition a is imposed at the outgoing boundary and the operator ϕ_i is inserted at the tip. With this choice of basis of the ground states, from Eq. (39.343), we obtain the following expression of the Witten index,

$$I(a,b)=\Pi_i^a\eta^{ij}\widetilde{\Pi}_j^b, \tag{39.345}$$

where η^{ij} is the inverse of the topological metric $\eta_{ij}=\langle i|j\rangle$. One could also use the ground states obtained through anti-topological twisting. If we do that for the $\langle\bar{j}|$ basis in Eq. (39.343), we have another expression $I(a,b)=\Pi_i^a g^{i\bar{j}}\widetilde{\Pi}_{\bar{j}}^b$, where $g^{i\bar{j}}$ is the inverse of the tt^* metric $g_{i\bar{j}}=\langle\bar{j}|i\rangle$.

The character of the overlaps depends crucially on which twist we choose. We discuss the two cases separately, along with the explicit expressions, in several examples.

39.5.3. A-brane with B-twist (B-brane with A-twist). Let us consider an A-brane a in a theory with conserved and integral axial R-charge where B-twist is possible. The operators ϕ_i we use to define the supersymmetric ground states are the *cc* ring elements. The boundary state $\langle a|$ obeys the boundary condition from Eq. (39.337). The overlaps $\Pi_i^a={}_{\rm RR}\langle a|\phi_i\rangle$ are invariant under the twisted F-term deformations of the theory;

$$\frac{\partial\Pi_i^a}{\partial t_{ac}}=0,\ \frac{\partial\Pi_i^a}{\partial\bar{t}_{ac}}=0. \tag{39.346}$$

As for the F-term deformations, the overlaps satisfy the equation

$$\begin{aligned}(\nabla_i\Pi^a)_j&=(D_i\delta_j^k+i\beta C_{ij}^k)\Pi_k^a=0,\\ (\nabla_{\bar{i}}\Pi^a)_{\bar{j}}&=(D_{\bar{i}}\delta_{\bar{j}}^{\bar{k}}-i\beta C_{\bar{i}\bar{j}}^{\bar{k}})\Pi_{\bar{k}}^a=0,\end{aligned} \tag{39.347}$$

where β is the circumference of the boundary circle S^1. Here D_i is the covariant derivative defined in Ch. 17 and C_{ij}^k are the structure constants of the chiral ring. The relations Eqs. (39.346)–(39.347) can be shown by the standard gymnastics in the tt^* equation. The essential point is that the contour integral of the supercurrent bounces back at the boundary of the cigar, with $\overline{G}_\pm$ turned into $\pm iG_\mp$ via the boundary condition shown in Eq. (39.337). Essentially the same relation holds for the other overlaps $\widetilde{\Pi}_i^a=\langle\phi_i|a\rangle$. They do not depend on the twisted F-term deformations, and, for the F-term deformations, satisfy

$$\begin{aligned}(\nabla_i\widetilde{\Pi}^a)_j&=(D_i\delta_j^k-i\beta C_{ji}^k)\widetilde{\Pi}_k^a=0,\\ (\nabla_{\bar{i}}\widetilde{\Pi}^a)_{\bar{j}}&=(D_{\bar{i}}\delta_{\bar{j}}^{\bar{k}}+i\beta C_{\bar{j}\bar{i}}^{\bar{k}})\widetilde{\Pi}_{\bar{k}}^a=0.\end{aligned} \tag{39.348}$$

Note that the sign in front of the chiral ring structure constants is the opposite of Eq. (39.347).

What is said here applies also for the overlap of B-branes and the RR ground states obtained from A-twist on the cigar.

Special Lagrangians in Calabi–Yau: Periods. Let us consider the D-brane wrapped on a special Lagrangian submanifold L of a Calabi–Yau manifold X. We will compute the overlap of the corresponding boundary states and the RR ground states. As the basis of the RR ground states, we use the one corresponding to chiral operators via B-twist. The chiral ring is identified as the classical ring $\oplus_{p,q}H^q(X,\bigwedge^pT_X)$, whose underlying vector space is identified as the space of harmonic forms on X by contraction with the holomorphic n-form Ω. We recall that the topological metric is given by

$$\eta_{IJ}=\int_X\Omega\cdot(\Omega^{-1}\omega_I\wedge\Omega^{-1}\omega_J)\wedge\Omega=(-1)^{n(q_I+1)}\int_X\omega_I\wedge\omega_J, \tag{39.349}$$

where $\Omega\cdot$ is the contraction with Ω and Ω^{-1} is its inverse. q_I is the anti-holomorphic degree of the harmonic form ω_I.

Since special Lagrangians preserve the R-symmetry coming from the bulk axial $U(1)$ R-symmetry, the boundary states have non-trivial overlaps only with the ground states with trivial axial R-charge. They are the ground states corresponding to n-forms. Thus, the only nonzero overlaps are those with the n-form $\Pi_i^L = \langle L|\omega_i\rangle$, where $\{\omega_i\}$ is a basis of $H^n(X)$. As the basis elements, we choose here those with definite Hodge degrees, $\omega_i \in H^{p_i, n-p_i}(X)$. Since L is an A-brane and $|\omega_i\rangle$ are obtained by B-twist, the overlaps are independent of the twisted chiral parameters. In particular, they are independent of the size of X, which is one of the Kähler parameters. Thus, one can exactly compute the overlaps by taking the large volume limit. In this limit, only the constant maps contribute in the cigar path-integral. Then we find

$$\Pi_i^L = c_i \int_L \omega_i, \tag{39.350}$$

$$\widetilde{\Pi}_i^L = \widetilde{c}_i \int_L \omega_i, \tag{39.351}$$

where c_i and $\widetilde{c}_i$ are numerical constants to be determined.

Let us see if the relations Eqs. (39.345)–(39.347)–(39.348) are satisfied. We first note the following. Under the variation of complex structure of X corresponding to a Beltrami-differential $\mu \in H^{0,1}(X, T_X)$, a harmonic (p,q) form ω varies as

$$\delta\omega = \mu \hat{\wedge} \omega + (p,q)\text{-form}, \tag{39.352}$$

where $\hat{\wedge}$ stands for contraction of holomorphic indices and wedge product on anti-holomorphic forms. Thus, $\mu\hat{\wedge}\omega$ is a $(p-1, q+1)$-form. Noting the algebraic relation $\Omega^{-1}(\mu\hat{\wedge}\omega) = (-1)^{p-1}\mu\wedge\Omega^{-1}\omega$, we find that Eqs. (39.347)–(39.348) are satisfied for an appropriate choice of c_i and $\widetilde{c}_i$. Also, again for an appropriate choice of c_i and $\widetilde{c}_i$, the relation given by Eq. (39.345) is satisfied by virtue of the Riemann bilinear identity:

$$\#(L_1 \cap L_2) = \sum_{i,j} \int_{L_1} \omega_i \, \widehat{\eta}^{ij} \int_{L_2} \omega_j, \tag{39.353}$$

where $\widehat{\eta}^{ij}$ is the inverse of $\widehat{\eta}_{ij} = \int_X \omega_i \wedge \omega_j$. The appropriate choices for the constants c_i and $\widetilde{c}_i$ are $c_i = c\, i^{p_i}(-1)^{\frac{p_i(p_i+1)}{2}}$ and $\widetilde{c}_i = \widetilde{c}\, i^{p_i}(-1)^{\frac{p_i(p_i-1)}{2}}$, with $c\widetilde{c} = i^n(-1)^{\frac{n(n+1)}{2}}$.

The overlap Π_0^L (or $\widetilde{\Pi}_0^L$) for the ground state corresponding to the holomorphic n-form is the standard period integral

$$\Pi_0^L = \int_L \Omega. \tag{39.354}$$

For a Calabi–Yau threefold, other overlaps can be obtained from this by variation of the complex structure. The Eq. (39.347) is then the first order form of the Picard–Fuchs differential equations on the period. The quantity $\Pi_0^L = \int_L \Omega$ is of special importance from space-time physics. It is the central charge of the space-time supersymmetry algebra, and its absolute value is the mass of the D-brane wrapped on L.

A-Branes in LG: Weighted period. Let us consider the LG model with the target space X and the superpotential W. We assume X to be a (non-compact) Calabi–Yau manifold so that B-twist is possible. We consider A-branes wrapped on the wave-front trajectories γ_a passing through the critical points p_a of W. As in the sigma model, by taking the large volume limit we can exactly compute the overlap of the A-brane boundary states and the ground states obtained via B-twist (i.e., those corresponding to chiral ring elements). In this limit, the theory reduces to the quantum mechanics where the Hilbert space of states are differential forms on X and the supercharges are realized as

$$\overline{Q}_+ = \overline{\partial} - \frac{i\beta}{2}\partial W\wedge, \quad Q_+ = *(-\partial + \frac{i\beta}{2}\overline{\partial W}\wedge)*, \tag{39.355}$$

$$Q_- = \partial + \frac{i\beta}{2}\overline{\partial W}\wedge, \quad \overline{Q}_- = *(-\overline{\partial} - \frac{i\beta}{2}\partial W\wedge)*\,. \tag{39.356}$$

There is a β dependence here because we are quantizing the system on the circle of circumference β. As noted in Sec. 10.4.4 and 10.4.5, the ground states are the middle-dimensional forms ω_i on X. We note that the topological metric is given by $\eta_{ij} = \int_X \omega_i \wedge *\omega_j$. Taking into account the shift of the action by the boundary term shown in Eq. (39.293), we find that Π_i^a and $\widetilde{\Pi}_i^a$ are the integrals over γ_a of $e^{-i\beta(W-\overline{W})/2}\omega_i$ and $e^{i\beta(W-\overline{W})/2} * \omega_i$ respectively. Note that the integrands are closed forms since the ground state wave-functions ω_i are annihilated by the supercharges Eqs. (39.355)–(39.356). In particular, the integrals are invariant under deformation of the integration region. We thus deform γ_a into $\gamma_a^{\mp}$ in such a way that the image in the W-plane is rotated by the small phases $e^{\mp i\epsilon}$ around $W(p_a)$, so that the exponential factors $e^{\mp i\beta(W-\overline{W})/2}$ are small at infinity. We thus obtain the expressions

$$\Pi_i^a = \int_{\gamma_a^-} e^{-i\beta(W-\overline{W})/2}\omega_i, \tag{39.357}$$

$$\widetilde{\Pi}_i^a = \int_{\gamma_a^+} e^{i\beta(W-\overline{W})/2} * \omega_i. \tag{39.358}$$

The relation given by Eq. (39.345) can be derived by using Riemann's bilinear identity and Poincaré duality in the *relative* (co)homology theory. To see this, we first note that the submanifold $\gamma_a^{\pm}$ has a boundary in the region $B_\pm$ of X where $\pm\mathrm{Im}\, W$ is large. Namely, it defines the relative homology class $[\gamma_a^{\pm}] \in H_n(X, B_\pm)$. The Poincaré dual P.D.$[\gamma_a^+]$ of $[\gamma_a^+]$ belongs to the dual space of $H_n(X, B_-)$, which is the cohomology group $H^n(X, B_-)$. This latter group is spanned by the basis vectors $e^{-i\beta(W-\overline{W})/2}\omega_i$. Thus the Poincaré dual is expanded as P.D.$[\gamma_a^+] = e^{-i\beta(W-\overline{W})/2}\omega_i c^i$. The coefficients c^i can be determined by taking the product with $e^{i\beta(W-\overline{W})/2} * \omega_j$ and integrating over X;

$$\int_X \mathrm{P.D.}[\gamma_a^+] \wedge e^{i\beta(W-\overline{W})/2} * \omega_j = c^i \int_X \omega_i \wedge *\omega_j. \tag{39.359}$$

By using $\int_X \omega_i \wedge *\omega_j = \eta_{ij}$, we find that the coefficients are

$$c^i = \eta^{ij} \int_{\gamma_a^+} e^{i\beta(W-\overline{W})/2} * \omega_j.$$

Thus, we find

$$\begin{aligned} \#(\gamma_a^- \cap \gamma_b^+) &= \int_{\gamma_a^-} \mathrm{P.D.}[\gamma_b^+] \\ &= \int_{\gamma_a^-} e^{-i\beta(W-\overline{W})/2}\omega_i\, \eta^{ij} \int_{\gamma_b^+} e^{i\beta(W-\overline{W})/2} * \omega_j. \end{aligned} \tag{39.360}$$

The right-hand side is just $\Pi_i^a \eta^{ij} \widetilde{\Pi}_j^b$. The left-hand side is zero if $\mathrm{Im}\, W(p_a) < \mathrm{Im}\, W(p_b)$ since the images in the W-plane do not intersect: $W(\gamma_a^-)$ is rotated clockwise and $W(\gamma_b^+)$ is rotated counterclockwise, from the 3-o'clock direction. On the contrary, if $\mathrm{Im}\, W(p_a) > \mathrm{Im}\, W(p_b)$, the W-images do intersect and the cycles γ_a^- and γ_b^+ may also intersect. The intersection number is $(-1)^{\frac{n(n-1)}{2}} \Delta_a \circ \Delta_b$. Finally, if $\gamma_a = \gamma_b$, the two cycles γ_a^- and γ_a^+ intersect transversely at the critical point p_a, with the sign $(-1)^{\frac{n(n-1)}{2}}$. Thus, the left-hand side of Eq. (39.360) is nothing but the Witten index $I(a,b)$ up to a sign (See Eq. (39.297)). The relation (39.345) is

thus shown to hold. A more general relation (39.343) can also be shown by replacing ω_i in (39.358) by another basis $\omega_{\bar{i}}$ of ground state wave-functions.

Finding the ground state wave-functions is a difficult problem. However, computation of the overlaps Eqs. (39.357)–(39.358) does not require the full information of ω_i. It is easy to see that Π_i^a depends only on the $(\overline{Q}_+ + Q_-)$-cohomomology class of ω_i, while $\widetilde{\Pi}_{\bar{i}}^a$ depends only on the $(Q_+ + \overline{Q}_-)$-cohomology class of $\omega_{\bar{i}}$. Note that the ground states can be written as

$$\omega_i = e^{-i\beta\overline{W}/2}\Omega_i + (\overline{Q}_+ + Q_-)\alpha_i, \tag{39.361}$$

$$\omega_{\bar{i}} = e^{-i\beta W/2}\overline{\Omega}_{\bar{i}} + (Q_+ + \overline{Q}_-)\beta_{\bar{i}}, \tag{39.362}$$

where Ω_i (resp. $\overline{\Omega}_{\bar{i}}$) are middle-dimensional holomorphic (resp. anti-holomorphic) forms. Then the overlaps can be expressed as

$$\Pi_i^a = \int_{\gamma_a^-} e^{-i\beta W/2}\Omega_i, \tag{39.363}$$

$$\widetilde{\Pi}_{\bar{i}}^a = \int_{\gamma_a^+} e^{-i\beta\overline{W}/2} * \overline{\Omega}_{\bar{i}}. \tag{39.364}$$

This formulation will be useful for practical purposes. In particular, we will use it when we study D-branes in $\mathcal{N}=2$ minimal models. For convenience, we record here the expression of the inner product of the ground states

$$g_{i\bar{j}} := \int \omega_i \wedge *\omega_{\bar{j}} = \int e^{-i\beta(W+\overline{W})/2}\Omega_i \wedge *\overline{\Omega}_{\bar{j}}. \tag{39.365}$$

This is derived by using the fact that $(\overline{Q}_+ + Q_-)^\dagger = (Q_+ + \overline{Q}_-)$.

Holomorphic Bundles (approximate). We next consider the D-brane wrapped on X and supporting a holomorphic vector bundle E over X. As the ground states, we take the ones corresponding to homology cycles C_i (or its Poincaré dual $[C_i]$) obtained via A-twist. Since E is a B-brane and the ground states $|[C_i]\rangle$ are obtained from A-twist, the overlaps $\Pi_i^E = \langle E|[C_i]\rangle$ are expected to depend on the size of X. Nevertheless, we will compute their behaviour in the large volume limit. Then the computation reduces to the quadratic approximation around the constant maps. This leads to the following approximate expressions.

$$\Pi_i^E = \int_{C_i} e^{B+i\omega}\mathrm{ch}(E^\vee)\sqrt{\mathrm{Td}(X)} + \cdots, \tag{39.366}$$

$$\widetilde{\Pi}_i^E = \int_{C_i} e^{-B-i\omega}\mathrm{ch}(E)\sqrt{\mathrm{Td}(X)} + \cdots, \tag{39.367}$$

where $+\cdots$ are contributions from non-constant maps. B is the B-field and ω is the Kähler form. Below, we explain how these expressions come about, and also provide some consistency checks.

The parts $\mathrm{ch}(E^\vee)$ or $\mathrm{ch}(E)$ simply emerge from the (Euclidean) path-integral of the quantum mechanics living on the boundary circles, precisely as in the discussion of Ch. 10. $\mathrm{e}^{\pm(B+i\omega)}$ comes out because of the absence of fermionic zero modes. By the parallel section condition shown in Eq. (39.347), the overlaps should obey $\partial_i\Pi^E_j \sim iC^k_{ij}\Pi^E_k$ and $\partial_i\widetilde{\Pi}^E_j \sim -iC^k_{ij}\widetilde{\Pi}^E_k$ for Kähler deformations, where C^k_{ij} are the matrix elements of the (quantum) wedging by the Kähler form ω_i,

$$C^k_{ij} = (\omega_i\wedge)^k_j + \text{instanton correction.} \tag{39.368}$$

The above expression for Π^E_i and $\widetilde{\Pi}^E_i$ is consistent with this. Also, we note that the B-field and the curvature F of the vector bundle enter into the expressions in the combination

$$B - \frac{i}{2\pi}F, \tag{39.369}$$

as they should. Finally, the relation given by Eq. (39.345) is satisfied by

$$\chi(E_1,E_2) = \sum_{i,j}\int_{C_i}\mathrm{e}^{B+i\omega}\mathrm{ch}(E_1^\vee)\sqrt{\mathrm{Td}(X)}\,\eta^{ij}\int_{C_j}\mathrm{e}^{-B-i\omega}\mathrm{ch}(E_2)\sqrt{\mathrm{Td}(X)}, \tag{39.370}$$

again by Riemann's identity.

39.5.4. A-brane with A-twist (B-brane with B-twist). Let us consider an A-brane a in a theory with conserved and integral vector R-charge where A-twist is possible. As the supersymmetric ground states, we take those corresponding via A-twist to the ac ring elements ϕ_i. In this case, the path-integral can be regarded as the path-integral for the topological field theory (A-model). By the standard argument, or again by the tt^* gymnastics, it is easy to show that the overlap is independent of the F-term deformations

$$\frac{\partial\Pi^a_i}{\partial t_{cc}} = 0,\ \frac{\partial\Pi^a_i}{\partial\bar{t}_{cc}} = 0, \tag{39.371}$$

and depends holomorphically on the twisted F-term deformations,

$$\frac{\partial\Pi^a_i}{\partial\bar{t}_{ac}} = 0. \tag{39.372}$$

The above applies also for the overlap of B-branes and the RR ground states obtained from B-twist on the cigar.

Lagrangian Submanifolds. Let us consider the non-linear sigma model on X and A-branes wrapped on Lagrangian submanifolds. We study the overlap of the boundary state $\langle L|$ and the RR ground states $|\omega_i\rangle$ corresponding to the cohomology classes $\omega_i\in H^*(X)$ obtained from A-twist on the cigar. The computation is done in the topological A-model, and the overlap is simply the disc amplitude with an insertion of $\mathcal{O}_{\omega_i}$ in the interior. This can be expressed as

$$\Pi^L_i = \sum_{\Phi\in H_2(X,L)} n_\Phi(\omega_i)\,\mathrm{e}^{-\int_\Phi\omega} \tag{39.373}$$

where $n_\Phi(\omega_i)$ is the number of holomorphic maps of the disc D^2 to X with the boundary ∂D^2 mapped to L, and with the origin mapped into the Poincaré dual of ω_i.

Holomorphic Branes in LG. As the final example, we consider B-branes in the LG model on a non-compact Calabi–Yau manifold X with superpotential W. We assume that all the critical points are nondegenerate. We choose a B-brane Z which is a complex submanifold of X on which W is a constant. Recall that the bulk chiral ring is the ring of functions on X subject to the constraint $\partial_i W = 0$, and its dimension is equal to the number N of the critical points. We denote by O_f the operator corresponding to a function f. The overlap $\Pi^Z_f = \langle Z|O_f\rangle$ is the disc topological amplitude with O_f insertion in the interior, and with the D-brane boundary condition corresponding to Z. As in the bulk theory, by the localization principle, the path-integral picks up contributions only at the constant maps to the critical points. Let us focus on one such point p_a. The nonzero mode integrals in the quadratic approximation at that point yields 1, as a consequence of the boson/fermion cancellation. Thus, we consider zero mode counting. For bosons ϕ^i, the zero modes are tangent directions to Z. For the fermions, $\theta_i = g_{i\bar{j}}(\overline{\psi}^{\bar{j}}_- - \overline{\psi}^{\bar{j}}_+)$, $\eta^{\bar{i}} = -(\overline{\psi}^{\bar{i}}_- + \overline{\psi}^{\bar{i}}_+)$, $\rho^i_z = \psi^i_-$, and $\rho^i_{\bar{z}} = \psi^i_+$, the respective zero modes are constant modes, constant modes, anti-holomorphic 1-forms, and holomorphic 1-forms that obey the D-brane boundary conditions. The D-brane boundary conditions shown in Eq. (39.317) dictate that $\eta^{\bar{i}}$ is tangent to Z, θ_j are normal to Z, and ρ^i_z and $\rho^i_{\bar{z}}$ are related by $\rho^i_z = \pm\rho^i_{\bar{z}}$ where $\pm$ depends on whether the index i is tangent or normal to Z. The number of ρ^i_μ zero modes is 1 (for a constant) minus the index shown in Eq. (39.219) for trivial $(E, E_{\mathbb{R}})$, which is $1-(2-2g-h)=2g+h-1$ on the worldsheet of genus g and hole h. Since we are considering the disc, $g=0, h=1$, the number is zero. Thus, (the contribution from p_a to) the amplitudes are

$$\int d\phi_{\rm t} d\eta_{\rm n} d\theta_{\rm t} \exp\left(-|dW|^2 - \partial_{\bar{i}}\partial_{\bar{j}}\overline{W}\eta^{\bar{i}} g^{\bar{j}i}\theta_i\right) f(p_a), \tag{39.374}$$

where the subscript "t" or "n" stands for tangent or normal. Note that θ and η are paired in the exponent. This means that the amplitude is vanishing unless the numbers of tangent and normal directions are the same, namely, unless Z is a middle-dimensional submanifold of X;

$$\Pi_f^Z = 0 \quad \text{if} \quad \dim Z \neq \frac{1}{2}\dim X. \tag{39.375}$$

If Z is middle dimensional, $\dim Z = m$ with $\dim X = 2m$, the fermion integral yields the determinant of the $m \times m$ matrix $\partial_{\bar{\imath}_{\rm t}}\partial_{\bar{\jmath}_{\rm n}}\overline{W}$ where $\bar{\imath}_{\rm t}$ runs over tangent indices and $\bar{\jmath}_{\rm n}$ runs over normal indices. The boson integral, on the other hand, yields the inverse of the determinant of the $m \times m$ matrix $\partial_{i_{\rm t}}\partial_k W g^{k\bar{l}}\partial_{\bar{l}}\partial_{\bar{\jmath}_{\rm t}}\overline{W}$. Since W is a constant on Z, we note that any number of tangent derivatives of W is zero on Z. Thus, the determinant is simply $|\det \partial_{i_{\rm t}}\partial_{j_{\rm n}} W|^2$. Combining the two, the cancellation of the anti-holomorphic determinant occurs, and we obtain simply $\det^{-1}\partial_{i_{\rm t}}\partial_{j_{\rm n}} W$. One might wonder if it depends on the choice of the coordinates. One can fix this ambiguity by using the holomorphic n-form Ω, and defining

$$\mathrm{Pf}^Z \partial_i\partial_j W := \Omega^{i_1\cdots i_m j_1^{\rm n}\cdots j_m^{\rm n}} \partial_{i_1}\partial_{j_1^{\rm n}} W \cdots \partial_{i_m}\partial_{j_m^{\rm n}} W. \tag{39.376}$$

Here $j_\alpha^{\rm n}$ are the normal indices, that is, we have chosen the coordinate system in such a way that Z is given by $z^{j_1^{\rm n}} = \cdots = z^{j_m^{\rm n}} = 0$. Thus, we have obtained the following expression for the overlap:

$$\Pi_f^Z = (-1)^m \widetilde{\Pi}_f^Z = \sum_{p_a:\,\text{crit pt in } Z} \frac{f(p_a)}{\mathrm{Pf}^Z \partial_i\partial_j W(p_a)}. \tag{39.377}$$

The factor $(-1)^m$ for $\widetilde{\Pi}_f^Z$ is a convention. Now one can check the relation Eq. (39.345) for the cases where the Witten index has been computed, and for other cases one can compute the Witten index using Eq. (39.345). First of all, one can deduce

$$I(Z_1, Z_2) = 0 \quad \text{if } \dim Z_1 \text{ or } \dim Z_2 \neq \tfrac{1}{2}\dim X. \tag{39.378}$$

This agrees, for example, with Eq. (39.327). For computation in the case $\dim Z_1 = \dim Z_2 = \frac{1}{2}\dim X$, it is convenient to choose the point basis for the bulk chiral ring. Let ϵ_a be holomorphic functions on X such that $\epsilon_a(p_b) = \delta_{ab}$. For such a basis we have

$$\eta_{ab} = \sum_{p_c:\,\text{crit pt}} \frac{\epsilon_a(p_c)\epsilon_b(p_c)}{\det \partial_i\partial_j W(p_c)} = \frac{\delta_{ab}}{\det \partial_i\partial_j W(p_a)}, \tag{39.379}$$

or $\eta^{ab} = \delta_{ab}\det\partial_i\partial_j W(p_a)$. The Hessian of W that appears here is related to Eq. (39.376) as $\det\partial_i\partial_j W = (-1)^m\left(\mathrm{Pf}^Z\partial_i\partial_j W\right)^2$. Then we find

$$\begin{aligned}
I(Z_1, Z_2) &= \Pi_a^{Z_1}\eta^{ab}\widetilde{\Pi}_b^{Z_2} \\
&= \sum_{p_c:\,\text{crit pt in } Z_1} \frac{\delta_{ac}}{\mathrm{Pf}^{Z_1}\partial_i\partial_j W(p_c)}\delta_{ab}\det\partial_i\partial_j W(p_a) \\
&\qquad \times \sum_{p_d:\,\text{crit pt in } Z_2} \frac{(-1)^m\delta_{bd}}{\mathrm{Pf}^{Z_2}\partial_i\partial_j W(p_d)} \\
&= \sum_{p_a:\,\text{crit pt in } Z_1\cap Z_2} \frac{(-1)^m \det\partial_i\partial_j W(p_a)}{\mathrm{Pf}^{Z_1}\partial_i\partial_j W(p_a)\mathrm{Pf}^{Z_2}\partial_i\partial_j W(p_a)} \\
&= \sum_{p_a:\,\text{crit pt in } Z_1\cap Z_2} (\pm 1)_a.
\end{aligned}$$

In the case $Z_1 = Z_2$, the signs are all positive and thus we have $I(Z, Z) = \#(\text{critical points in } Z)$, reproducing Eq. (39.336).

39.6. D-Brane Charge and Monodromy

In superstring theory, the closed string supersymmetric ground states generate space-time gauge potentials, called Ramond-Ramond potentials. The overlaps of the boundary states and the RR ground states thus measure the charges of the D-brane with respect to the RR potentials. In other words, a D-brane with a non-trivial overlap is a soure of RR gauge fields. Even if the bulk theory does not give rise to a superstring background and the space-time interpretation is absent, the overlaps introduce a particular *linear relation* among the totality of D-branes. In any case, we can always talk about the "D-brane charge" and the *charge lattice* $\Lambda_{\rm D}$ of D-branes. The charge lattice is roughly the dual of the space of supersymmetric ground states. The relation from Eq. (39.343) shows that the Witten index $I(a, b)$ for the a-b open string depends only on the charges of the D-branes a and b. Namely, the Witten index defines an integral bilinear form on $\Lambda_{\rm D}$.

We have seen that the overlaps obey the Eq. (39.347) or (39.348) with respect to the variation of the parameters of the bulk theory. This means that they are parallel sections of the "improved connection" ∇ of the vacuum bundle over the parameter space $\mathcal{M}$ of the bulk theory. Let $\{\gamma_a\}$ be a set of D-branes such that the overlaps $\Pi^a = (\Pi_i^{\gamma_a})$ span the space of parallel sections. Then for any D-brane γ, the overlap Π^γ can be expressed as a linear combination of Π^a's, with the coefficients being constant (i.e., invariant under the deformation). Let us consider a contour $\mathcal{C}$ in the parameter space $\mathcal{M}$, starting from and ending at the same point $P \in \mathcal{M}$. After travelling along the contour, the overlaps Π^a change into something else, $\Pi^a|_{\mathcal{C}}$ in general, but remain as the parallel sections of the improved connection. Thus, they can be spanned by the old basis:

$$\Pi^a|_{\mathcal{C}} = (M_{\mathcal{C}})^a_{\ b}\Pi^a. \tag{39.380}$$

The matrix $M_{\mathcal{C}}$ is called the monodromy along $\mathcal{C}$ for the system of Eq. (39.347). It is the holonomy of the (flat) improved connection ∇ along $\mathcal{C}$. It is also interpreted as the monodromy of the D-brane charge. The same thing can be said for the other overlaps $\widetilde{\Pi}^a$. Because of the interpretation as the monodromy of D-brane charge, the monodromy matrix should be the same as for Π_a. Thus, we have

$$\widetilde{\Pi}^a|_{\mathcal{C}} = (M_{\mathcal{C}})^a_{\ b}\widetilde{\Pi}^a. \tag{39.381}$$

The open string Witten index is monodromy invariant:

$$(M_{\mathcal{C}})^a_{\ c} I(c,d)(M_{\mathcal{C}})^b_{\ d} = I(a,b). \tag{39.382}$$

This follows from the fact that $I(a,b)$ is deformation invariant, or by using the bilinear relation given by Eq. (39.343).

Below, we present the charge lattice $\Lambda_{\rm D}$ and the monodromy for the three classes of D-branes considered in Sec. 39.5.3.

39.6.1. Special Lagrangians in Calabi–Yau. Let us consider special Lagrangian D-branes in a Calabi–Yau manifold X. The charge lattice is the middle dimensional homology group

$$\Lambda_{\rm D} = H_n(X;\mathbb{Z}). \tag{39.383}$$

The sign flip in the lattice corresponds to the orientation flip. The bilinear form is the intersection form, which is symmetric for even n and anti-symmetric for odd n.

The overlaps of the boundary states and the RR ground states are parallel sections of the vacuum bundle over the moduli space $\mathcal{M}_c$ of complex

structures. (The dependence on the Kähler moduli is trivial, as we have seen in Eq. (39.346).) The moduli space $\mathcal{M}_c$ contains a discriminant locus Δ where X develops a singularity. Let us consider a non-trivial contour $\mathcal{C}$ in $\mathcal{M}_c \setminus \Delta$. Under the transport along $\mathcal{C}$, the middle dimensional homology group $H_n(X,\mathbb{Z})$ undergoes an integral linear transformation $[\gamma_a] \to [\gamma_b]M^b_{\ a}$. The matrix M is clearly identified as the monodromy matrix $M_{\mathcal{C}}$. On a discriminant locus of complex codimension one, what happens typically is that an n-dimensional cycle γ_V shrinks to zero size. Then the monodromy along a contour $\mathcal{C}$ encircling such a locus is given by

$$[\gamma] \mapsto [\gamma] + \#(\gamma\cap\gamma_V)[\gamma_V]. \tag{39.384}$$

For the case $\dim X = 3$, the system of Eq. (39.347) is identified as the first-order form of the Picard-Fuchs equation. Thus, the monodromy $M_{\mathcal{C}}$ can also be considered as the monodromy of the Picard-Fuchs equation.

39.6.2. A-Branes in Landau–Ginzburg. Let us next consider A-branes in a Landau–Ginzburg model. They are wave-front trajectories γ_a emanating from the critical points p_a and extending in the positive real directions in the image in the W-plane. The charge lattice is the relative homology group

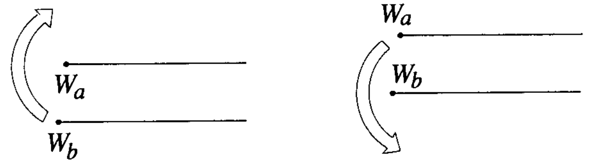
$$\Lambda_{\rm D} = H_n(X, B_W;\mathbb{Z}). \tag{39.385}$$

where B_W is a region in X such that ${\rm Re}\, W$ is large. The homology classes of γ_a provide a distinguished basis of $\Lambda_{\rm D}$ for any values of the parameters of the theory. Thus, we can consider the problem of monodromy even between different points of the parameter space. The bilinear form is given by

$$I([\gamma_a],[\gamma_b]) = \#(\gamma_a^-\cap\gamma_b^+), \tag{39.386}$$

where $\gamma^\pm$ is defined as in Sec. 39.5.3. This is neither symmetric nor anti-symmetric.

To see the monodromy action, we look at a situation where some change is expected. Let γ_a and γ_b be wave-front trajectories emanating from critical points p_a and p_b, where ${\rm Im}\, W(p_a)$ is slightly larger than ${\rm Im}\, W(p_b)$. By (39.297), the Witten index is given by $I(\gamma_a,\gamma_b) = \#(\gamma_a^-\cap\gamma_b^+) = \Delta_a\circ\Delta_b$ while $I(\gamma_b,\gamma_a) = 0$. Let us change the parameters of the superpotential $W \to W'$ so that the critical values move as in Fig. 13(L). After the move, the relation of the critical values is reversed, ${\rm Im}\, W'(p_b') > {\rm Im}\, W'(p_a')$. Let



(L) (R)

FIGURE 13. Moves of Critical Values

us denote by γ_a' and γ_b' the wave-front trajectories emanating from p_a' and p_b'. By Picard–Lefschetz theory, the homology classes of γ_a and γ_b turn into the following classes in $H_n(X, B_{W'}; \mathbb{Z})$:

$$\text{Fig. 13(L)}: \quad \begin{aligned} &[\gamma_a] \longrightarrow [\gamma_a'], \\ &[\gamma_b] \longrightarrow [\gamma_b'] + (\Delta_a \circ \Delta_b)\,[\gamma_a']. \end{aligned} \tag{39.387}$$

We may describe the latter as "creation" of the brane γ_a' of multiplicity $(\Delta_a \circ \Delta_b)$ when p_a passes through γ_b. For the index to be invariant along the move, the Witten index for the branes γ_a' and γ_b' must be given by $I(\gamma_a', \gamma_a') = I(\gamma_b', \gamma_b') = 1$, $I(\gamma_a', \gamma_b') = 0$ and $I(\gamma_b', \gamma_a') = -(\Delta_a \circ \Delta_b)$, which indeed obeys the general structure as in Eq. (39.297). Similarly, for the move depicted in Fig. 13(R), the homology classes transform as

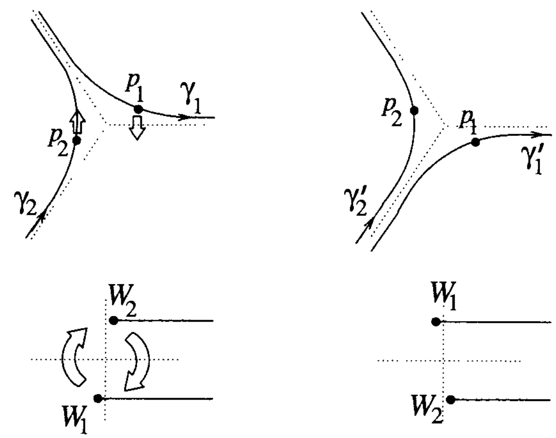
$$\text{Fig. 13(R)}: \quad \begin{aligned} &[\gamma_a] \longrightarrow [\gamma_a''] + (\Delta_a \circ \Delta_b)\,[\gamma_b''], \\ &[\gamma_b] \longrightarrow [\gamma_b''], \end{aligned} \tag{39.388}$$

where $[\gamma_a'']$, $[\gamma_b'']$ are the basis after the move. The monodromy of any other move can be considered as a combination of these simple moves. Therefore, the information of the monodromy is completely determined by the data $\Delta_a \circ \Delta_b$, or the soliton numbers.

For illustration, let us consider the single-variable theory with the superpotential

$$W = \frac{1}{3}\Phi^3 - q\Phi. \tag{39.389}$$

The region B is the union of the directions $\phi \in \mathbb{R}_+$, $e^{2\pi i/3}\mathbb{R}_+$, $e^{-2\pi i/3}\mathbb{R}_+$ for any value of q. The critical points are at $\phi_1 = \sqrt{q}$ and $\phi_2 = -\sqrt{q}$ with the critical values $W(\phi_1) = -\frac{2}{3}q^{3/2}$ and $W(\phi_2) = \frac{2}{3}q^{3/2}$. We start with the



(1) (2)

FIGURE 14. The wave-front trajectories

value $q = e^{i\epsilon}$ where the cycles γ_1 and γ_2 are depicted in Fig. 14(1). We note that $\Delta_2 \circ \Delta_1 = \gamma_2^- \cap \gamma_1^+ = -1$. We now decrease the phase of q so that $q = e^{-i\epsilon}$. The new wave-front trajectories are depicted in Fig. 14(2). One can explicitly see that $\gamma_2 \to \gamma_2'$ and $\gamma_1 \to \gamma_1' - \gamma_2'$, which obeys Eq. (39.387).

39.6.3. R-charge and Monodromy. As an application, we show how the ground state R-charge of a *conformal* LG model is related to the monodromy matrix of the D-branes, which in turn is expressed in terms of the soliton numbers in a massive deformation of the theory. Let us consider a homogeneous superpotential $W(\Phi)$ with respect to some one-parameter family of holomorphic maps f_t, $W(f_t\Phi) = e^{2it}W(\Phi)$. Then the LG model has a vector R-symmetry given by the transformation

$$e^{i\alpha F_V} : \Phi(x^\mu, \theta^\pm, \overline{\theta}^\pm) \to f_\alpha \Phi(x^\mu, e^{-i\alpha}\theta^\pm, e^{i\alpha}\overline{\theta}^\pm). \tag{39.390}$$

The theory is expected to flow in the IR limit to a non-trivial superconformal field theory. The vector R-symmetry rotates the superpotential as $e^{i\alpha F_V} W(\phi)\, e^{-i\alpha F_V} = e^{2i\alpha}W(\phi)$. Thus, if $|\gamma\rangle$ is the boundary state corresponding to a D-brane γ with $W|_\gamma \in \mathbb{R}$, then $e^{-i\alpha F_V}|\gamma\rangle$ is the boundary state corresponding to the D-brane $f_\alpha\gamma$ with $W|_{f_\alpha\gamma} \in e^{2i\alpha}\mathbb{R}$. This can also be understood by noting that

$$\mathrm{e}^{-i\alpha F_V}(\overline{Q}_+ + Q_-)\,\mathrm{e}^{i\alpha F_V} = \mathrm{e}^{-i\alpha}\overline{Q}_+ + \mathrm{e}^{i\alpha}Q_- =: Q_A^\alpha, \tag{39.391}$$

and that the W-image of a D-brane preserving this supercharge must be parallel to $\mathrm{e}^{2i\alpha}\mathbb{R}$. In particular, as we change α from 0 to π, the W-image of the brane rotates by 360°. Note also that $\mathrm{e}^{-i\alpha F_V}$ rotates the fermions, and at $\alpha=\pi$ it acts as the (vectorial) fermion number operator $(-1)^{F_V}$. If we perform the GSO projection $(-1)^{F_A}=1$ (which we do in many cases), the operator $(-1)^{F_V}$ simply corresponds to the sign flip of RR-sector states. In particular, when acting on the boundary states, it reverses the sign of RR-coefficients. This corresponds to the flip of the orientation of the brane. Thus, we find

$$\mathrm{e}^{-i\pi F_V} : |\gamma\rangle \to |\overline{f_\pi\gamma}\rangle. \tag{39.392}$$

By considering various branes γ, the action of $\mathrm{e}^{-i\pi F_V}$ on the supersymmetric ground states is fully recovered by this equation. In this way, the R-charge of the RR-ground states is related to the monodromy problem $\gamma \to \overline{f_\pi\gamma}$.

One can actually express the monodromy $\gamma \to \overline{f_\pi\gamma}$ in terms of the soliton number in the massive deformation of the theory. We deform W (without changing its asymptotic behaviour) so that all the critical points, say N of them, are non-degenerate. As discussed before we can associate N A-branes to these vacua, one for each critical point. The image in the W-plane is a straight line emanating from the critical point and going to infinity along a line whose slope depends on the combination of A-type supercharges we are preserving: for Q_A^α in Eq. (39.391) they make an angle α relative to the real axis.

Let us start with $\alpha=0$ and order the N D-branes according to the lower value for $\mathrm{Im}(W)$, as depicted in Fig. 15. Let us further assume that the critical values have a convexity compatible with the ordering of $\mathrm{Im}(W)$ as shown in the figure. This can be done, by deforming the coefficients of W if necessary. As we increase α from 0 to $\pi/2$ we rotate the image of branes by 180° in the W-plane counter clockwise. As discussed above, during this process we "create" new branes as in Eq. (39.387). In particular the action of brane creation in the basis of branes emanating from the critical points γ_a is rather simple: The rotation of branes by 180° in the W-plane causes the γ_b brane to cross all the other γ_a branes with $a>b$ exactly once. Moreover

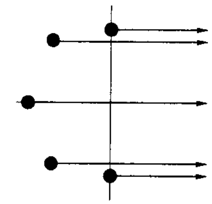


FIGURE 15. A convex arrangement of critical values in the W-plane is the most convenient one for deriving the monodromy action by 2π rotation.

during this crossover it creates $(\Delta_a\circ\Delta_b)$ new γ_a branes. This action of 180° rotation of branes is thus realized by an $N\times N$ upper triangular matrix with 1 on the diagonal and $\Delta_a\circ\Delta_b$ for each $a>b$. Thus we find

$$f_{\frac{\pi}{2}}\gamma_b^{(+)} = \gamma_b^{(-)} + \sum_{a>b}\gamma_a^{(-)}(\Delta_a\circ\Delta_b) = \sum_{a=1}^{N}\gamma_a^{(-)}I_{ab}, \tag{39.393}$$

where $\gamma_b^{(+)}$ and $\gamma_a^{(-)}$ are the basis of branes before and after the move: the former extending to real positive infinity and the latter extending to real negative infinity.

Now consider instead going from $\alpha=0$ to $\alpha=-\pi/2$. In this case for each $a>b$ we get $\Delta_a\circ\Delta_b$ brane creation of γ_b, according to Eq. (39.388). Thus the process is

$$f_{-\frac{\pi}{2}}\gamma_a^{(+)} = \gamma_a^{(-')} + \sum_{a>b}\gamma_b^{(-')}(\Delta_a\circ\Delta_b) = \sum_{b=1}^{N}\gamma_b^{(-')}I_{ab}, \tag{39.394}$$

where $\gamma_b^{(-')}$ are the basis after the move. Since both $\gamma_a^{(-)}$ and $\gamma_a^{(-')}$ emanate from the critical point p_a, they are the same setwise. However, they may differ in orientation. The relation of the orientations can be seen by tracing $f_\alpha\gamma_a$ in the neighborhood of the critical point p_a. In the W-image, the two are related by 360° rotation. However, since the superpotential is quadratic near p_a, it is 180° rotation in the pre-image. Thus, the two are related simply by the orientation flip

$$\gamma_a^{(-')} = \overline{\gamma_a^{(-)}}. \tag{39.395}$$

Thus, we find

$$\overline{f_\pi\gamma_b^{(+)}} = \overline{f_{-\frac{\pi}{2}}^{-1} f_{\frac{\pi}{2}}\gamma_b^{(+)}} = \overline{f_{-\frac{\pi}{2}}^{-1}\gamma_a^{(-)}} I_{ab} = f_{-\frac{\pi}{2}}^{-1}\gamma_a^{(-')} I_{ab} = \gamma_c^{(+)} I^{ac} I_{ab}. \tag{39.396}$$

Namely, the relevant monodromy for our problem is given by the matrix $M = {}^tI^{-1}I$. Combining what we have obtained, we see $\langle\bar{\jmath}|\,\mathrm{e}^{-i\pi F_V}|\gamma_b\rangle = \langle\bar{\jmath}|\gamma_c\rangle I^{ac}I_{ab}$. Using the definition of Π_i^a and $\widetilde{\Pi}_{\bar{\jmath}}^a$, and also the relation (39.344),

$$\langle\bar{\jmath}|i\rangle = \widetilde{\Pi}_{\bar{\jmath}}^c I^{ca}\Pi_i^a, \tag{39.397}$$

we find

$$\langle\bar{\jmath}|\,\mathrm{e}^{-i\pi F_V}|i\rangle = \widetilde{\Pi}_{\bar{\jmath}}^c I^{ac}\Pi_i^a. \tag{39.398}$$

This is the relation we wanted to show.

EXERCISE 39.6.1. *It is instructive to work out all these in the example (39.389), where $q\to 0$ is the conformal point. Start with Fig. 14(1) and show that $f_{\frac{\pi}{2}}\gamma_1 = -\widetilde{\gamma}_1 - \widetilde{\gamma}_2$, $f_{\frac{\pi}{2}}\gamma_2 = \widetilde{\gamma}_2$, as well as $f_{-\frac{\pi}{2}}\gamma_1 = \widetilde{\gamma}_1$, $f_{-\frac{\pi}{2}}\gamma_2 = -\widetilde{\gamma}_1 - \widetilde{\gamma}_2$, where $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ are as depicted in Fig. 16. In the notation of the present discussion, we have $\gamma_1^{(-)} = -\widetilde{\gamma}_1$, $\gamma_2^{(-)} = \widetilde{\gamma}_2$, and $\gamma_1^{(-')} = \widetilde{\gamma}_1$, $\gamma_2^{(-')} = -\widetilde{\gamma}_2$. Indeed the two basis sets have opposite orientations.*

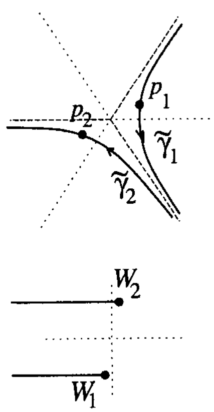


FIGURE 16. The cycles $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$

39.6.4. Holomorphic Bundles. Finally, we consider B-branes of the non-linear sigma model on a Kähler manifold X. We have seen that, for a D-brane wrapped totally on X and supporting holomorphic vector bundles, the overlaps depend on the Chern character of the bundle, Eqs. (39.366)–(39.367). This motivates us to propose that the charge lattice is the topological K-theory

$$\Lambda_{\mathrm{D}} = K(X), \tag{39.399}$$

which has a homomorphism $\mathrm{ch}: K(X)\to H^{\mathrm{even}}(X;\mathbb{Q})$. K-theory contains elements corresponding to lower-dimensional D-branes such as D-branes wrapped on complex submanifolds and supporting vector bundles on them, not just the maximum dimensional branes. Such objects can also be considered as coherent sheaves on X, or more precisely, objects in the derived category of sheaves in X. The bilinear form is given by Eq. (39.305):

$$I([E_1],[E_2]) = \int_X \mathrm{ch}([E_1^\vee])\mathrm{ch}([E_2])\mathrm{Td}(X). \tag{39.400}$$

For a Calabi–Yau n-fold, this is symmetric for even n and anti-symmetric for odd n. For a non-Calabi–Yau, it is neither symmetric nor anti-symmetric.

The overlaps with the RR ground states are parallel sections of the vacuum bundle over the moduli space $\mathcal{M}_k$ of Kähler classes and B-field classes. Classically, $\mathcal{M}_k$ is a cone in $H^2(X,\mathbb{R})$ times the torus $H^2(X,\mathbb{R})/H^2(X,\mathbb{Z})$, and therefore a non-contractible loop is associated with the integral shift of the B-field, $B\to B+\sigma$, $\sigma\in H^2(X,\mathbb{Z})$. As can be seen from the expressions Eqs. (39.366)–(39.367) for the overlaps, this induces a linear map of the charge lattice,

$$\mathcal{L}_\sigma\otimes : K(X)\to K(X), \tag{39.401}$$

which is induced from the tensor product by a line bundle with first Chern class $c_1(\mathcal{L}_\sigma) = \sigma$. This preserves the bilinear form given by Eq. (39.400). This class of monodromy is the only one that is present in the large radius limit. However, as we have learned many times in the previous chapters, the structure of the moduli space $\mathcal{M}_k$ is much more complicated than a cone times a torus. There are loci where the theory develops a singularity or has some special property. There can be other non-trivial loops, and the problem of monodromy is more non-trivial than it seems. Below, we present another class of changes of bundles that are related to the monodromy problem, in the case where X is Fano, $c_1(X) > 0$.

Mutation of Exceptional Sheaves. Let X be a Fano manifold, $c_1(X)>0$. A sheaf E on X is called *exceptional* if $\mathrm{Ext}^i(E,E)=\delta_{i,0}\mathbb{C}$. A pair of exceptional sheaves (E,F) is called an *exceptional pair* if $\mathrm{Ext}^i(E,F)=0$, $i\neq i_0$ for some i_0, and $\mathrm{Ext}^i(F,E)=0\ \forall i$. Recall that, in such a case, the space of supersymmetric ground states for the E-F open string is given by $\mathrm{Ext}^\bullet(E,F)$. For an exceptional pair, one can define new exceptional sheaves L_EF and R_FE, called the left mutation of F with respect to E and the right mutation of E with respect to F, such that (L_EF,E) and (F,R_FE) are both exceptional pairs. In the cases (i) $\mathrm{Ext}^0(E,F)\neq 0$ and $\mathrm{Ext}^0(E,F)\otimes E\to F$ is surjective, (ii) $\mathrm{Ext}^0(E,F)\neq 0$ and $\mathrm{Ext}^0(E,F)\otimes E\to F$ is injective, or (iii) $\mathrm{Ext}^1(E,F)\neq 0$, they are defined by the following exact sequences:

$$(i)\qquad \begin{aligned} &0\to L_EF\to \mathrm{Ext}^0(E,F)\otimes E\to F\to 0,\\ &0\to E\to \mathrm{Ext}^0(E,F)^*\otimes F\to R_FE\to 0,\end{aligned}$$

$$(ii)\qquad \begin{aligned} &0\to \mathrm{Ext}^0(E,F)\otimes E\to F\to L_EF\to 0,\\ &0\to R_FE\to E\to \mathrm{Ext}^0(E,F)^*\otimes F\to 0,\end{aligned}$$

$$(iii)\qquad \begin{aligned} &0\to F\to L_EF\to \mathrm{Ext}^1(E,F)\otimes E\to 0,\\ &0\to \mathrm{Ext}^1(E,F)^*\otimes F\to R_FE\to E\to 0.\end{aligned}$$

The K-theory classes are therefore related by

$$[F]=\pm[L_EF]+\chi(E,F)[E], \tag{39.402}$$

$$[E]=\pm[R_FE]+\chi(E,F)[F], \tag{39.403}$$

where the signs are $-$, $+$, $+$ for the cases (i), (ii), (iii) respectively. It is known that $R_EL_EF=\pm F$. Let us consider the example on $X=\mathbb{CP}^1$ with $E=\mathcal{O}$ and $F=\mathcal{O}(1)$. We know that $\mathcal{O}(1)$ has two holomorphic sections, denoted by X_0 and X_1, and hence $\mathrm{Ext}^0(\mathcal{O},\mathcal{O}(1))\cong\mathbb{C}^2$. The map $\mathrm{Ext}^0(\mathcal{O},\mathcal{O}(1))\otimes\mathcal{O}\to\mathcal{O}(1)$ is given by $(f,g)\mapsto fX_0+gX_1$ and is surjective. Thus, case (i) applies and we have the exact sequences

$$\begin{aligned} &0\to\mathcal{O}(-1)\to\mathrm{Ext}^0(\mathcal{O},\mathcal{O}(1))\otimes\mathcal{O}\to\mathcal{O}(1)\to 0,\\ &0\to\mathcal{O}\to\mathrm{Ext}^0(\mathcal{O},\mathcal{O}(1))^*\otimes\mathcal{O}(1)\to\mathcal{O}(2)\to 0.\end{aligned}$$

The first map of the first sequence is given by $\sigma\mapsto(X_1\sigma,-X_0\sigma)$. In the second sequence, the first map is given by $f\mapsto(fX_0,fX_1)$ and the last map is $(\lambda,\mu)\mapsto\lambda X_1-\mu X_0$. Thus, we find $L_{\mathcal{O}}\mathcal{O}(1)=\mathcal{O}(-1)$ and $R_{\mathcal{O}(1)}\mathcal{O}=\mathcal{O}(2)$.

An *exceptional collection* is an ordered set of exceptional sheaves $\{E_1,\dots,E_\ell\}$ such that (E_i,E_j) is an exceptional pair for any $i<j$. A *helix of period* N is an ordered set of infinitely many exceptional sheaves $\{E_i;i\in\mathbb{Z}\}$, such that $\{E_{i+1},\dots,E_{i+N}\}$ is an exceptional collection for any i and such that

$$E_{i+N}=R_{E_{i+N-1}}\cdots R_{E_{i+1}}(E_i),\qquad E_i=E_{i+N}\otimes K_X.$$

The collection $\{E_{i+1},\dots,E_{i+N}\}$ is called a foundation of the helix. A foundation of a helix generates the derived category of sheaves, and can be used as the basis of the charge lattice. Change of basis is done by mutations. For $X=\mathbb{CP}^1$, $E_i=\mathcal{O}(i)$ give the helix of period 2, as one can see from $R_{\mathcal{O}(1)}\mathcal{O}(0)=\mathcal{O}(2)$ and $\mathcal{O}(0)=\mathcal{O}(2)\otimes K_{\mathbb{CP}^1}$. For $X=\mathbb{CP}^{N-1}$, $\{\mathcal{O}(i);i\in\mathbb{Z}\}$ is a helix of period N.

We note here the similarity between A-branes in Landau–Ginzburg models and B-branes in Fano sigma models. Namely, the wave-front trajectories emanating from the critical points look very similar to the exceptional collection of sheaves in the following sense. Both are ordered sets of branes where the ordering determines the asymmetry of the open string Witten index. For A-branes in LG models the ordering is given by the value $\mathrm{Im}(W)$ of the W-image, while the ordering is a part of the data for an exceptional collection of sheaves. Furthermore, one can change the basis of the charge lattice. For an LG model, it is organized by Picard–Lefschetz theory. For exceptional collections, it is done by mutation of bundles. In fact, mutations of the pair $(E,F)\to(L_EF,E)$ or (F,R_FE) look very similar to the change of A-branes in the LG model $(\gamma_a,\gamma_b)\to(\gamma_b',\gamma_a')$ or (γ_b'',γ_a'') discussed above: compare Eqs. (39.387) and (39.402) as well as Eqs. (39.388) and (39.403). The change of branes in the two class of theories appears to be related in the way suggested in Fig. 17. In fact, this is not a coincidence: Mirror

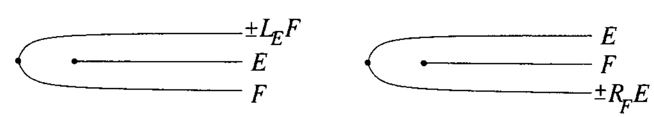


FIGURE 17. LG Interpretation of Mutations

symmetry maps the B-brane associated with the exceptional bundles of X to the A-branes wrapped on the wave-front trajectories in the mirror LG model. Then mutation of bundles indeed corresponds to change of basis of A-branes involving brane creation.

39.7. D-Branes in $\mathcal{N}=2$ Minimal Models

In this section, we study D-branes in the simplest non-trivial $\mathcal{N}=2$ superconformal field theory — $\mathcal{N}=2$ minimal models. This model is an exactly solvable model and is expected to arise as the infrared fixed point of the Landau–Ginzburg model of a single chiral superfield X with superpotential

$$W = X^{k+2}. \tag{39.404}$$

We first review the construction of D-branes following Cardy. We then find the corresponding D-branes in the Landau–Ginzburg description. In particular, we will find a beautiful geometric realization of the Verlinde ring for $SU(2)$ level k Wess–Zumino–Witten models as well as a simple understanding of the $\tau \to -\frac{1}{\tau}$ modular transformation matrix S_i^j.

39.7.1. The Model. $\mathcal{N}=2$ minimal models are unitary $(2,2)$ superconformal field theories in two dimensions with central charge $c=\frac{3k}{k+2}$, where k is a positive integer. They can be viewed as an $SU(2)/U(1)$ super-GKO construction at level k. The superconformal primary fields are labeled by three integers (j,n,s) such that

$$\begin{aligned} & j = 0, \frac{1}{2}, 1, \ldots, \frac{k}{2}, \\ & n \in \mathbb{Z}_{2(k+2)},\ (0,1,\ldots,2k+3 \bmod 2(k+2)), \\ & s \in \mathbb{Z}_4,\ (-1,0,1,2 \bmod 4), \end{aligned} \tag{39.405}$$

with the constraint $2j+n+s \equiv 0 \pmod 2$ and field identification $(j,n,s) = (\frac{k}{2}-j, n+k+2, s+2)$. We denote the set of these labels (j,n,s) by $\widehat{M}_k$. $s=0,2$ in the NS sector and $s=\pm 1$ in the Ramond sector. The two different values of s denote the parity of the fermion number in the Ramond or NS sector. The conformal weights and the $U(1)$ charges of the primary fields are

$$h_{j,n,s} = \frac{j(j+1)-n^2/4}{k+2} + \frac{s^2}{8} \bmod 1, \quad q_{j,n,s} = \frac{s}{2} - \frac{n}{k+2} \bmod 2. \tag{39.406}$$

The characters $\chi_{j,n,s}(\tau) = \mathrm{Tr}_{\mathcal{H}_{j,n,s}} q^{L_0 - \frac{c}{24}}$ obey the modular transformation relation

$$\chi_{j,n,s}(-\tfrac{1}{\tau}) = \sum_{(j',n',s')\in\widehat{M}_k} S_{j,n,s}^{j',n',s'} \chi_{j',n',s'}(\tau), \tag{39.407}$$

where

$$S_{j,n,s}^{j',n',s'} = \frac{1}{k+2}\sin\!\left(\pi\frac{(2j+1)(2j'+1)}{k+2}\right) e^{\frac{i\pi n n'}{k+2}} e^{-\frac{i\pi s s'}{2}}. \tag{39.408}$$

We consider the model with a vector-like GSO projection, with the space of states

$$\mathcal{H} = \bigoplus_{(j,n,s)\in\widehat{M}_k} \mathcal{H}_{j,n,s} \otimes \mathcal{H}_{j,n,s}. \tag{39.409}$$

There are $\mathcal{N}=2$ chiral primary states $(j,-2j,0)$ in the NS sector. The related Ramond states $(j,-(2j+1),-1)$ can be reached by spectral flow. This model can also be described by the IR fixed point of the LG model with a single chiral superfield X with superpotential (39.404). The chiral primary fields X^l correspond to the states $(\frac{l}{2},-l,0)$ and provide a representation of the chiral ring. We denote the corresponding RR-ground states in $(\frac{l}{2},-(l+1),-1)$ by $|l\rangle$:

$$|l\rangle \longleftrightarrow X^l, \qquad (l=0,1,\ldots,k). \tag{39.410}$$

Note that there are only $k+1$ chiral primary fields or $k+1$ supersymmetric ground states. However there are a total of $(k+1)(k+2)$ primary states in the Ramond sector.

39.7.2. Cardy's Construction. There are both A-branes and B-branes that preserve half of the $(2,2)$ super*conformal* symmetry. The conditions on the corresponding boundary states are the generalizations of Eqs. (39.337) and (39.338):

$$\overline{\widetilde{G}}_{-r} - i\,e^{i\alpha} G_r = \widetilde{G}_r - i\,e^{-i\alpha}\overline{G}_{-r} = J_n - \widetilde{J}_{-n} = 0 \ \text{ for A-brane}, \tag{39.411}$$

$$\overline{\widetilde{G}}_r - i\,e^{i\beta}\overline{G}_{-r} = \widetilde{G}_{-r} - i\,e^{-i\beta} G_r = J_n + \widetilde{J}_{-n} = 0 \ \text{ for B-brane}. \tag{39.412}$$

For a rational conformal field theory, the boundary states are linear combinations of "Ishibashi states" on which the left and the right generators of the superconformal algebra are linearly related. Let us define A- and B-type Ishibashi states by

$$|A_{j,n,s}\rangle\!\rangle := \sum_N |j,n,s;N\rangle \otimes \Omega_M U|j,n,s;N\rangle \text{ “}\in \mathcal{H}_{j,n,s}\otimes\mathcal{H}_{j,n,s}\text{”}, \tag{39.413}$$

$$|B_{j,n,s}\rangle\!\rangle := \sum_N |j,n,s;N\rangle \otimes U|j,n,s;N\rangle \text{ “}\in \mathcal{H}_{j,n,s}\otimes\mathcal{H}_{j,-n,-s}\text{”}. \tag{39.414}$$

Here the states $|j,n,s;N\rangle$ form an orthonormal basis of $\mathcal{H}_{j,n,s}$. U is an anti-linear operator $\mathcal{H}_{j,n,s}\to\mathcal{H}_{j,-n,-s}$ such that $U^{-1}G_rU=-i\epsilon\overline{G}_r$ on $\mathcal{H}_{j,n,s}$ with $\epsilon=1$ for $s=-1,0$ and $\epsilon=-1$ for $s=1,2$. Ω_M is the mirror automorphism of the $\mathcal{N}=2$ superconformal algebra. Note that all of $|A_{j,n,s}\rangle\!\rangle$ survive the GSO projection while $|B_{j,n,s}\rangle\!\rangle$ survives it only for the values $(j,0,0)$, $(j,0,2)$ (j integer) and also $(\frac{k}{4},\frac{k+2}{2},\pm1)$, where the latter is possible only for even k. The boundary states are particular linear combinations of Ishibashi states. For A-branes, it is simplest to follow Cardy's prescription

$$|A_{j,n,s}\rangle = \sum_{(j',n',s')\in\widehat{M}_k} \frac{S_{j,n,s}^{\;j',n',s'}}{\sqrt{S_{0,0,0}^{\;j',n',s'}}}|A_{j',n',s'}\rangle\!\rangle\,. \tag{39.415}$$

It is easy to show that this obeys the A-type condition Eq. (39.411) with $e^{i\alpha}=1$ if $s=-1,1$ and $e^{i\alpha}=1$ if $s=0,2$. All of the A-branes have non-trivial RR-components, and therefore the corresponding D-branes are charged under RR-potentials.

To obtain the B-branes, we note that the present model can be obtained from the model with the opposite GSO projection, with the Hilbert space $\oplus\mathcal{H}_{j,n,s}\otimes\mathcal{H}_{j,-n,-s}$, by orbifolding the discrete $\mathbb{Z}_{k+2}\times\mathbb{Z}_2$ symmetry. The boundary states are obtained by taking the average:

$$|B_{j,s}\rangle = \sqrt{2(k+2)} \sum_{j'\in\mathbb{Z},s'=0,2} \frac{S_{j,-2j-s,s}^{\;j',0,s'}}{\sqrt{S_{0,0,0}^{\;j',0,s'}}}|B_{j',0,s'}\rangle\!\rangle \qquad s=0,1,\; j=0,\frac{1}{2},\ldots,\frac{1}{2}\left[\frac{k}{2}\right]. \tag{39.416}$$

They obey Eq. (39.412) with $e^{i\beta}=-1$ for $s=0$ and $e^{i\beta}=1$ for $s=1$. Actually, the ones with $j=\frac{k}{4}$ (possible only if k is even) split into two

boundary states:

$$|B_s\rangle = \frac{1}{2}|B_{\frac{k}{4},s}\rangle + \frac{\sqrt{k+2}}{2}e^{-i\pi\frac{s^2}{2}}\sum_{s'=1,-1}e^{-i\pi\frac{ss'}{2}}|B_{\frac{k}{4},\frac{k+2}{2},s'}\rangle\!\rangle,\quad s=-1,0,1,2. \tag{39.417}$$

Note that the D-branes corresponding to $|B_{j,s}\rangle$ are neutral under RR-potentials, but those corresponding to $|B_s\rangle$ are charged.

The boundary states at the outgoing boundary circle are found by the prescription $\langle a| = \left[e^{i\pi J_0}|a\rangle\right]^\dagger$. Since $e^{i\pi J_0}=e^{-\pi i(\frac{n}{k+2}-\frac{s}{2})}$ on $|A_{j,n,s}\rangle\!\rangle$ and $|B_{j,n,s}\rangle\!\rangle$, we find

$$\langle A_{j,n,s}| = |A_{j,n-1,s-1}\rangle^\dagger, \tag{39.418}$$

$$\langle B_{j,s}| = |B_{j,s-1}\rangle^\dagger, \tag{39.419}$$

$$\langle B_{-1}| = |B_0\rangle^\dagger,\;\; \langle B_0| = |B_{-1}\rangle^\dagger,\;\; \langle B_1| = |B_2\rangle^\dagger,\;\; \langle B_2| = |B_1\rangle^\dagger. \tag{39.420}$$

Overlaps with supersymmetric ground states. Let us compute the overlaps of the boundary states with the RR ground states, $\langle a|\,l\,\rangle$ and $\langle\,\bar{l}\,|a\rangle$, where $|\,l\,\rangle$ is the normalized ground state corresponding to the chiral operator X^l, and $\langle\,\bar{l}\,| = |\,l\,\rangle^\dagger$. We find for A-branes

$$\langle A_{j,n,s}|\,l\,\rangle = \frac{1}{\sqrt{k+2}}\frac{\sin\left(\pi\frac{(2j+1)(l+1)}{k+2}\right)}{\sqrt{\sin\left(\pi\frac{l+1}{k+2}\right)}}e^{i\pi\frac{(n-1)(l+1)}{k+2}}e^{-i\pi\frac{s-1}{2}}, \tag{39.421}$$

$$\langle\,\bar{l}\,|A_{j,n,s}\rangle = \frac{1}{\sqrt{k+2}}\frac{\sin\left(\pi\frac{(2j+1)(l+1)}{k+2}\right)}{\sqrt{\sin\left(\pi\frac{l+1}{k+2}\right)}}e^{-i\pi\frac{n(l+1)}{k+2}}e^{i\pi\frac{s}{2}}, \tag{39.422}$$

while for the charged B-branes (for even k)

$$\langle B_s|\,l\,\rangle = \langle\,\bar{l}\,|B_s\rangle = \delta_{l,\frac{k}{2}}\frac{\sqrt{k+2}}{2},\quad s=-1,0, \tag{39.423}$$

$$\langle B_s|\,l\,\rangle = \langle\,\bar{l}\,|B_s\rangle = -\delta_{l,\frac{k}{2}}\frac{\sqrt{k+2}}{2},\quad s=1,2. \tag{39.424}$$

Witten Index. Consider an open string in the (a,b) sector. As we discussed in Sec. 39.4.3, the index $I(a,b)=\mathrm{Tr}_{a,b}(-1)^F e^{-\beta H}$ corresponds in the closed string channel to an overlap in the Ramond sector boundary states $I(a,b)=\mathrm{Tr}_{a,b}(-1)^F={}_{\mathrm{RR}}\langle a|b\rangle_{\mathrm{RR}}$. In the present case, however, since we are performing the GSO projection, we should regard the tree-channel amplitude $\langle a|q^{H_c}|b\rangle$ as the GSO projected open string partition function

$\mathrm{Tr}_{ab}(\frac{1+(-1)^F}{2}q_o^{H_o})$. Thus, the correct identification is

$$I(a,b) = \mathrm{Tr}_{ab}(-1)^F q_o^{H_o} = 2\,{}_{\mathrm{RR}}\langle a|q_c^{H_c}|b\rangle_{\mathrm{RR}} = 2\,{}_{\mathrm{RRG}}\langle a|b\rangle_{\mathrm{RRG}}, \tag{39.425}$$

where $|b\rangle_{\mathrm{RRG}}$ is the projection to the RR-ground states. We start with the A-branes $A_{j_1,n_1,s_1}, A_{j_2,n_2,s_2}$ with s_1, s_2 both odd (so that the two preserve the same $\mathcal{N}=2$ supersymmetry);

(39.426)

$$\begin{aligned}
&I(A_{j_1,n_1,s_1}, A_{j_2,n_2,s_2}) = \\
&= 2\,[|A_{j_1,n_1-1,s_1-1}\rangle_{\mathrm{RR}}]^\dagger\, q_c^{H_c} ||A_{j_2,n_2,s_2}\rangle_{\mathrm{RR}} \\
&= 2\sum_{s\,\mathrm{odd}} \frac{(S^{\,j,n,s}_{j_1,n_1-1,s_1-1})^* S^{\,j,n,s}_{j_2,n_2,s_2}}{S^{\,j,n,s}_{0,0,0}} \chi_{j,n,s}(q_c) \\
&= \frac{1}{k+2}\sum_{2j+n\,\mathrm{odd}} \frac{\sin(\frac{\pi(2j_1+1)(2j+1)}{k+2})\sin(\frac{\pi(2j_2+1)(2j+1)}{k+2})}{\sin(\frac{\pi(2j+1)}{k+2})}\, e^{\frac{i\pi n(n_2-n_1+1)}{k+2}} \\
&\qquad\qquad \times e^{-i\pi(s_2-s_1+1)}(\chi_{j,n,1}-\chi_{j,n,-1}) \\
&= \frac{2\,e^{\frac{i\pi(s_1-s_2)}{2}}}{k+2} \\
&\qquad \times \sum_j \frac{\sin(\frac{\pi(2j_1+1)(2j+1)}{k+2})\sin(\frac{\pi(2j_2+1)(2j+1)}{k+2})\sin(\frac{\pi(n_2-n_1+1)(2j+1)}{k+2})}{\sin(\frac{\pi(2j+1)}{k+2})} \\
&= (-1)^{\frac{s_1-s_2}{2}} N_{j_1,j_2}^{\frac{n_2-n_1}{2}},
\end{aligned}$$

where $N_{j_1,j_2}^{j_3}$ are the $SU(2)_k$ fusion coefficients,

$$N_{j_1,j_2}^{j_3} = \begin{cases} 1 & \text{if } \;|j_1-j_2| \le j_3 \le \min\{j_1+j_2, k-j_1-j_2\}, \\ 0 & \text{otherwise}. \end{cases} \tag{39.427}$$

If $\Delta n = n_2 - n_1$ is negative, it is understood that $N_{j,j'}^{\frac{\Delta n}{2}} = -N_{j,j'}^{\frac{|\Delta n|-2}{2}}$. The B-branes $B_{j,s}$ are not charged under RR-potentials, and so the Witten index vanishes. For the charged B-branes B_s, we find

$$I(B_1,B_1) = 2[|B_2\rangle_{\mathrm{RRG}}]^\dagger |B_1\rangle_{\mathrm{RRG}} = \frac{k+2}{2}, \tag{39.428}$$

$$I(B_{-1},B_1) = 2[|B_0\rangle_{\mathrm{RRG}}]^\dagger |B_1\rangle_{\mathrm{RRG}} = -\frac{k+2}{2}. \tag{39.429}$$

Finally, the index for the open string stretched between B_s and $A_{j,n,s}$ is given by

$$\begin{aligned}
I(B_{\pm1}, A_{j,n,s}) &= 2[|B_{2/0}\rangle_{\mathrm{RRG}}]^\dagger |A_{j,n,s}\rangle_{\mathrm{RRG}} \\
&= \mp 2\frac{\sqrt{k+2}}{2}\frac{S^{\frac{k}{4},\frac{k+2}{2},1}_{j,n,s}}{\sqrt{S^{\frac{k}{4},\frac{k+2}{2},1}_{0,0,0}}} = \mp e^{\frac{i\pi(n-s)}{2}}\sin(\pi(j+\tfrac{1}{2})).
\end{aligned} \tag{39.430}$$

It is 1 or -1 if j is an integer, but it vanishes if j is $\frac{1}{2}$ plus an integer.

39.7.3. Landau–Ginzburg Description of A-Branes.

39.7.3. Landau–Ginzburg Description of A-Branes. As mentioned, the $\mathcal{N}=2$ minimal model is realized as the IR fixed point of the LG model of a single variable X with superpotential $W = X^{k+2}$. Here we study A-branes in this LG model by considering its massive deformations. We have seen that, in a massive LG model, the gradient flow lines of $\mathrm{Re}\,(W)$ starting from a critical point sweep out a middle-dimensional Lagrangian submanifold that defines an A-brane. In the model of a single variable X, such A-branes are simply the preimages of the straight lines in the W-plane, starting from the critical values and going to $+\infty$ in the real positive direction.

The superpotential $W = X^{k+2}$ has a single critical point $X=0$ of multiplicity $(k+1)$ with critical value $w_* = 0$. If we consider deforming the superpotential by lower powers of X we will generically obtain $(k+1)$ non-degenerate critical points $\{a\}$ with distinct critical values $\{w_a\}$. We assume that $\mathrm{Im}\,(w_a)$ are separate from one another. Then we would get $(k+1)$ A-branes γ_a, one associated to each of the critical points, whose W-images are lines starting from w_a and extending straight to infinity in the positive real direction. For large values of X the deformation terms are irrelevant and the D-branes approach the preimages of the positive real axis $X^{k+2} \in \mathbb{R}_{\ge 0} \subset \mathbb{C}$, namely

$$X = r\cdot\exp\left(\frac{2\pi i\nu}{k+2}\right),\quad \nu = 0,\cdots,k+1,\quad r\in[0,\infty). \tag{39.431}$$

Thus we see that the X-plane is divided up into $k+2$ wedge shaped regions by the $k+2$ lines going from the origin to infinity making an angle of $\frac{2\pi\nu}{k+2}$ with the positive real axis; we will denote such a line by $\mathcal{L}_\nu$.

Any A-brane γ_a of the deformed theory is a curve in the X-plane that will asymptote to a pair of such lines, say $\mathcal{L}_{\nu_1}$ and $\mathcal{L}_{\nu_2}$ with $\nu_1 \ne \nu_2$. To see this, we note that the deformed superpotential W is approximately quadratic around any (non-degenerate) critical point a and the preimage of the straight line emanating from $W(a)$ in the W-plane splits into trajectories of two points (wave-fronts) starting from a. The two wave-fronts approach the lines $\mathcal{L}_{\nu_1}$ and $\mathcal{L}_{\nu_2}$ as they move away from the critical point. Since there is only one preimage of $W \in \mathbb{R}_{\geq 0}$ in the vicinity of each $\mathcal{L}_\nu$ (near infinity), the two asymptotic regions should be different for $\nu_1 \neq \nu_2$.

Thus, in the deformed theory there are $(k+1)$ distinct A-branes, each labeled by a pair of (modulo $(k+2)$) integers $\{\nu_1, \nu_2\}$ which label the asymptotes it makes. Taking into acconut the orientation, there are $2(k+1)$ of them labeled by the ordered pair (ν_i, ν_f). We will label such a D-brane by $\gamma_{\nu_i \nu_f}$.

The $(k+1)$ D-branes we obtain in this way will not intersect with each other, as their images in the W-plane do not intersect one another. Nevertheless, as discussed in Sec. 39.4.3 the index $I(\gamma_a, \gamma_b) = \mathrm{Tr}_{a,b}(-1)^F$ is not in general zero (Eq. (39.297)). This in turn is given by the intersection number of the cycles γ_a^- and γ_b^+ which are obtained from γ_a and γ_b by slightly tilting the W-images in the negative and positive imaginary directions respectively. Let us consider two allowed branes γ_{ν_i,ν_f} and $\gamma_{\nu_i',\nu_f'}$. We are interested in computing the Witten index of the theory of the oriented open string starting from γ_{ν_i,ν_f} and ending on $\gamma_{\nu_i'\nu_f'}$. If none of the ν_i, ν_f and ν_i', ν_f' are equal, the branes γ_{ν_i,ν_f}^- and $\gamma_{\nu_i',\nu_f'}^+$ do not intersect, and thus the index is zero. The more subtle case is when one of the ν_i, ν_f is equal to one of the ν_i', ν_f'. If they are both equal, then we get the Witten index to be 1 as discussed before. Without loss of generality we can order the branes so that $\nu_i < \nu_f$ and $\nu_i' < \nu_f'$ (otherwise the intersection number is multiplied by a minus sign for each switch of order). Thus there are only four more cases to discuss. Let us also assume that $\nu_i' + \nu_f' > \nu_i + \nu_f$ (by $\nu_f' = \nu_i$ we mean $\nu_f' = \nu_i + k + 2$). It is a simple exercise to see that in such cases

$$I(\gamma_{\nu_i,\nu_f}, \gamma_{\nu_i',\nu_f'}) = \begin{cases} 1 \text{ if } \nu_i = \nu_i' \text{ or } \nu_f = \nu_f' \\ 0 \text{ otherwise.} \end{cases} \tag{39.432}$$

Now we come to the D-branes at the conformal point. Since the $(k+1)$ D-branes make sense arbitrarily close to the conformal point, they survive in the limit of the conformal point as well. The resulting cycles are the broken straight lines with a sharp corner at $X = 0$:

$$\gamma_{\nu_i,\nu_f} = \mathcal{L}_{\nu_f} - \mathcal{L}_{\nu_i}. \tag{39.433}$$

We have different allowed pairs (ν_i, ν_f) in the massive theory, depending on the choice of the deformation polynomials: *all* of them survive at the conformal point. Actually, any pair (ν_i, ν_f) is realized in terms of a D-brane for *some* deformation of W. Thus, the A-brane γ_{ν_i,ν_f} exists at the conformal point for any pair (ν_i, ν_f) of mod $(k+2)$ integers. This gives us a total of $(k+2)(k+1)$ A-branes that are pairwise the same up to orientation. Here we encounter an interesting effect: *The number of D-branes jumps as we go from the conformal point to the massive theory.*

The fact that we have obtained $(k+2)(k+1)$ such branes at the conformal point is very encouraging as that is exactly the same as the predicted number of Cardy states, as already established. Moreover, if we consider the range of parameters where $0 \leq \nu_i < \nu_f \leq k+1$ we see that $|\nu_f - \nu_i| \in \{1, \ldots, k+1\}$ and $\nu_i + \nu_f \in \{0, \ldots, 2k+2\}$. The range of these parameters exactly corresponds to the quantum numbers $(2j, n)$ labelling Cardy's A-branes with $s = \pm 1$ that preserve the A-type supersymmetry with $e^{i\alpha} = 1$. This motivates us to claim that the A-branes γ_{ν_i,ν_f} are identified with Cardy's A-brane $A_{j,n,s}$, where the labels are related by

$$|\nu_f - \nu_i| = 2j + 1, \qquad \nu_f + \nu_i = n, \qquad s = \mathrm{sign}(\nu_f - \nu_i), \tag{39.434}$$

or

$$\left.\begin{aligned} \nu_f &= \tfrac{n+2j+1}{2} \\ \nu_i &= \tfrac{n-2j-1}{2} \end{aligned}\right\} s = 1, \qquad \left.\begin{aligned} \nu_f &= \tfrac{n-2j-1}{2} \\ \nu_i &= \tfrac{n+2j+1}{2} \end{aligned}\right\} s = -1. \tag{39.435}$$

Note that this is compatible with the field identification $(2j, n, s) = (k - 2j, n + k + 2, s + 2)$: If we change $\nu_i \to \nu_f$ and $\nu_f \to \nu_i + k + 2$ we get the same brane back up to a flip in the orientation (reflected in $s \to s + 2$).

Below, we provide further evidence for this identification. Along the way we find a simple geometric interpretation of Verlinde ring and modular S-matrix for $SU(2)$ level k.

Geometric Interpretation of Verlinde Algebra. We would like to compute the Witten index at the conformal point for the open string stretched between two A-branes $\gamma_{\nu_i^1,\nu_f^1}$ and $\gamma_{\nu_i^2,\nu_f^2}$ and reproduce the index formula

Eq. (39.426). One aspect of the formula is clear. The intersection number will not change if we rotate both branes by integral multiples of $\frac{2\pi}{k+2}$, which implies that the index will depend on $n_2 - n_1$ but not on the other combinations of n_1 and n_2 (as n_1 and n_2 shift by the same amount under the rotation). Moreover the appearance of $(-1)^{\frac{s_2-s_1}{2}}$ in the index is also natural as that correlates with the choice of orientation on the D-branes. So without loss of generality we set $s_1 = s_2 = 1$, i.e., as before we choose $\nu_f^1 > \nu_i^1$ and $\nu_f^2 > \nu_i^2$. Also in checking Eq. (39.426) it suffices to consider the case where $n_2 - n_1 \geq 0$, which is the same case as $\nu_f^2 + \nu_i^2 \geq \nu_f^1 + \nu_i^1$. It is now clear that the index is 1 if

$$\nu_i^1 \leq \nu_i^2 < \nu_f^1 \leq \nu_f^2 < \nu_i^1 + k + 2 \tag{39.436}$$

and it vanishes for all other cases. Note that the condition of getting non-vanishing results in the case of equality follows from Eq. (39.432).

Now we use Eq. (39.435) to rewrite Eq. (39.436) as

$$\begin{aligned} n_1 - 2j_1 - 1 &\leq n_2 - 2j_2 - 1 < n_1 + 2j_1 + 1 \\ &\leq n_2 + 2j_2 + 1 < n_1 - 2k_1 + 2k + 3. \end{aligned}$$

These four conditions can also be written as

$$|j_2 - j_1| \leq \frac{n_2 - n_1}{2} \leq \min[j_1 + j_2, k - j_1 - j_2]. \tag{39.437}$$

(We used the fact that $n_2 - n_1$ and $2j_2 + 2j_1$ are equal mod 2.) This is precisely the condition for non-vanishing of the $SU(2)$ level k fusion coefficient, and we have thus derived Eq. (39.426) from a purely LG point of view.

Overlap with RR ground states as Period Integral. We next compute the overlaps of the boundary states and the supersymmetric ground states. As shown in Eq. 39.5.3, they are given by the weighted period integrals Eqs. (39.357)–(39.358), or by (anti-)holomorphic integrals Eqs. (39.363)–(39.364) if the $(\overline{Q}_+ + Q_-)$ or $(Q_+ + \overline{Q}_-)$-cohomology classes of the ground states are expressed as in Eqs. (39.361)–(39.362). We will consider the ground state wave-functions of the form Eqs. (39.361)–(39.362) where the (anti-)holomorphic representatives are given by

$$\Omega_l = c_l X^l dX, \tag{39.438}$$

$$\overline{\Omega}_{\bar{l}} = \overline{c_l}\,\overline{X}^l d\overline{X}. \tag{39.439}$$

For a suitable choice of c_l, they define an orthonormal basis of the supersymmetric ground states. To see this, we compute the inner product using the

formula Eq. (39.365) (here we set the circumference $\beta = 2$ for simplicity):

$$\begin{aligned} g_{l\bar{l}'} &= \int e^{-i(W+\overline{W})}\Omega_l \wedge *\overline{\Omega}_{\bar{l}'} \\ &= c_l\overline{c_{l'}} \int e^{-i(X^{k+2}+\overline{X}^{k+2})} X^l \overline{X}^{l'} dX \wedge *d\overline{X} \\ &= 2c_l\overline{c_{l'}} \int e^{-2ir^{k+2}\cos((k+2)\theta)}\, e^{i(l-l')\theta} r^{l+l'+1} dr d\theta. \end{aligned}$$

At this stage, we see that they are orthogonal, $g_{l\bar{l}'} = 0$ if $l \neq l'$ (we assume $0 \leq l, l' \leq k$). Expanding the exponential and performing the θ integral, we find

$$\begin{aligned} g_{l\bar{l}} &= 4\pi|c_l|^2 \int_0^\infty \sum_{m=0}^\infty \frac{(-1)^m}{(m!)^2}(r^{k+2})^{2m} r^{2l+1} dr \\ &= 4\pi|c_l|^2 \int_0^\infty J_0(2r^{k+2}) r^{2l+1} dr \\ &= |c_l|^2 \frac{2}{k+2}\left[\Gamma\left(\frac{l+1}{k+2}\right)\right]^2 \sin\left(\frac{\pi(l+1)}{k+2}\right), \end{aligned}$$

where $J_0(x) = \sum_{m=0}^\infty (-1)^m (x/2)^{2m}/(m!)^2$ is the Bessel function, and we have used the integral formula

$$\int_0^\infty x^{\mu-1} J_0(ax)dx = 2^{\mu-1}[\Gamma(\mu/2)]^2 \sin(\pi\mu/2)/(\pi a^\mu).$$

Thus, they form an orthonormal basis if c_l are chosen as

$$c_l = e^{i\gamma_l} \frac{\sqrt{k+2}}{\Gamma(\frac{l+1}{k+2})\sqrt{2\sin(\frac{\pi(l+1)}{k+2})}}, \tag{39.440}$$

where $e^{i\gamma_l}$ is some phase.

Now, let us perform the overlap integrals. We consider the brane $\gamma_{\nu_i\nu_f} = \gamma_{jns}$ corresponding to A_{jns}, where ν_i, ν_f are related to jns via Eq. (39.435). We consider in particular the one with $s = 1$. So γ_{jns} is a broken line coming from infinity in the direction $z_i = e^{\frac{\pi i(n-2j-1)}{k+2}}$, cornering at $X = 0$, and then going out to infinity in the direction $z_f = e^{\frac{\pi i(n+2j+1)}{k+2}}$.

The overlap is

$$\begin{aligned}\Pi_l^{\gamma_{jn1}} &= c_l \int_{\gamma_{jn1}^-} e^{-iX^{k+2}} X^l dX \\ &= c_l(z_f^{l+1}-z_i^{l+1})\int_0^{+\infty-i\cdot 0} e^{-it^{k+2}} t^l dt \\ &= c_l \frac{1}{k+2} e^{\frac{\pi i(n-1/2)(l+1)}{k+2}} 2i\sin\left[\frac{\pi(2j+1)(l+1)}{k+2}\right]\Gamma\left(\frac{l+1}{k+2}\right)\end{aligned}$$

and

$$\begin{aligned}\widetilde{\Pi}_l^{\gamma_{jn1}} &= \overline{c_l}\int_{\gamma_{jn1}^+} e^{-i\overline{X}^{k+2}}\overline{X}^l * d\overline{X} \\ &= i\overline{c_l}\overline{\int_{\gamma_{jn1}^+} e^{iX^{k+2}} X^l dX} \\ &= i\overline{c_l}\, e^{\frac{-\pi i(n+1/2)(l+1)}{k+2}}(-2i)\sin\left[\frac{\pi(2j+1)(l+1)}{k+2}\right]\Gamma\left(\frac{l+1}{k+2}\right).\end{aligned}$$

Using the normalization formula Eq. (39.440) with the phase choice $e^{i\gamma_l} = -i\, e^{\frac{-\pi i(l+1)}{2(k+2)}}$, we find that the above reproduce the overlaps obtained earlier (Eqs. (39.421)–(39.422)):

(39.441) $$\Pi_l^{\gamma_{jn1}} = \sqrt{2}\,\langle A_{jn1}|\,l\,\rangle,\quad \widetilde{\Pi}_{\bar{l}}^{\gamma_{jn1}} = \sqrt{2}\,\langle\,\bar{l}\,|A_{jn1}\rangle.$$

Note that $\sqrt{2}$ is the standard factor associated with the GSO projection: compare Eqs. (39.343) and (39.425).

Among the $(k+2)(k+1)$ A-branes, let us choose $(k+1)$ of them, $\gamma_{a,a+1}$ with $a=0,1,\ldots,k$, which generate the charge lattice $H_1(\mathbb{C},B)$. It is straightforward to see that the index $I_{ab}=I(\gamma_{a,a+1},\gamma_{b,b+1})$ and its inverse I^{ab} are given by

(39.442) $$I_{ab}=\delta_{a,b}-\delta_{a,b+1},\qquad I^{ab}=\begin{cases}1 & \text{if } a\geq b,\\ 0 & \text{if } a<b.\end{cases}$$

Using this, and using the above expressions for Π_l^a and $\widetilde{\Pi}_{\bar{l}'}^a$ for $\gamma_{a,a+1}=\gamma_{0,2a+1,1}$, it is straightforward to compute

(39.443) $$\sum_{a,b=0}^{k}\widetilde{\Pi}_{\bar{l}'}^a I^{ab}\Pi_l^b = \delta_{l,\bar{l}'},$$

(39.444) $$\sum_{a,b=0}^{k}\widetilde{\Pi}_{\bar{l}'}^a I^{ba}\Pi_l^b = -\,e^{-2\pi i\frac{l+1}{k+2}}\delta_{l,\bar{l}'}.$$

This confirms the relations (39.397) and (39.398). Note that the ground state $|\,l\,\rangle$ is characterized by its vector R-charge $F_V = 2\frac{l+1}{k+2}-1$, or $e^{-\pi i F_V} = -\,e^{\frac{-2\pi i(l+1)}{k+2}}$. The latter Eq. (39.444) can also be regarded as confirming that the ground state wave-function with holomorphic representative $\Omega_l = c_l X^l dX$ indeed corresponds to the state $|\,l\,\rangle$. Finally, the same equation can also be regarded as the diagonalization of the fusion coefficients using the modular S-matrix: I_{ab} is essentially the fusion coefficients $N_{jj'}^{\ j''}$ while Π_l^a is the modular S-matrix $S_j^{\ j'}$.

39.7.4. Landau–Ginzburg Description of B-Branes. We now study B-branes of the LG model $W=X^{k+2}$ by considering its massive deformations.

We have seen that a B-brane in a massive LG model is wrapped on a complex submanifold on which the superpotential W is constant. Furthermore, in order for the supersymmetry to be unbroken, it has to pass through some of the critical points of W. In the LG model with a single variable, this means that only D0-branes localized at the critical points are possible. For a generic deformation of $W=X^{k+2}$, there are $(k+1)$ critical points. Thus, we find $(k+1)$ D0-branes at those points. As we have seen, the Witten index of any pair of such branes is zero. Also, since the dimension of the cycle, 0, is not one-half of the number of LG fields (which is $1/2$), they have vanishing overlaps with RR-ground states (see Eq. (39.375)). This means that these B-branes are not charged under RR-potentials.

Actually, one can consider another LG realization of the $\mathcal{N}=2$ minimal model. It is to add one variable, Y, with the quadratic superpotential Y^2. Namely, we consider the model of two variables X,Y with the superpotential

(39.445) $$W_0 = X^{k+2}-Y^2.$$

In the IR limit, the Y sector simply goes away and we again obtain the $\mathcal{N}=2$ minimal model. We will discuss its massive deformations, obtained by adding lower-order terms in X to the superpotential W_0. An advantage in this realization is that one can now have middle-dimensional complex submanifolds, that is, submanifolds of complex dimension 1. We are interested in submanifolds passing through some critical points of W. Let us focus on a neighborhood of a critical point $(X,Y)=(X_*,0)$, where W behaves as $W=W(X_*)+c(X-X_*)^2-Y^2+\cdots$. Then the submanifold defined by $W=W(X_*)$ has a double point singularity — it is a union of two curves $Y = \pm\sqrt{c}(X - X_*) + \cdots$ intersecting at the critical point $(X,Y) = (X_*,0)$. Our previous discussion is not sufficient to study such a D-brane. The only way to avoid this difficulty is to select either one of the signs $Y = \pm\sqrt{c}(X-X_*)+\cdots$. This is consistent only if the superpotential globally factorizes as $W = (Y+\sqrt{c}(X-X_*)+\ldots)(-Y+\sqrt{c}(X-X_*)+\ldots)$. To avoid linear terms in Y the dot terms $+\ldots$ must be the same, and therefore W should be of the form

$$W = (X^{\frac{k+2}{2}}+\cdots)^2 - Y^2 = \Big((X-a_1)\cdots(X-a_{\frac{k+2}{2}})\Big)^2 - Y^2, \tag{39.446}$$

up to addition of a constant. This is possible only if k is even. There are $\frac{k+2}{2}$ critical points at $X=a_i, Y=0$ plus $\frac{k}{2}$ more at other places. The subsets

$$Z_\pm = \{\, Y = \pm(X-a_1)\cdots(X-a_{\frac{k+2}{2}})\,\} \tag{39.447}$$

are smooth, middle-dimensional submanifolds obeying $W=$ constant. Thus, they define B-branes possibly with non-vanishing overlaps with RR-ground states. They both pass through exactly $\frac{k+2}{2}$ critical points at $X=a_i, Y=0$.

One can apply the method developed in Sections 39.4.5 and 39.5.4 to study some of the properties of these branes. First, the Witten index of an open string ending on the same brane is the number of critical points included in the brane. Thus,

$$I(Z_\pm, Z_\pm) = \frac{k+2}{2}. \tag{39.448}$$

We next compute the overlap of the corresponding boundary states and RR-ground states. Since we consider B-brane boundary states and the ground states from B-twist, the overlaps are given by topological disc amplitudes. To find the normalization of the ground states, we first compute the topological metric (S^2 amplitudes with two operator insertions). We fix the holomorphic two -form as $\Omega = dYdX$. An elementary way to compute $\langle O_1O_2\rangle$ is to first compute the Hessian at each of the critical points $\det\partial_i\partial_jW$ and then sum $O_1O_2/\det\partial_i\partial_jW$ over the critical points. The contribution of the Y-sector to the Hessian is simply -2. The X-sector contribution would be complicated to obtain, but we will use the following technique: For holomorphic functions $f(X)$ and $g(X)$ we have

$$\sum_{g=0}\frac{f(X)}{g'(X)} = \frac{1}{2\pi i}\oint_{\text{Large circle}}\frac{f(X)dX}{g(X)}, \tag{39.449}$$

where the contour is a large circle encircling all zeroes of g. Applying this to $g=\partial_XW = (k+2)X^{k+1}+\cdots$ and $f = X^{l_1}X^{l_2}$, we find

$$\begin{aligned}\langle X^{l_1}X^{l_2}\rangle &= \sum_{\partial_XW=0}\frac{X^{l_1}X^{l_2}}{(-2)\partial_X^2W}\\ &= \frac{1}{(-2)(k+2)}\delta_{l_1+l_2,k}+\cdots \xrightarrow{W\to W_0} \frac{1}{(-2)(k+2)}\delta_{l_1+l_2,k},\end{aligned} \tag{39.450}$$

where (-2) is from the Y-sector, and $+\cdots$ are the terms that vanish in the conformal limit $W\to W_0 = X^{k+2}-Y^2$. Thus, the properly normalized operators in the conformal limit are

$$O_l = \sqrt{2(k+2)}X^l. \tag{39.451}$$

To compute the disc amplitudes, we must compute the "Pfaffian," Eq. (39.376), at each critical point included in $Z_\pm$. Let us first consider Z_+ and use the normal coordinate $z = (X-a_1)\ldots(X-a_{\frac{k+2}{2}})-Y$. Then we find $\Omega = dXdz$ and $\partial_X\partial_zW = 2\partial_X((X-a_1)\ldots(X-a_{\frac{k+2}{2}}))$. Thus, the disc amplitude is

$$\begin{aligned}\Pi_l^{Z_+} &= \sum_{i=1}^{\frac{k+2}{2}}\left.\frac{\sqrt{2(k+2)}X^l}{2\partial_X((X-a_1)\ldots(X-a_{\frac{k+2}{2}}))}\right|_{X=a_i}\\ &= \frac{1}{2\pi i}\oint_{\text{Large circle}}\frac{\sqrt{2(k+2)}X^l}{2(X-a_1)\ldots(X-a_{\frac{k+2}{2}})} \xrightarrow{W\to W_0} \frac{\sqrt{2(k+2)}}{2}\delta_{l,\frac{k}{2}}.\end{aligned} \tag{39.452}$$

where we have used Eq. (39.449). Together with the topological metric $\eta_{l,l'} = -\delta_{l+l',k}$ and the other overlap $\widetilde{\Pi}_l^{Z_+} = -\Pi_l^{Z_+}$, we reproduce the index Eq. (39.448) via $\Pi_l^{Z_+}\eta^{ll'}\widetilde{\Pi}_{l'}^{Z_+} = \frac{k+2}{2}$. We also find $\Pi_l^{Z_-} = -\widetilde{\Pi}_l^{Z_-} = -\delta_{l,\frac{k}{2}}\sqrt{2(k+2)}/2$, and

$$I(Z_+,Z_-) = I(Z_-,Z_+) = -\frac{k+2}{2}. \tag{39.453}$$

As far as the ground state sector is concerned, we see that Z_+ and Z_- differ only in their orientations. Moreover, in the conformal limit, Z_+ and Z_- are geometrically the same in the vicinity of $X=Y=0$. This suggests that they indeed differ only in orientation in this limit.

Let us now compare with the result obtained by Cardy's method. There we have found k (resp. $(k+1)$) uncharged B-branes if k is even (resp. k is odd), half of them ($s=0$) preserving the opposite B-type supersymmetry to that of the other half ($s=1$). Some of them may correspond to the $(k+1)$ D0-brane discovered in the LG description. However, it is hard to make the relation more precise, since there is no non-trivial quantity that is computable in the LG description. If k is even we have also found four charged B-branes B_s ($s=-1,0,1,2$) where B_{s+2} and B_s differ just by orientation, and B_{even} preserves the opposite supersymmetry to that of B_{odd}. For a suitable boundary condition for the fermions, we see that Z_+ or Z_- corresponds to B_{-1} or B_1. Indeed, the Witten indices agree, and we also see

$$\Pi_l^{Z_\pm}=\sqrt{2}\langle B_{\mp 1}|\,l\,\rangle,\ \ \widetilde{\Pi}_l^{Z_\pm}=-\sqrt{2}\langle\,\bar{l}\,|B_{\mp 1}\rangle. \tag{39.454}$$

The factor $\sqrt{2}$ is the one associated with the GSO projection. The minus sign of the second equation can be traced back to the relation of the topological states $\langle\,l\,|$ and the conjugate states $\langle\,\bar{l}\,|=-\langle(k-l)|$ (which follows from the comparison of the metrics).

39.8. Mirror Symmetry

In this section, we describe how D-branes are related by mirror symmetry. Recall that mirror symmetry exchanges the supercharges Q_- and $\overline{Q}_-$. Since A-branes preserve the combination of the supercharges $\overline{Q}_++Q_-$ while B-branes preserve $\overline{Q}_++\overline{Q}_-$, mirror symmetry exchanges A-branes and B-branes. We will study how Lagrangian A-branes are related to holomorphic B-branes. We will consider the correspondence between holomorphic branes in sigma models and Lagrangian branes in Landau–Ginzburg mirrors, as well as Lagrangian branes in sigma models and holomorphic branes in LG.

39.8.1. Holomorphic Bundles on X and Lagrangian Branes in LG. In Sec. 39.6, where we have studied monodromy properties of D-branes, we have seen a similarity between B-branes in Fano sigma models and A-branes in Landau–Ginzburg models. Here we show that this is not a coincidence — these branes are mapped to each other under mirror symmetry.

Let us take the example $X=\mathbb{CP}^{N-1}$. We have seen that this is mirror to the Landau–Ginzburg model of $N-1$ periodic variables $Y_1,\dots,Y_{N-1}$ of periodicity $2\pi i$ with superpotential

$$W=\mathrm{e}^{-Y_1}+\cdots+\mathrm{e}^{-Y_{N-1}}+\mathrm{e}^{-t+Y_1+\cdots+Y_{N-1}}. \tag{39.455}$$

Here $t=r-i\theta$ corresponds to the complexified Kähler class parameter of $\mathbb{CP}^{N-1}$, where θ determines the B-field, Eq. (15.71). The superpotential has N critical points p_ℓ ($\ell\equiv\ell+N$), where p_ℓ is at $\mathrm{e}^{-Y_1}=\cdots=\mathrm{e}^{-Y_{N-1}}=\mathrm{e}^{-t/N}\,\mathrm{e}^{2\pi i\ell/N}$.

Let us first find, in the theory with $\theta=0$ ($B=0$), the mirror of the D-brane wrapped totally on $\mathbb{CP}^{N-1}$ and supporting the trivial $U(1)$ gauge field. We recall that the mirror symmetry was derived using the linear sigma model — $U(1)$ gauge theory with charge 1 chiral superfields $\Phi_1,\dots,\Phi_N$. The variables Y_i ($i=1,\dots,N$) were obtained by T-dualizing along the phase of Φ_i. The D-brane wrapped on $\mathbb{CP}^{N-1}$ with trivial gauge field is lifted in the linear sigma model to a pure Neumann boundary condition. What is relevent here is the fact (see Sec. 39.1.3) that T-duality maps the Neumann boundary condition on S^1 to the Dirichlet boundary condition on the dual circle $\widetilde{S}^1$, where the gauge field on S^1 corresponds to the position of the dual brane on $\widetilde{S}^1$. When this rule is applied to the linear sigma model fields, we find that the pure Neumann boundary condition on Φ_i is mapped to the D-brane boundary condition $\operatorname{Im}Y_i=0$ on Y_i. Thus, we find that the D-brane wrapped on $\mathbb{CP}^{N-1}$ with trivial gauge field on it is mirror to the D-brane wrapped on the middle-dimensional submanifold $\operatorname{Im}Y_i=0$. The latter is indeed a Lagrangian submanifold and its W-image is the straight line emanating from $W(p_0)=N$ and extending in the positive real direction.

Let us now apply to this mirror pair an axial R-rotation. The axial R-rotation shifts the theta angle of the linear sigma model as $\theta\to\theta-2N\alpha$. This reduces in the non-linear sigma model to the shift of the B-field. Thus the trivial B-brane on $\mathbb{CP}^{N-1}$ is now under the influence of a non-trivial B-field. On the other hand, on the LG side this induces a rotation of the critical points and their critical values: $W(p_\ell)=N\,\mathrm{e}^{2\pi i\ell/N}\to N\,\mathrm{e}^{-2i\alpha+2\pi i\ell/N}$. In fact this axial R-rotation shifts the fields Y_i as $Y_i\to Y_i+2i\alpha$. Thus, the A-brane is deformed to $\operatorname{Im}Y_i=2\alpha$. The image in the W-plane is a straight line emanating from the new critical value $N\,\mathrm{e}^{-2i\alpha}$ and making an angle -2α with the real axis. Now, consider the special value $\alpha=\pi/N$. At this value, the set of critical points is the same as the starting configuration: p_ℓ has moved to the position where $p_{\ell-1}$ was. Now we recall the fact that a D-brane in the theory with a quantized value B_0 of the B-field is identical to a D-brane in a theory with trivial B-field supporting a non-trivial $U(1)$ gauge bundle whose first Chern class c_1 is related to B_0. In the present case,

$\theta = -2\pi$ corresponds to a gauge bundle with $\int_{\mathbb{CP}^1} c_1 = 1$. The holomorphic line bundle with this topology is simply $\mathcal{O}(1)$. Thus, we have seen that $\mathcal{O}(1)$ is mirror to the A-brane through p_{-1} whose W-image is the straight line emanating from $W(p_{-1}) = N\,\mathrm{e}^{-2\pi i/N}$ and extending straight to infinity with angle $-2\pi/N$ with the real axis. Repeating this procedure, we find that $\mathcal{O}(-\ell)$ is mirror to the A-brane through p_ℓ whose W-image is the straight line emanating from $W(p_\ell) = N\,\mathrm{e}^{2\pi i\ell/N}$ and extending straight to infinitly with angle $2\pi\ell/N$ with the real axis.

Since the axial rotation rotates the supercharges, in general, D-branes obtained in this way preserve different combinations of supersymmetry compared to the original one which is $\overline{Q}_+ + Q_-$. In the LG side, a D-brane preserving $\overline{Q}_+ + Q_-$ is obtained by deforming the cycle so that the W-image is rotated around the critical value and becomes parallel to the real axis. This is depicted in Fig. 18 in the example of $\mathcal{O}(-1)$ and $\mathcal{O}(1)$ on $\mathbb{CP}^5$. For the bundles $\mathcal{O}(2)$, $\mathcal{O}(3), \ldots$ or $\mathcal{O}(-2)$, $\mathcal{O}(-3), \ldots$, it is impossible to rotate the cycle so that the images in the W-plane are parallel to the real axis without passing through other critical values. One can avoid this cross-over by bending the branes, although it results in the breaking of the supersymmetry. Bending in the clockwise direction, we obtain the collection of bundles $\{\mathcal{O}(0), \ldots, \mathcal{O}(n+1)\}$, which are exceptional collections as we have seen above. By partially changing the direction of bending, we can obtain other exceptional collections as shown in Fig. 19 for the case of $\mathbb{P}^5$. If we order the lines in terms of decreasing asymptotic imaginary part, then the exceptional collection in Fig. 19(b) is $\{\mathcal{O}(-3), \mathcal{O}(-2), \mathcal{O}(-1), \mathcal{O}(0), \mathcal{O}(1), \mathcal{O}(2)\}$ and the exceptional collection in Fig. 19(c) is $\{\mathcal{O}(-2), \mathcal{O}(-1), \mathcal{O}(0), \mathcal{O}(1), \mathcal{O}(2), \mathcal{O}(3)\}$. These exceptional collections are related to one another by mutations. The similarity of the monodromy property observed in Sec. 39.6 can thus be understood as a

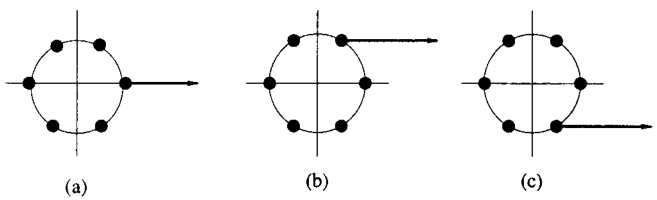


FIGURE 18. a) $\mathcal{O}(0)$, b) $\mathcal{O}(-1)$, c) $\mathcal{O}(1)$

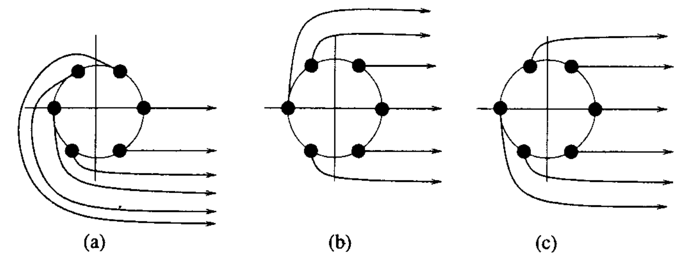


FIGURE 19. Three exceptional collections.

consequence of mirror symmetry. For bundles E_i corresponding to supersymmetric (straight) A-branes γ_i, the spaces of open string supersymmetric ground states match with each other

$$\mathrm{Ext}^\bullet(E_i, E_j) \cong HF^\bullet_W(\gamma_i, \gamma_j). \tag{39.456}$$

39.8.2. Lagrangians in X and Holomorphic Branes in LG. We next consider A-branes in toric sigma models and B-branes in LG models. In both cases, the combination of supercharges that defines the topological theory is preserved by the boundary condition. In Sec. 39.4, we have seen that the relation $Q^2 = 0$ can be modified: it can be broken by an anomaly in non-linear sigma models while it can be modified at the classical level in LG models. In the case where the sigma model and the LG model are mirror to each other, the two modifications of $Q^2 = 0$ should match. In other words, one can use mirror symmetry to predict when we expect the quantum anomaly of $Q^2 = 0$ in the sigma model.

Let us consider the general toric manifold X, defined by the charges Q_i^a ($i = 1, \ldots, N$, $a = 1, \ldots, k$). One can construct a boundary interaction in the linear sigma model that includes a boundary analogue of the F-term that depends holomorphically on parameters $s_1, \ldots, s_N$ obeying $\sum_{i=1}^N Q_i^a s_i = t^a$. We denote $s_j = c_j - ia_j$. In the sigma model limit $e_a \to \infty$, the boundary interaction reduces to the one for the D-brane wrapped on the torus T_c^{N-k} defined by $|\phi_i|^2 = c_i$ with the Wilson line

$$A_a = \sum_{i=1}^N \left[a_i d\varphi_i - \theta^a M_{ab} Q_i^b c_i d\varphi_i \right], \tag{39.457}$$

where $\varphi_i = \arg\phi_i$ and M_{ab} is the inverse matrix of $\sum_{i=1}^N Q_i^a Q_i^b c_i$. Eq. (39.457) is well defined on X if and only if the constraints $\sum_{i=1}^N Q_i^a s_i = t^a$ are satisfied. Dualizing the boundary theory, we find that the mirror of (T_c, A_a) is a D0-brane in the LG model located at $\mathrm{e}^{-Y_i} = \mathrm{e}^{-s_i}$.

Now, we come to an interesting point. Let us consider such A-branes, $(T_{c^{(1)}}, A_{a^{(1)}})$ and $(T_{c^{(2)}}, A_{a^{(2)}})$. They are mirror to a pair of D0-branes at $\mathrm{e}^{-Y} = \mathrm{e}^{-s^{(1)}}$ and $\mathrm{e}^{-s^{(2)}}$. We have seen that, for such a pair of B-branes in the LG model, the open string supercharge squares to

$$Q^2 \propto W(\mathrm{e}^{-s^{(2)}}) - W(\mathrm{e}^{-s^{(1)}}). \tag{39.458}$$

The quantum anomaly for the A-branes in X vanishes if and only if the right-hand side vanishes. This also applies to the most general case where the mirror B-branes are not necessarily D0-branes. Thus, we find the condition: *for a pair of Lagrangians with flat connections, the quantum anomaly of $Q^2 = 0$ vanishes if and only if the W-values of the mirror B-branes match with each other.*

Supersymmetry is spontaneously broken unless the point $\mathrm{e}^{-Y_i} = \mathrm{e}^{-s_i}$ is a critical point of the superpotential W. This is a strong constraint on T_c and A_a. For example, consider $X = \mathbb{CP}^1$ where the mirror is the $\mathcal{N}=2$ sine-Gordon model $W = \mathrm{e}^{-Y} + \mathrm{e}^{-t+Y}$. There are two critical points $\mathrm{e}^{-Y} = \pm\,\mathrm{e}^{-t/2}$. The corresponding branes are wrapped on the equator $T_{r/2}$ and have Wilson line 0 or π.

One can use this mirror symmetry to solve the open topological sigma model for (T_c, A_a) in terms of the open topological LG model for the corresponding D0-brane. Let $(\theta)^n := \overline{\theta}^1 \cdots \overline{\theta}^n$ be the chiral ring element of the boundary LG model. This corresponds to the volume element of T_c. The cylinder amplitude with $(\theta)^n$ insertion at both boundaries is given by $\det \partial_i \partial_j W$ at the D0-brane location p. The disk amplitude with boundary $(\theta)^n$ insertion and a bulk chiral ring element O inserted in the interior is the value of O at p. It is a simple exercise to show that, for $X = \mathbb{CP}^1$, this gives the correct number of holomorphic disks with the boundary lying in the equator $T_{r/2}$.

CHAPTER 40

References

Part I

Chapter 1

The material in this section is standard. There are many books on differential geometry. The author accepts any blame, however, for the extremely loose "derivation" of covariant differentiation in Sec. 1.4.2, which is similar to that of [**117**].

Chapter 2

This chapter has borrowed *heavily* from [**121**]. A more detailed explanation of toric varieties is given in the chapter titled "Toric Varieties for String Theory," and references are given therein.

Chapter 3

This chapter is a compilation of material that is described eloquently and much more thoroughly in other references.

The material on cohomology and Poincaré duality is described in [**30**]. [**193**] is an excellent treatment of Morse theory. The point of view of Chern classes described here can be found in [**121**]. The axiomatic definition of Chern classes is taken from [**194**], to which the reader is referred for a much more comprehensive treatment. The differential geometric view is discussed in many references; some with a physical viewpoint are [**79**] and Ref. [**117**], which also includes the monopole example. The Grothendieck–Riemann–Roch formula and Serre duality are staples of algebraic geometry, and are proved for example in [**128**].

Chapter 4

This chapter is basically a summary of the paper by Atiyah and Bott, [**10**]. Kontsevich was the first to compute Gromov–Witten invariants by localization techniques in [**163**].

Chapter 5

The author learned about almost complex structures and the integrability condition from R. Bott. The conventions of the exercise on the Nijenhuis tensor were chosen to agree with Eq. (15.2.4) of [**120**]). A good discussion of the Kähler condition is given in [**121**]. A readable discussion of the Calabi–Yau condition (and its importance for physics) is given in chapter 15 of [**120**].

Chapter 6

Newlander and Nirenberg showed that complex coordinates could be found if the Nijenhuis tensor is zero in [**202**]. The theory of deformations of complex structures was developed by Kodaira and Spencer in Ref. [**161**]. Tian (Ref. [**235**]) and Todorov (Ref. [**241**]) proved the unobstructedness ("local Torelli theorem") of Calabi-Yau moduli space, while the coordinatization by the z^a periods was shown by Bryant and Griffiths in Ref. [**38**]. Griffiths developed the theory of variation of Hodge structure. Many people worked on results involving Calabi–Yau moduli space. Our treatment in section three follows the work of Candelas and de la Ossa in Ref. [**43**]. After some early hints at such a phenomenonon, mirror symmetry was proved (from a physical point of view) for the quintic in Ref. [**122**]. The first solution of the Picard–Fuchs equations and interpretation in terms of rational curves was performed by Candelas, de la Ossa, Green and Parkes in Ref. [**45**]. The geometry of the conifold singularity was discussed in Ref. [**44**].

Chapter 7

This chapter is an expanded version of lectures by Sheldon Katz at the CMI Mirror Symmetry Summer School.

Pedagogical treatments of toric varieties can be found in [**153, 203, 99**]. Another treatment of toric varieties with a view towards application to mirror symmetry can be found in Ch. 3 of [**63**].

Details of the construction of toric varieties from fans using homogeneous coordinates are given in [**62**]. See any of Sec. 2.1 of Ref. [**99**], Theorem 1.10 of Ref. [**203**], or Theorem 4 of Ref. [**153**] for the omitted details in the proof of the smoothness criterion.

See [**265**] for details regarding the gauged linear sigma model. See Sec. 3.3, 3.4 of Ref. [**63**] and the references therein for the relationship between the constructions given here and the symplectic quotient construction of toric varieties. See Sec. 3.4 of Ref. [**63**] for a description of the GKZ decomposition in the language of toric geometry.

See Sec. 5.1–5.2 of [**99**] for a description of the Chow group in terms of toric data. Details about the connection between the charges and the intersection numbers are in Sec. 3.3.2 of Ref. [**63**]. Theorem 7.4.4 is proven in Prop. 1.6 of Ref. [**219**].

The statement used in the proof of the orbifold criterion can be in Corollary 2 in [**32**].

Example 7.7.4 is from [**46**].

Part II

Chapter 8

This chapter gives a basic overview of QFT. For a more advanced/detailed version, see [**209**], [**68**], and [**251**].

Chapter 9

No specific references.

Chapter 10

For quantum mechanics in general, [**168**] is an excellent textbook. For supersymmetric quantum mechanics, there are three basic papers of Witten, [**260**], [**258**] and [**259**]. The localization formula for the Atiyah–Singer index theorem is discussed in [**94**] and [**3**]. An excellent account on instantons is given by Coleman in [**60**]. The discussion of degenerate Morse theory is based on [**13**]. The proof in (10.5.7) that moment maps for torus actions are perfect Morse functions is due to a private communication from K. Fukaya and Y.-G. Oh. Other proofs can be found in F. Kirwan's book [**156**] and also in H. Nakajima's notes [**200**, Ch. 5.1].

Chapter 11

T-duality was found by Kikkawa and Yamasaki, [**154**] and by Sakai and Senda [**223**]. Path-integral derivation of T-duality is due to Buscher, Ref. [**39**]. For further study of σ-models on tori and the Narain lattice, see Ref. [**201**]. For a review and more references see Ref. [**111**].

Boson-fermion equivalence in $1+1$ dimensions was originally discovered in an interacting system — sine-Gordon model and massive Thirring model, see [56] and [185]. The particular version discussed in this Chapter is in [82].

Chapter 12

For general definitions and expository treatment of superspace, see [92]. More extensive references on superspace, supersymmetry, and supergravity are [252], [103], and [89]. The twisted chiral multiplet of (2, 2) supersymmetry in 1 + 1 dimensions was found in [104]. For the statement of mirror symmetry, see the original paper [173] of Lerche–Vafa–Warner.

Chapter 13

The relation of $\mathcal{N} = 2$ supersymmetry and Kähler geometry is essentially due to Zumino [272]. The Landau–Ginzburg model is the dimensional reduction of the very first example of supersymmetric field theory found by Wess and Zumino in [253] and is sometimes called the Wess–Zumino model. The chiral anomaly in the path-integral formalism is due to Fujikawa [95]. T-duality in the (2, 2) superspace is due to Roček and Verlinde [220]

Chapter 14

For general ideas of the renormalization group, see [254] and [255]. See also accounts by Coleman [58].

For renormalization of non-linear sigma models, see [93]. For renormalization of supersymmetric non-linear σ-models, we refer the reader to [5]. That the Kähler class is renormalized only at one-loop was originally found in [4], but our argument greatly simplifies the proof.

The non-renormalization theorem of the superpotential was originally found in perturbation theory by Grisaru, Roček and Siegel [125]. A non-perturbative proof using holomorphy, symmetry and the asymptotic condition is due to [225].

The basic paper on conformal field theory in two dimensions is by Belavin, Polyakov and Zamolodchikov [23]. For $\mathcal{N} = (2, 2)$ superconformal field theories, see [173]. For IR fixed points of $\mathcal{N} = 2$ Landau–Ginzburg models, see [189] and [247].

Chapter 15

For basics on gauge symmetry, spontaneous symmetry breaking and Higgs mechanism, see the textbooks on quantum field theory. An account by Coleman [59] is also recommended. Absence of spontaneous breaking

of continuous symmetry in 1 + 1 dimensions is due to Coleman [55] and to Mermin and Wagner [192].

The discussion of the linear sigma model is based totally on the paper by Witten [265]. For early papers on supersymmetric linear sigma models, see [257] and [65]. See also the more recent paper [196] with a lot of examples. For linear sigma models with non-abelian gauge groups, see for example [266].

For dynamics of gauge theory in 1+1 dimensions, see Coleman [57] and [61]. The $\Sigma \log \Sigma$-type effective action was first obtained by D'Adda et al [66].

For the connections between superconformal theories and the Calabi–Yau sigma model, we refer the reader to Gepner's construction, [107]. For the Landau–Ginzburg model in this context see [124] and [189]. For the sigma model in this context see [256].

Chapter 16

For construction of chiral rings, see the original paper [173].

For background on topological field theories, see Witten's paper on topological sigma models [261].

For topological Landau–Ginzburg models, see the paper of Vafa of the same title [244]. For A-twist of the linear sigma model, see [196]. For more information about the topological B-model, see the paper of Vafa [245] and the paper of Witten [264].

Chapter 17

See the original paper by Cecotti and Vafa [50].

Chapter 18

For the structure of BPS kinks see [51]. For more detail on the CFIV index see the original paper [49].

Chapter 19

D-branes were discovered by Dai–Leigh–Polchinski [67] and by Hořava [131]. D-branes in superstring theory were recognized as the source of Ramond–Ramond gauge potential in the seminal paper [210]. For review articles on D-branes and superstrings, see [211] and [214].

Part III

Chapter 20

The original source for the physics proof of mirror symmetry is Hori-Vafa [**135**].

For the construction of mirror pairs as LG orbifolds, see [**123**]. For enumerative predictions using the Greene–Plesser models, see [**45**]. Mirror symmetry for the $\mathbb{C}P^n$ σ-model is discussed in [**88**]. Also related is the work of Eguchi, Hori and Xiong [**80**].

Part IV

Chapters 21–34 contain the content of Rahul Pandharipande's lectures at the Clay Institute Mirror Symmetry Summer School, as well as detailed background material from [**102**], [**118**], [**207**], [**208**].

The Mirror symmetry predictions of the genus 0 invariants of Calabi–Yau hypersurfaces have been proven from a different perspective by Lian, Liu, and Yau in [**180**]–[**183**]. An expository mathematical development of Gromov–Witten theory and Mirror symmetry can be found in [**63**]. We refer the reader to [**63**] for a history of the subject and a literature survey.

The perspective of these chapters in Part IV is algebro-geometric. A foundational treatment in symplectic geometry can be found in [**222**].

Chapter 21

Sec. 21.1.2. *Deligne–Mumford stacks.* The theory of Deligne–Mumford stacks was introduced in [**69**] to study the moduli space of curves. The forthcoming book [**21**] should provide an accessible introduction.

Chapter 22

Sec. 22.3. *Differentials on nodal curves.* An accessible sketch of why this is a reasonable generalization of differential is given in [**14**, p. 675].

Chapter 23

Sec. 23.3.1. *Degenerations of non-singular curves.* The important fact that a family of stable curves over the punctured complex disc can be completed (after base change) is the *Stable Reduction Theorem* (see [**127**] Section 3E or [**14**] Section 1 for more information).

Chapter 24

Moduli spaces $\overline{\mathcal{M}}_{g,n}(X,\beta)$ of stable maps. A more complete (and mathematically rigorous) introduction is given in [**102**]. Much of the discussion here closely follows [**102**].

Sec. 24.1. *Example: The Grassmannian.* See [**101**] for a discussion of the Fulton–MacPherson configuration space.

Chapter 25

Sec. 25.1.1. *Recursions for rational plane curves.* The presentation is taken from [**102**]. For a check that the right-hand side of Eq. (25.1) (in *Exercise 25.1.1*) counts maps correctly, see [**102**] Lemma 14.

Sec. 25.2. *The tautological line bundles $\mathbb{L}_i$ and the classes ψ_i.* The boundary lemma is well known; see [**63**] Lemma 9.2.2 for a proof. *Exercise 25.2.8*: See [**152**] for a discussion of the intersection theory of $\overline{\mathcal{M}}_{0,n}$, which is essentially combinatorial in nature.

Sec. 25.2.1. *Aside: Witten's conjecture (Kontsevich's theorem).* For more on Witten's conjecture, see the original sources [**263**], [**162**]. For more on the Virasoro conjecture, see [**63**] Section 10.1.4.

Chapter 26

Sec. 26. *The virtual fundamental class.* See [**178**] and [**20**], [**22**] for discussion and proofs. See [**100**] Section 6.3 for a proof of Theorem 26.1.2, the Excess intersection formula.

Sec. 26.2. *Gromov–Witten invariants and descendant invariants.* The presentation here is taken directly from [**102**] and [**207**].

Sec. 26.4. *Descendant invariants from Gromov–Witten invariants in genus 0.* This section is taken directly from [**207**]. Proposition 26.4.1 (Genus 0 descendant reconstruction) was originally proved in [**77**].

Secs. 26.5–26.5.2 are taken directly from [**102**]. See also [**165**], [**222**].

Chapter 27

An introduction to equivariant cohomology and localization. See [**10**] Sections 2 and 3 for a more detailed (and beautifully explained) introduction to equivariant cohomology and localization. The algebraic version of the theory is developed in [**78**].

Sec. 27.1. *The equivariant cohomology of projective space.* See [**30**] equation (20.6) or [**100**] Ex. 8.3.4 for discussion of the splitting formula.

Sec. 27.3. *Localization on the moduli space of maps: Determination of fixed loci.* Kontsevich introduced the localization method to stable maps in [**163**]. The presentation in this section is taken directly from [**118**].

Sec. 27.5. *The Aspinwall–Morrison formula.* Manin's proof of the Aspinwall–Morrison formula appears in [**186**]. A proof in symplectic geometry is given in [**249**].

Sec. 27.6. *Virtual localization.* The virtual localization formula is proven in [**118**]. The presentation of this section is taken directly from that source. For Faber's algorithm, see [**83**]. In [**84**], a natural sequence of differential operators are found annihilating generating series of Hodge and ψ-classes. Theoretical techniques for manipulating Hodge integrals have been developed in [**84**], [**85**], [**86**].

Sec. 27.7. *The full multiple cover formula for* $\mathbb{P}^1$. See [**84**] for a full proof of Theorem 27.7.1. The contributions $C(1,d)$ were first computed in [**118**].

Chapter 28

The presentation here is taken directly from [**207**].

Chapter 29

The Mirror conjecture for hypersurfaces I — the Fano case. While the most general context for such relationships has not yet been understood, tremendous progress has recently been made, see [**180**]–[**183**], [**155**], [**207**], [**26**], [**27**], [**170**], [**105**], [**108**]–[**110**]. The presentation in this chapter is taken directly from [**207**].

Sec. 29.1 *Overview of the conjecture.* The prediction of Candelas, de la Ossa, Green, and Parkes appeared in [**45**].

Sec. 29.1.2 *The Clemens Conjecture.* See [**248**] for the reason why there exist nodal degree 5 rational curves on a generic quintic. For the quintic, the numbers n_d are enumerative at least for $d \leq 9$ ([**145**], [**157**]; for a discussion of higher degrees, see [**158**]). See [**63**] Section 7.4 for more discussion of the Clemens conjecture.

Chapter 30

The Mirror conjecture for hypersurfaces II — the CY case. This proof is taken from [**207**] Section 4. The method follows Givental's approach [**108**].

Part V

Chapter 31

Topological strings were first introduced by Witten in [**262**] .

For topological twisting for CFTs, and topological strings examined in the context of CFT, see Eguchi and Yang [**81**]. For topological strings and the holomorphic anomaly, see the BCOV papers [**24**] and [**25**].

For the Dijkgraaf–Verlinde–Verlinde formula on associativity of the chiral ring, see their paper [**70**] on topological strings In $D < 1$.

Chapter 32

The references for the material in this chapter are the two papers on M-theory and topological strings [**114**] as well as [**147**].

Chapter 33

A related but slightly different approach to a mathematical of the GV invariants has been proposed in [**138**]. Our approach follows [**147**] with some extensions.

The K3 situation was discussed in [**271**]. Related mathematical work in the K3 case appears in [**18, 87**].

Our deformation analysis of smooth curves is from Ch. 8 of [**47**]. There is a more general analysis for Calabi–Yau threefolds with ordinary double points in [**54**].

See [**164**] for an introduction to homological mirror symmetry.

Example 33.0.6 is from [**35**], where different computational techniques were used.

The classification of singularity types of a general surface section of the contraction of a curve is from [**150**].

Chapter 34

Ch. 34 *Multiple covers, integrality, and Gopakumar–Vafa invariants.* A sketch of a natural approach to Conjecture 34.0.1 in the symplectic category can be found in [**63**]. In the quintic case, a number theoretic approach is also possible via hypergeometric series (see [**179**]). The Gopakumar–Vafa invariants were first defined in [**112**], [**113**]. There should be a relationship between Eq. (34.3) and multiple cover contributions, see [**208**], [**37**]. For compelling evidence of Gopakumar and Vafa's mathematical construction of n^g_β using a moduli space of sheaves on X, see [**147**], [**137**]. A refined version of Conjecture 34.0.2 for local Calabi–Yau geometries is proposed in [**37**].

Sec. 34.1 *The Gromov–Witten theory of threefolds.* The presentation follows [**208**]. The Hodge integral $\int_{\overline{\mathcal{M}}_g} \lambda_{g-1}^g$ was computed in [**84**]. For more motivation for Pandharipande's conjecture 34.2.1, see [**208**]. In the case $\int_\beta c_1(X) = 0$, the integrality constraint is believed to be "equivalent" to a suitably defined cover formula; see [**208**], [**37**] for a discussion.

Sec. 34.3 *Consequences for algebraic surfaces.* See [**199**] for the explanation of why λ_2 satisfies a boundary relation in $\mathcal{M}_2$.

Sec. 34.4 *Elliptic rational surfaces.* See [**34**] for further details, including a proof of Conjecture 34.3.2 for the classes L_k. The proof that $F(q)$ wis a modular form is due to Bryan and Leung [**36**].

Chapter 35

Most of the material presented in this section goes back to the papers [**24**] and [**25**]. The derivation of the partition function of string theory on the torus may require some standard string theory background, see e.g., [**212**]. The duality group for string theory on tori is reviewed in [**111**]. The evaluation of F_1 on the torus has some phenomenological interest as it captures threshold corrections to the gauge couplings below the string scale in orbifold compactifications, see e.g., [**72**], where useful explicite calculations can be found.

For the mathematical discussion of the Ray–Singer torsion and the Quillen anomaly see [**218**] and [**216**]. Reference [**28**] deals specifically with the differential equations for the determinants needed to calculate F_1 on Calabi–Yau spaces.

The open string action is presented e.g., in [**268**] and [**171**]. The calculation of the determinants for the open string can be done following [**6**].

The necessary material concerning the mirror map and the genus 0 amplitudes on Calabi–Yau spaces was reviewed in Sec. 6, but see also [**45**], [**136**]. The same formalism for the local case was specifically considered in [**146**],[**53**] and [**160**]. The explicit higher-genus ($g > 2$) calculations were done in [**160**] and [**147**]. Some direct calculations of the n_d^g for the quintic and the $\mathcal{O}(-3) \to \mathbb{P}^2$ case can be found in [**147**].

Chapter 36

For further detail on geometric engineering of QFTs see [**146**] and [**149**] and the review articles of Klemm [**159**], Lerche [**172**] and Mayr [**190**].

For the large N limit of Chern–Simons theory and the relation to topological strings see the original paper of Gopakumar and Vafa [**115**] .

For the connection to knot invariants, see [**205**], [**166**], and [**167**].

Chapter 37

For an introduction to effective actions from string theory, read [**120**], Chs. 13, 14, and [**212**], Chs. 12, 17, 18. Non-perturbative mirror symmetry first appeared in [**90**]. The condition of conformality on bosonic D-branes was derived in [**67**]. Polchinski showed that D-branes carry non-perturbative Ramond–Ramond charge in [**210**]. BPS states in general supersymmetry algebras can be found in [**252**]. Boundary states and the open/closed string channel are discussed in D-brane cycle conditions were derived from the low energy point of energy point of view in [**19**], and from the conformal view in Ref. [**204**] (no gauge field). D-brane conditions including B-field and gauge field on the brane are derived in Ref. [**188**]. The special Lagrangian fibration conjecture appeared in Ref. [**234**]. The moduli space of special Lagrangian submanifolds was first discussed in Ref. [**191**]. The complexified moduli space of special Lagrangian cycles was discussed in Ref. [**234**] and Ref. [**130**]. Mirror Calabi–Yau manifolds are shown to have dual torus fibrations by M. Gross in Ref. [**126**]; also by W.-D. Ruan in Ref. [**221**]. Mirror symmetry with bundles is due to Vafa (Ref. [**246**]). Witten derived the result that Chern–Simons theory (resp. holomorphic C–S theory) is the string field theory of the open string A-model (resp. B-model) in Ref. [**267**]. A gerbes to mirror symmetry can be found in [**130**], and is due to Hitchin and Chatterjee. Also in that paper is a discussion of D-brane (A-cycle) moduli space. The "definitive" description of B-cycles is still lacking. The role of the tachyon in D-brane annihilation was conjectured by Sen in [**230, 231**], and emphasized by Witten in Ref. [**269**]. Superconnections are due to Quillen (see, e.g., Ref. [**217**]. Douglas and collaborators have studied the physics of the derived category; a review can be found in Ref. [**75**]. Kontsevich's conjecture can be found in Ref. [**164**]. Fukaya developed the A_∞ category of Lagrangians in [**96, 98**]. The example of the elliptic curve is worked out in detail in [**215**]. The The real Fourier transform functor appears from the topological point of view (for tori) in Ref. [**7**] and from the geometric point of view developed in this chapter in [**176**]. Kontsevich discusses the relation between his conjecture and the usual mirror symmetry in his original paper, [**164**].

Chapter 38

The splitting of boundary conditions on the worldsheet into purely complex and symplectic sides is due to Witten [**267**], and interpreted in the context of mirror symmetry exchanging the two is in [**246**]. The reformulation of mirror symmetry as the statement of equivalence of two associated categories of D-branes (long before D-branes were invented) is in [**164**]. That considering the complex and symplectic structures independently reflects the physics is referred to as the "decoupling conjecture" in Ref. [**33**], and checked there in a number of cases. For standard theory of hermitian metrics and connections on holomorphic vector bundles see Sec. 5.2.1 and, for example, Section 0.5 of Ref. [**121**]. A proof of the of a $\bar{\partial}$-operator with $\bar{\partial}^2 = 0$, giving a bundle a holomorphic structure, is in Chapter 2 of [**73**].

Symplectic structures and reduction on the infinite dimensional space of connections on a bundle were introduced in [**9**]. Using the gradient flow of minus the norm square of the moment map was exploited mainly by Donaldson in infinite dimensions, for instance to prove (38.2.2) on algebraic surfaces in [**73**] (where there are more references to original sources). Uhlenbeck and Yau prove Theorem 38.2.2 in Ref. [**243**]. For deformed instanton equations such as those of Ref. [**188**] little analytical work has yet been done, but the corresponding results for a related equation (associated with Gieseker stability) are in Ref. [**174**], with similar symplectic structures to those discussed here in Ref. [**175**]. The stability condition we describe should presumably be related to the the 'Π-stability' of Ref. [**76**]. The algebro-geometric construction of moduli of stable sheaves (as opposed to the more standard Gieseker stable sheaves) was accomplished for a complex surface in Ref. [**177**].

That D-branes should take their charges in K-theory was proposed in [**195**] and [**269**]; the role of tachyons between branes was discovered in [**231**]. For an introduction to derived categories, their motivation and geometric meaning see [**238**]; for proper rigorous accounts try [**106**], [**144**], [**250**]. Reconstruction of a variety (with ample (anti-)canonical bundle) from its derived category is proved in Ref. [**29**], and the fact that all equivalences are Fourier–Mukai transforms ([**197**]) is proved in [**206**], based on work in Ref. [**198**]. The main result of Ref. [**12**], expressing moduli of string theories on $K3$ surfaces in terms of just the Mukai lattice, can be reinterpreted as saying that the derived category determines the string theory, by a result of Ref. [**206**]. That a flop induces an equivalence of derived categories is proved in many cases in Ref. [**29**] and in general in Ref. [**31**]. The Fourier–Mukai transforms by mirror symmetry and the monodromy transformations around rational double points of the mirror (symplectic) manifold are studied in Ref. [**229**]. Fourier–Mukai transforms and derived featured in the physics literature in, for instance, [**134**], [**11**], [**15**]. Since the summer school took place the role of the derived category in string theory has been dramatically clarified in [**75**] as the category of topological boundary conditions of the conformal field theory.

The second half of the chapter, on Lagrangians, is based on [**239**] and [**240**] (in which there are more details), in turn inspired by [**164**] and [**98**]. Graded Lagrangians are due to Kontsevich [**164**] and studied by Seidel in [**227**] (where it is shown that Lagrangian connect sums are gradeable if and only if the Floer index of their intersection is 1). Their definition is the one we gave in Sec. 37.7.1; the one given in Sec. 38.4 is slightly different in using the complex structure, and using orientations: we pass to the $\mathbb{Z}_2$-orientation cover of the Lagrangian Grassmannian, then to its universal $\mathbb{Z}$-cover; this is equivalent to the standard definition. The Ricci-flat metric on a Calabi–Yau manifold is provided by the fundamental result of Yau [**270**]. Special Lagrangians were introduced in [**129**], and derived from supersymmetry in [**19**]. The deformation theory of sLags is due to McLean [**191**].

Floer homology for Lagrangians originated in [**91**]. The obstruction theory for (special) Lagrangians and their Floer theory has now been worked out in [**97**]. Explicit mirror calculations of obstructions appear in [**97**] and [**142**]. That stable holomorphic bundles (six-branes) can be obstructed is shown in [**237**].

The holomorphic Chern–Simons functional was introduced in [**267**], and its geometry studied in [**236**], where the holomorphic Casson invariant is also defined, as suggested in [**74**], and calculated in examples such as the one mentioned in the footnote. The mirror relation between the real and holomorphic Chern–Simons functionals is suggested in [**246**], which has been further developed after completion of the school in [**1, 2**]. This mirror relationship has also been proved in a special case in [**176**]. The complexified functional that allows the extension to non-Lagrangian submanifolds is due independently to Chen and Tian [**52**] and to [**239**]. Joyce's study of counting sLag homology three-spheres is begun in [**140**]. Perhaps the first person to suggest the mirror equality of real and holomorphic Casson invariants was Tyurin [242].

Mean curvature flow for Lagrangians is studied in [233]; see also [224], [240] for more details about minimising volume of Lagrangians. Lagrangian connect sum and the symplectic monodromy results quoted are in Ref. [226]. Joyce's examples are in Ref. [140], based on work of Lawlor [169]. That slope should be mirror to slope is a slogan of Ref. [215]; see also Ref. [76] for the more complicated situation. The mirror monodromy calculations in the derived category (Eqs. (38.29) and (38.30)) are from Ref. [229]. Interpreting Joyce's obstruction as a marginal for sLags is due to Ref. [76], and an earlier example of this phenomenon is in Ref. [232]. Stability for Lagrangians, moment maps, and the conjecture, are from Ref. [239]. Mean curvature flow, the limiting Jordan–Hölder sLag decomposition and other details are from Ref. [240]. There examples are worked out proving the conjecture for mean curvature flow to the sLags of Ref. [232], and uniqueness of sLags in hamiltonian deformation classes is proved under some Floer homology conditions. Atiyah's classification of sheaves on T^2 is from Ref. [8], and convergence of mean curvature flow is due to Grayson in [119].

Chapter 39

Materials recorded in this Chapter are partly classic, partly new but known, and partly original. In paricular, the discussion on charged B-branes in Landau–Ginzburg models has not appeared in the literature.

First of all, we again cite the fundamental papers on D-branes [67] and [131]. Boundary states are introduced in [40, 213].

See Ooguri–Oz–Yin [204] for a discussion of the supersymmetric condition for D-branes in Calabi–Yau spaces. The condition of supersymmetry for D-branes in Landau–Ginzburg models is found in [116] and [133]. A-branes that are not Lagrangian were extensively studied in [143].

The paper [148] was useful in preparing the discussion on R-anomaly.

Supersymmtric ground states for open strings stretched between Lagrangian submanifolds is extensively studied in [97]. Open string ground states for a pair of A-branes in Landau–Ginzburg models and for a pair of B-branes in non-linear sigma models are studied in [133]. A part of the discussion on the open string ground states for a pair of B-branes in LG models is in [132], but the rest is original.

Properties of overlaps of boundary states and RR ground states were originally discussed in [204] and [133].

For background references for D-branes in rational conformal field theory, see [139] and [48]. The basic reference for D-branes in $\mathcal{N}=2$ minimal model is [184]. For minimal model D-branes in Landau–Ginzburg description, we again refer to [133] and [132], but the discussion on charged B-branes is original.

Mirror symmetry between B-branes in toric manifold and A-branes in LG models is studied in [133] and [228]. Mirror symmetry between A-branes in toric manifolds and B-branes in LG models is studied in [132]. For mirror symmetry between special Lagrangians in toric Calabi–Yau and holomorphic branes in the mirror CY, with computation of space-time superpotential, see [1] and [2].

Bibliography

[1] M. Aganagic and C. Vafa, *Mirror symmetry, D-branes and counting holomorphic discs*, preprint, 2000 (hep-th/0012041).

[2] M. Aganagic, A. Klemm and C. Vafa, *Disk instantons, mirror symmetry and the duality web*, Z. Naturforsch. A **57** (2002), 1–28.

[3] L. Alvarez-Gaumé, *Supersymmetry and the Atiyah–Singer index theorem*, Comm. Math. Phys. **90** (1983), 161–173.

[4] L. Alvarez-Gaumé, S. R. Coleman and P. Ginsparg, *Finiteness of Ricci flat $N = 2$ supersymmetric σ-models*, Comm. Math. Phys. **103** (1986), 423–430.

[5] L. Alvarez-Gaumé and D. Z. Freedman, *Geometrical structure and ultraviolet finiteness in the supersymmetric σ-model*, Comm. Math. Phys. **80** (1981), 443–451; *The background field method and the ultraviolet structure of the supersymmetric nonlinear σ-model*, Ann. Phys. **134** (1981), 85–109.

[6] L. Alvarez-Gaumè, G. Moore and C. Vafa, *Theta functions, modular invariance and strings*, Comm. Math. Phys. **106** (1986), 1–40.

[7] D. Arinkin and A. Polishchuk, *Fukaya category and Fourier transform*, in *Winter school on mirror symmetry, vector bundles and Lagrangian submanifolds* (Cambridge, MA, 1999) (C. Vafa and S.-T. Yau, eds.), AMS/IP Stud. Adv. Math. vol. 23, Amer. Math. Soc., Providence, RI, 2001, pp. 261–274.

[8] M. F. Atiyah, *Vector bundles over an elliptic curve*, Proc. London Math. Soc. **7** (1957), 414–452.

[9] M. F. Atiyah and R. Bott, *The Yang-Mills equations over Riemann surfaces*, Phil. Trans. Roy. Soc. London Ser. A **308** (1983), 523–615.

[10] M. F. Atiyah and R. Bott, *The moment map and equivariant cohomology*, Topology **23** (1984), 1–28.

[11] P. S. Aspinwall and R. Y. Donagi, *The heterotic string, the tangent bundle, and derived categories*, Adv. Theor. Math. Phys. **2** (1998), 1041–1074.

[12] P. S. Aspinwall and D. R. Morrison, *String theory on $K3$ surfaces*, in *Mirror Symmetry*, II (B. Greene and S.-T. Yau, eds.), AMS/IP Stud. Adv. Math. vol. 1, Amer. Math. Soc., Providence, RI, 1997, pp. 703–716.

[13] D. M. Austin and P. J. Braam, *Morse-Bott theory and equivariant cohomology*, in *The Floer Memorial Volume* (H. Hofer et al, eds.), Progr. Math. vol. 133, Birkhäuser, Boston, MA, 1995, pp. 123–183.

[14] F. Bardelli, *Lectures on stable curves*, in Lectures on Riemann Surfaces (Trieste, 1987) (M. Cornalba, X. Gómez-Mont and A. Verjovsky, eds.), World Sci., Teaneck, NJ, 1989, pp. 648–704.

[15] C. Bartocci, U. Bruzzo, D. Hernandez Ruiperez and J. Munoz Porras, *Mirror symmetry on $K3$ surfaces via Fourier-Mukai transform*, Comm. Math. Phys. **195** (1998), 79–93.

[16] V. Batyrev, *Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties*, J. Algebraic Geom. **3** (1994), 493–535.

[17] V. V. Batyrev and L. A. Borisov, *On Calabi-Yau complete intersections in toric varieties*, in *Higher-dimensional complex varieties* (Trento, 1994) (M. Andreatta and T. Peternell, eds.), Walter de Gruyter, Berlin, 1996, pp. 39–65.

[18] A. Beauville, *Counting rational curves on K3 surfaces*, Duke Math. J. **97** (1999), 99–108.

[19] K. Becker, M. Becker and A. Strominger, *Fivebranes, membranes and non-perturbative string theory*, Nuclear Phys. B **456** (1995), 130–152.

[20] K. Behrend, *Gromov-Witten invariants in algebraic geometry*, Invent. Math. **127** (1997), 601–617.

[21] K. Behrend, D. Edidin, B. Fantechi, W. Fulton, L. Göttsche and A. Kresch, *Introduction to stacks*, in preparation.

[22] K. Behrend and B. Fantechi, *The intrinsic normal cone*, Invent. Math. **128** (1997), 45–88.

[23] A. A. Belavin, A. M. Polyakov and A. B. Zamolodchikov, *Infinite conformal symmetry in two-dimensional quantum field theory*, Nuclear Phys. B **241** (1984), 333–380.

[24] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, *Holomorphic anomalies in topological field theories*, Nuclear Phys. B **405** (1993), 279–304.

[25] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, *Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes*, Comm. Math. Phys. **165** (1994), 311–427.

[26] A. Bertram, *Another way to enumerate rational curves with torus actions*, Invent. Math. **142** (2000), 487–512.

[27] G. Bini, C. de Concini, M. Polito and C. Procesi, *On the work of Givental relative to mirror symmetry*, in Appunti dei Corsi Tenuti da Docenti della Scuola, Scuola Normale Superiore, Pisa, 1998.

[28] J.-M. Bismut, H. Gillet and C. Soulé, *Analytic torsion and holomorphic determinant Bundles I. Bott-Chern Forms and Analytic Torsion*, Comm. Math. Phys. **115** (1988), 49–78.

[29] A. I. Bondal and D. O. Orlov, *Semiorthogonal decomposition for algebraic varieties*, preprint, 1995 (alg-geom/9506012).

[30] R. Bott and L. W. Tu, *Differential forms in algebraic topology*, Springer–Verlag, New York, 1986.

[31] T. Bridgeland, *Flops and derived categories*, Invent. Math. **147** (2002), 613–632.

[32] M. Brion, *Rational smoothness and fixed points of torus actions*, Transf. Groups **4** (1999), 127–156.

[33] I. Brunner, M. R. Douglas, A. Lawrence and C. Römelsberger, *D-branes on the quintic*, J. High Energy Phys. (2000), no. 8, paper 15, 72 pp.

[34] J. Bryan, *Evidence for a conjecture of Pandharipande*, Turkish J. Math. **26** (2002), 69–73.

[35] J. Bryan, S. Katz and N. C. Leung, *Multiple covers and the integrality conjecture for rational curves in Calabi–Yau threefolds*, J. Algebraic Geom. **10** (2001), 549–568.

[36] J. Bryan and N. C. Leung, *The enumerative geometry of K3 surfaces and modular forms*, J. Amer. Math. Soc. **13** (2000), 371–410.

[37] J. Bryan and R. Pandharipande, *BPS states of curves in Calabi–Yau 3-folds*, Geom. Topol. **5** (2001), 287–318.

[38] R. Bryant and P. Griffiths, *Some observations on the infinitesimal period relations for regular threefolds with trivial canonical bundle*, in *Arithmetic and geometry, vol. II* (M. Artin and J. Tate, eds.), Progr. Math. vol. 36, Birkhäuser, Boston, 1983, pp. 77–102.

[39] T. H. Buscher, *A symmetry Of the string background field equations*, Phys. Lett. B **194** (1987), 59–62; *Path integral derivation of quantum duality in nonlinear σ-models*, Phys. Lett. B **201** (1988), 466–472.

[40] C. G. Callan, C. Lovelace, C. R. Nappi and S. A. Yost, *Adding holes and cross-caps to the superstring*, Nuclear Phys. B **293** (1987), 83–113.

[41] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, *Vacuum configurations for superstrings*, Nuclear Phys. B **258** (1985), 46–74.

[42] P. Candelas, M. Lynker and R. Schimmrigk, *Calabi–Yau manifolds in weighted* $\mathbf{P}_4$, Nuclear Phys. B **341** (1990), 383–402.

[43] P. Candelas and X. C. de la Ossa, *Moduli space of Calabi–Yau manifolds*, Nuclear Phys. B **255** (1991), 455–481.

[44] P. Candelas and X. C. de la Ossa, *Comments on conifolds*, Nuclear Phys. B **342** (1990), 246–258.

[45] P. Candelas, X. C. de la Ossa, P. Green and L. Parkes, *A pair of Calabi–Yau manifolds as an exactly soluble superconformal field theory*, Nuclear Phys. B **359** (1991), 21–74.

[46] P. Candelas and A. Font, *Duality between the webs of heterotic and type II vacua*, Nuclear Phys. B **511** (1998), 295-325.

[47] P. Candelas, A. Font, S. Katz and D. Morrison, *Mirror symmetry for two parameter models* II, Nuclear Phys. B **429** (1994), 626–674.

[48] J. L. Cardy, *Boundary conditions, fusion rules and Verlinde formula*, Nuclear Phys. B **324** (1989), 581–596.

[49] S. Cecotti, P. Fendley, K. A. Intriligator and C. Vafa, *A New supersymmetric index*, Nuclear Phys. B **386** (1992), 405–452.

[50] S. Cecotti and C. Vafa, *Topological antitopological fusion*, Nuclear Phys. B **367** (1991), 359–461.

[51] S. Cecotti and C. Vafa, *On classification of $N = 2$ supersymmetric theories*, Comm. Math. Phys. **158** (1993), 569–644.

[52] J. Chen, *Lagrangian sections and holomorphic $U(1)$-connections*, Pacific J. Math. **203** (2002), 139–160.

[53] T. M. Chiang, A. Klemm, S.-T. Yau and E. Zaslow, *Local mirror symmetry: Calculations and interpretations*, Adv. Theor. Math. Phys. **3** (2000), 495-565.

[54] H. Clemens and H. Kley, *Counting curves which move with threefolds*, J. Algebraic Geom. **9** (2000), 175–200.

[55] S. R. Coleman, *There are no Goldstone bosons in two-dimensions*, Comm. Math. Phys. **31** (1973), 259–264.

[56] S. R. Coleman, *Quantum sine-Gordon equation as the massive Thirring model*, Phys. Rev. D **11** (1975), 2088–2097.

[57] S. R. Coleman, *More about the massive Schwinger model*, Ann. Phys. **101** (1976), 239–267.

[58] S. Coleman, *Dilatation*, in Aspects of symmetry, Cambridge Univ. Press, Cambridge, 1985.

[59] S. Coleman, *Secret symmetry: An introduction to spontaneous symmetry breakdown and gauge fields*, in *Aspects of symmetry*, Cambridge Univ. Press, Cambridge, 1985.

[60] S. Coleman, *The uses of instantons*, in *Aspects of Symmetry*, Cambridge Univ. Press, Cambridge, 1985.

[61] S. R. Coleman, R. Jackiw and L. Susskind, *Charge shielding and quark confinement in the massive Schwinger model*, Ann. Phys. **93** (1975), 267–275.

[62] D. Cox, *The homogeneous coordinate ring of a toric variety*, J. Algebraic Geom. **4** (1995), 17–50.

[63] D. Cox and S. Katz, *Mirror symmetry and algebraic geometry*, Mathematical surveys and Monographs vol. 68, Amer. Math. Soc., Providence, RI, 1999.

[64] D. A. Cox, S. Katz, and Y.-P. Lee, *Virtual fundamental classes of zero loci*, in *Advances in algebraic geometry motivated by physics* (Lowell, MA, 2000) (E. Previato, ed.), Contemp. Math. vol. 276, Amer. Math. Soc., Providence, RI, 2001, pp. 157–166.

[65] A. D'Adda, P. Di Vecchia and M. Luscher, *Confinement and chiral symmetry breaking in $\mathbb{CP}^{N-1}$ models with quarks*, Nuclear Phys. B **152** (1979), 125–144.

[66] A. D'Adda, A. C. Davis, P. Di Vecchia and P. Salomonson, *An effective action for the supersymmetric $\mathbb{CP}^{N-1}$ model*, Nuclear Phys. B **222** (1983), 45–70.

[67] J. Dai, R. Leigh, and J. Polchinski, *New connections between string theories*, Mod. Phys. Lett. A **4** (1989), 2073–2083.

[68] P. Deligne, P. Etingof, D. S. Freed, L. C. Jeffrey, D. Kazhdan, J. W. Morgan, eds., D. R. Morrison, E. Witten, *Quantum fields and strings: A course for mathematicians*, Amer. Math. Soc., Providence, RI, Inst. Adv. St., Princeton, NJ, 1999.

[69] P. Deligne and D. Mumford, *The irreducibility of the space of curves of given genus*, Inst. Hautes Études Sci. Publ. Math. **36** (1969), 75–109.

[70] R. Dijkgraaf, H. Verlinde and E. Verlinde, *Topological strings in $d < 1$*, Nuclear Phys. B **352** (1991), 59–86.

[71] L. Dixon, *Some world-sheet properties of superstring compactifications, on orbifolds and otherwise*, in *Superstrings, unified theories and cosmology 1987* (Trieste, 1987), ICTP Ser. Theoret. Phys. vol. 4, World Sci., Teaneck, NJ, 1988, pp. 67–126.

[72] L. J. Dixon, V. S. Kaplunovsky and J. Louis, *Moduli-dependence of string loop corrections to the gauge coupling constants*, Nuclear Phys. B **355** (1991), 649–688.

[73] S. K. Donaldson and P. B. Kronheimer, *The geometry of four-manifolds*, Oxford Univ. Press, Oxford, 1990.

[74] S. K. Donaldson and R. P. Thomas, *Gauge theory in higher dimensions*, in *The geometric universe: Science, geometry and the work of Roger Penrose*, (S. A. Huggett et al, eds.), Oxford Univ. Press, Oxford, 1998, pp. 31–47.

[75] M. R. Douglas, *D-branes, categories and $N = 1$ supersymmetry*, J. Math. Phys.textbf42 (2001), 2818–2843.

[76] M. R. Douglas, B. Fiol, and C. Römelsberger, *Stability and BPS branes*, preprint (hep-th/0002037).

[77] B. Dubrovin, *The geometry of 2D topological field theories*, in *Integrable Systems and Quantum Groups*, Lecture Notes in Math. vol. 1620, Springer–Verlag, Berlin, 1996, pp. 120–348.

[78] D. Edidin, W. Graham, *Localization in equivariant intersection theory and the Bott residue formula*, Amer. J. Math. **120** (1998), 619–636.

[79] T. Eguchi, P. B. Gilkey and A. J. Hanson, *Gravitation, gauge theories and differential geometry*, Physics Reports vol. 66., North-Holland, Amsterdam, 1980.

[80] T. Eguchi, K. Hori and C. S. Xiong, *Gravitational quantum cohomology*, Internat. J. Modern Phys. A **12** (1997), 1743–1782.

[81] T. Eguchi, and S. K. Yang, *$N = 2$ superconformal models as topological field theories*, Mod. Phys. Lett. A **5** (1990), 1693–1701.

[82] S. Elitzur, E. Gross, E. Rabinovici and N. Seiberg, *Aspects of bosonization in string theory*, Nuclear Phys. B **283** (1987), 413–432.

[83] C. Faber, *Algorithms for computing intersection numbers, with an application to the class of the locus of Jacobians*, in *New Trends in Algebraic Geometry* (Warwick, 1996), London Math. Soc. Lecture Note Ser. vol .264, Cambridge Univ. Press, Cambridge, 1999, pp. 93-109.

[84] C. Faber and R. Pandharipande, *Hodge integrals and Gromov–Witten theory*, Invent. Math. **139** (2000), 173–199.

[85] C. Faber and R. Pandharipande, *Hodge integrals, partition matrices, and the λ_g conjecture*, Ann. Math. **157** (2003), 97–124.

[86] C. Faber and R. Pandharipande, *Logarithmic series and Hodge integrals in the tautological ring* (with an appendix by Don Zagier), Michigan Math. J. **48** (2000), 215–252.

[87] B. Fantechi, L. Göttsche, and D. van Straten, *Euler number of the compactified Jacobian and multiplicity of rational curves*, J. Algebraic Geom. **8** (1999), 115–133.

[88] P. Fendley and K. A. Intriligator, *Scattering and thermodynamics in integrable $N = 2$ theories*, Nuclear Phys. B **380** (1992), 265–290.

[89] S. Ferrara, ed., *Supersymmetry*, North–Holland/World Sci., Amsterdam/Singapore, 1987.

[90] S. Ferrara, J. A. Harvey, A. Strominger, and C. Vafa, *Second-quantized mirror symmetry*, Phys. Lett. B **361** (1995), 59–65.

[91] A. Floer, *Morse theory for Lagrangian intersections*, J. Differential Geom. **28** (1988), 513–547.

[92] D. S. Freed, *Five lectures on supersymmetry*, Amer. Math. Soc., Providence, RI, 1999.

[93] D. H. Friedan, *Nonlinear models in $2+\varepsilon$ dimensions*, Phys. Rev. Lett. **45** (1980), 1057–1060.

[94] D. Friedan and P. Windey, *Supersymmetry and index theorems*, Phys. D **15** (1985), 71–74.

[95] K. Fujikawa, *Chiral anomaly and the Wess-Zumino condition*, Phys. Rev. D (3) **31** (1985), 341–351.

[96] K. Fukaya, *Morse homotopy, A^∞-category, and Floer Homologies*, in *The proceedings of the 1993 GARC workshop on geometry and topology*, (H. J. Kim, ed.), Lecture Notes Ser. vol. 18, Seoul National University, Seoul, 1993.

[97] K. Fukaya, Y.-G. Oh, H. Ohta and K. Ono, *Lagrangian intersection Floer theory — anomaly and obstruction*, to appear, Internat. Press.

[98] K. Fukaya and P. Seidel, *Floer homology, A^∞-categories and topological field theory*, in *Geometry and physics* (Aarhus, 1995), Lecture Notes in Pure and Appl. Math. vol. 184, Dekker, New York, 1997, pp. 9–32.

[99] W. Fulton, *Introduction to toric varieties*, Princeton Univ. Press, Princeton, 1993.

[100] W. Fulton, *Intersection Theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete vol. 2, Springer–Verlag, Berlin, 1998.

[101] W. Fulton and R. MacPherson, *A compactification of configuration spaces*, Ann. Math. **139** (1994), 183–225.

[102] W. Fulton and R. Pandharipande, *Notes on stable maps and quantum cohomology*, in *Algebraic geometry* (Santa Cruz, 1995), Proc. Sympos. Pure Math. vol. 62, Part 2, Amer. Math. Soc., Providence, RI, 1997, pp. 45–96.

[103] S. J. Gates, et al., *Superspace, or one thousand and one lessons in supersymmetry*, Frontiers in physics vol. 58, Benjamin/Cummings, Reading, MA, 1983.

[104] S. J. Gates, C. M. Hull and M. Roček, *Twisted multiplets and new supersymmetric nonlinear σ-models*, Nuclear Phys. B **248** (1984), 157–186.

[105] A. Gathmann, *Relative Gromov-Witten invariants and the mirror formula*, prerpint, 2000 (math.AG/0009190), to appear in Math. Ann.

[106] S. Gelfand and Yu. Manin, *Methods of homological algebra*, Springer–Verlag, Berlin, 1996.

[107] D. Gepner, *Exactly solvable string compactifications on manifolds of $SU(N)$ holonomy*, Phys. Lett. B **199** (1987), 380–388; *Space-time supersymmetry in compactified string theory and superconformal models*, Nuclear Phys. B **296** (1988), 757–778.

[108] A. Givental, *Equivariant Gromov–Witten invariants*, Int. Math. Res. Notices **13** (1996), 613–663.

[109] A. Givental, *A mirror theorem for toric complete intersections*, in *Topological field theory, primitive forms and related topics* (Kyoto, 1996) (M. Kashiwara et al, eds.), Progr. Math. vol 160, Birkhäuser Boston, Boston, MA, 1998, pp. 141–175.

[110] A. Givental, *The mirror formula for quintic three-folds*, in *Northern California symplectic geometry seminar*, Amer. Math. Soc. Transl. Ser. 2 vol. 196, Amer. Math. Soc., Providence, RI, 1999, pp. 49–62.

[111] A. Giveon, M. Porrati and E. Rabinovici, *Target space duality in string theory*, Phys.Rept. **244** (1994), 77–202.

[112] R. Gopakumar and C. Vafa, *Topological gravity as large N topological gauge theory*, Adv. Theor. Math. Phys. **2** (1998), 413–442.

[113] R. Gopakumar and C. Vafa, *Branes and fundmental groups*, Adv. Theor. Math. Phys. **2** (1998), 339–411.

[114] R. Gopakumar and C. Vafa, *M-theory and Topological Strings* I & II, preprints, 1998 (hep-th/9809187, hep-th/9812127).

[115] R. Gopakumar and C. Vafa, *On the gauge theory/geometry correspondence*, Adv. Theor. Math. Phys. **3** (1999), 1415–1443.

[116] S. Govindarajan, T. Jayaraman and T. Sarkar, *Worldsheet approaches to D-branes on supersymmetric cycles*, Nuclear Phys. B **580** (2000), 519–547.

[117] M. Göckeler and T. Schücker, *Differential, geometry, gauge theories, and gravity*, Cambridge Univ. Press, Cambridge, 1989.

[118] T. Graber and R. Pandharipande, *Localization of virtual classes*, Invent. Math. **135** (1999), 487–518.

[119] M. Grayson, *Shortening embedded curves*, Ann. Math. **129** (1989), 71–111.

[120] M. B. Green, J. H. Schwarz and E. Witten, *Superstring theory*, vol. 2, Cambridge Univ. Press, Cambridge, 1987.

[121] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley-Interscience, New York, 1978.

[122] B. Greene and M. R. Plesser, *Duality in Calabi-Yau moduli space*, Nuc. Phys. B **338** (1990), 15–37.

[123] B. R. Greene, M. R. Plesser and S. S. Roan, *New constructions of mirror manifolds: Probing moduli space far from Fermat points*, in *Essays on mirror manifolds* (S.-T. Yau, ed.), Internat. Press, Hong Kong, 1992, pp. 347-389.

[124] B. R. Greene, C. Vafa and N. P. Warner, *Calabi-Yau manifolds and renormalization group flows*, Nuclear Phys. B **324** (1989), 371–390.

[125] M. T. Grisaru, W. Siegel and M. Roček, *Improved methods for supergraphs*, Nuclear Phys. B **159** (1979), 429–450.

[126] M. Gross, *Topological mirror symmetry*, Invent. Math **144** (2001), 75-137.

[127] J. Harris and I. Morrison, *Moduli of curves*, Graduate Texts in Math. vol. 187, Springer–Verlag, New York, 1998.

[128] R. Hartshorne, *Algebraic geometry*, Springer–Verlag, New York, 1997.

[129] F. R. Harvey and H. B. Lawson, *Calibrated geometries*, Acta Math. **148** (1982), 47–157.

[130] N. Hitchin, *Lectures on special Lagrangian submanifolds*, *Winter school on mirror symmetry, vector bundles and Lagrangian submanifolds* (Cambridge, MA, 1999) (C. Vafa and S.-T. Yau, eds.), AMS/IP Stud. Adv. Math. vol. 23, Amer. Math. Soc., Providence, RI, 2001, pp. 151–182.

[131] P. Horava, *Strings On world sheet orbifolds*, Nuclear Phys. B textbf327 (1989), 461–484; *Background duality of open string models*, Phys. Lett. B **231** (1989), 251–257.

[132] K. Hori, *Linear models of supersymmetric D-branes*, in *Symplectic geometry and mirror symmetry* (Seoul, 2000) (K. Fukaya et al, eds.), World Sci., River Edge, NJ, 2001, pp. 111–186.

[133] K. Hori, A. Iqbal and C. Vafa, *D-branes and mirror symmetry*, preprint, 2000 (hep-th/0005247).

[134] K. Hori and Y. Oz, *F-Theory, T-Duality on $K3$ surfaces and $N = 2$ supersymmetric gauge theories in four dimensions*, Nuclear Phys. B **501** (1997), 97–108.

[135] K. Hori and C. Vafa, *Mirror symmetry*, preprint, 2000 (hep-th/0002222).

[136] S. Hosono, A. Klemm, S. Theisen and S.-T. Yau, *Mirror symmetry, mirror map and applications to Calabi–Yau hypersurfaces*, Comm. Math. Phys. **167** (1995) 301–350; *Mirror symmetry, mirror map and applications to Calabi–Yau hypersurfaces*, Nuclear Phys. B **433** (1995) 501–554.

[137] S. Hosono, M.-H. Saito and A. Takahashi, *Holomorphic anomaly equation and BPS state counting of rational elliptic surface*, Adv. Theor. Math. Phys. **1** (1999), 177–208.

[138] S. Hosono, M.-H. Saito and A. Takahashi, *Relative Lefschetz action and BPS state counting*, Internat. Math. Res. Notices **15** (2001), 783–816.

[139] N. Ishibashi, *The boundary and crosscap states in conformal field theories*, Mod. Phys. Lett. A **4**, 1989, 251–264.

[140] D. D. Joyce, *On counting special Lagrangian homology 3-spheres*, in *Topology and Geometry: Commemorating SISTAG* (A. J. Berrick, M. C. Leung and X. W. Xu, eds.), Cont. Math. vol .314, Amer. Math. Soc., Providence, RI, 2002, pp. 125–151.

[141] S. Kachru, S. Katz, A. Lawrence, J. McGreevy, *Open string instantons and superpotentials*, Phys. Rev. D(3) **62** (2000), 026001, 9 pp.

[142] S. Kachru, S. Katz, A. Lawrence, J. McGreevy, *Mirror symmetry for open strings*, Phys. Rev. D(3) **62** (2000), 126005, 14 pp.

[143] A. Kapustin and D. Orlov, *Remarks on A-branes, mirror symmetry, and the Fukaya category*, preprint, 2001 (hep-th/0109098).

[144] M. Kashiwara and P. Schapira, *Sheaves on manifolds*, Springer–Verlag, Berlin, 1990.

[145] S. Katz, *On the finiteness of rational curves on quintic threefolds*, Compositio Math. **60** (1986), 151–162.

[146] S. Katz, A. Klemm and C. Vafa, *Geometric engineering of quantum field theories*, Nuclear Phys. B **497** (1997), 173–195.

[147] S. Katz, A. Klemm, and C. Vafa, *M-Theory, topological strings and spinning black holes*, Adv. Theor. Math. Phys. **3** (1999), 1445–1537.

[148] S. Katz and C. C. Liu, *Enumerative geometry of stable maps with Lagrangian boundary conditions and multiple covers of the disc*, Adv. Theor. Math. Phys. **5** (2001), 1–49.

[149] S. Katz, P. Mayr and C. Vafa, *Mirror symmetry and exact solution of $4D$ $N = 2$ Gauge Theories* I, Adv. Theor. Math. Phys. **1** (1998), 53–114.

[150] S. Katz and D. R. Morrison, *Gorenstein threefold singularities with small resolutions via invariant theory for Weyl groups*, J. Algebraic Geom. **1** (1992), 449–530.

[151] S. Katz and C. Vafa, *Geometric engineering of $N = 1$ quantum field theories*, Nuclear Phys. B **497** (1997), 196–204.

[152] S. Keel, *Intersection theory of moduli space of stable n-pointed curves of genus zero*, Trans. Amer. Math. Soc. **330** (1992), 545–574.

[153] G. Kempf, F. Knudsen, D. Mumford and B. Saint-Donat, *Toroidal Embeddings* I, Lect. Notes in Math. vol. 339, Springer–Verlag, Berlin, 1973.

[154] K. Kikkawa and M. Yamasaki, *Casimir effects in superstring theories*, Phys. Lett. B **149** (1984), 357–360.

[155] B. Kim, *Quantum hyperplane section theorem for homogeneous spaces*, Acta Math. **183** (1999), 71–99.

[156] F. Kirwan, *Cohomology of quotients in symplectic and algebraic geometry*, Mathematical Notes vol. 31, Princeton Univ. Press, Princeton, 1984.

[157] S. Kleiman and T. Johnsen, *Rational curves of degree at most* 9 *on a general quintic threefold*, Comm. Algebra **24** (1996), 2721–2753.

[158] S. Kleiman and T. Johnsen, *Toward Clemens' conjecture in degrees between* 10 *and* 24, Serdica Math. J. **23** (1997), 131–142.

[159] A. Klemm, *On the geometry behind $N = 2$ supersymmetric actions in four dimensions*, preprint, 1997 (hep-th/9705131).

[160] A. Klemm and E. Zaslow, *Local mirror symmetry at higher genus*, in *Winter school on mirror symmetry, vector bundles and Lagrangian submanifolds* (Cambridge, MA, 1999) (C. Vafa and S.-T. Yau, eds.), AMS/IP Stud. Adv. Math. vol. 23, Amer. Math. Soc., Providence, RI, 2001, pp. 183–207.

[161] K. Kodaira and D. C. Spencer, *on Deformations of Complex Analytic Structures*, I-II, III, Ann. Math. **67** (1958), 328–466; **71** (1960), 43–76.

[162] M. Kontsevich, *Intersection theory on the moduli space of curves and the matrix Airy function*, Comm. Math. Phys. **147** (1992), 1–23.

[163] M. Kontsevich, *Enumeration of rational curves via torus actions*, in *The Moduli Space of Curves* (Texel Island, 1994) (R. Dijkgraaf, C. Faber and G. van der Geer, eds.), Progr. Math. vol. 129, Birkhäuser, Boston, 1995, pp. 335–368.

[164] M. Kontsevich, *Homological algebra of mirror symmetry*, in *Proceedings of the International Congress of Mathematicians* (Zürich, 1994), Birkhäuser, Boston, 1995, pp. 120–139.

[165] M. Kontsevich and Yu. Manin, *Gromov-Witten classes, quantum cohomology, and enumerative geometry*, Comm. Math. Phys. **164** (1994), 525–562.

[166] J. M. Labastida and M. Marino, *A new point of view in the theory of knot and link invariants*, J. Knot Theory Ramifications **11** (2002), 173–197.

[167] J. M. Labastida, M. Marino and C. Vafa, *Knots, links and branes at large N*, J. High Energy Phys. (2000), no 11, paper 7, 42 pp.

[168] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics, non-relativistic theory*, Pergamon Press, Oxford, 1977.

[169] G. Lawlor, *The angle criterion*, Invent. Math. **95** (1989), 437–446.

[170] Y.-P. Lee, *Quantum Lefschetz hyperplane theorem*, Invent. Math. **145** (2001) 121–149.

[171] R. G. Leigh, *Dirac-Born-Infeld action from Dirichlet σ-model*, Mod. Phys. Lett. A **4** (1989), 2767–2772.

[172] W. Lerche, *Lecture on $N = 2$ supersymmetric gauge theory*, in *Symétries quantiques* (Les Houches. 1995) (A. Connes, K. Gawedzki and J. Zinn-Justin, eds.), North-Holland, Amsterdam, 1998, pp. 613–640.

[173] W. Lerche, C. Vafa and N. P. Warner, *Chiral rings in $N = 2$ superconformal theories*, Nuclear Phys. B **324** (1989), 427–474.

[174] N. C. Leung, *Einstein type metrics and stability on vector bundles*, Jour. Differential Geom. **45** (1997), 514–546.

[175] N. C. Leung, *Symplectic structures on gauge theory*, Comm. Math. Phys. **193** (1998), 47–67.

[176] N. C. Leung, S.-T. Yau, and E. Zaslow, *From special Lagrangian to Hermitian-Yang-Mills via real Fourier-Mukai transform*, in *Winter school on mirror symmetry, vector bundles and Lagrangian submanifolds* (Cambridge, MA, 1999) (C. Vafa and S.-T. Yau, eds.), AMS/IP Stud. Adv. Math. vol. 23, Amer. Math. Soc., Providence, RI, 2001, pp. 209–225

[177] J. Li, *Algebraic geometric interpretation of Donaldson's polynomial invariants*, J. Differential Geom. **37** (1993), 417–466.

[178] J. Li and G. Tian, *Virtual moduli cycles and Gromov-Witten invariants of algebraic varieties*, J. Amer. Math. Soc. **11** (1998), 119–174; *Virtual moduli cycles and Gromov-Witten invariants of general symplectic manifolds*, in *Topics in symplectic 4-manifolds* (Irvine, CA, 1996) (R. J. Stern, ed.), Internat. Press, Cambridge, MA, 1998, pp. 47–83.

[179] B. Lian and S.-T. Yau, *Arithmetic properties of mirror map and quantum coupling*, Comm. Math. Phys. **176** (1996), 163–191.

[180] B. Lian, K. Liu, and S.-T. Yau, *Mirror principle* I, Asian J. of Math, **1** (1997), 729–763.

[181] B. Lian, K. Liu, and S.-T. Yau, *Mirror principle* II, Asian J. Math. **3** (1999), 109–146.

[182] B. Lian, K. Liu, and S.-T. Yau, *Mirror principle* III, Asian J. Math. **3** (1999), 771–800.

[183] B. Lian, K. Liu, and S.-T. Yau, *Mirror principle* IV, Surv. Differ. Geom. VII, Internat. Press, Somerville, MA, 2000, pp. 475–496.

[184] J. Maldacena, G. W. Moore and N. Seiberg, *Geometrical interpretation of D-branes in gauged WZW models*, J. High Energy Phys. (2001), no. 7, paper 46, 63 pp.

[185] S. Mandelstam, *Soliton operators for the quantized sine-Gordon equation*, Phys. Rev. D **11** (1975), 3026–3030.

[186] Yu. Manin, *Generating functions in algebraic geometry and sums over trees*, in *The moduli space of curves* (Texel Island, 1994) (R. Dijkgraaf, C. Faber and G. van der Geer, eds.), Progr. Math. vol. 129, Birkhäuser Boston, Boston, 1995, pp. 401–417.

[187] Yu. Manin, *Frobenius manifolds, quantum cohomology, and moduli spaces*, AMS Colloq. Publ. vol. 47, Amer. Math. Soc., Providence, RI, 1999.

[188] M. Mariño, R. Minasian, G. Moore, and A. Strominger, *Nonlinear instantons from supersymmetric p-branes*, J. High Energy Phys. (1999), no. 11, paper 21, 29 pp.

[189] E. J. Martinec, *Algebraic geometry and effective Lagrangians*, Phys. Lett. B **217** (1989), 431–437; *Criticality, catastrophes and compactifications*, in *Physics and mathematics of strings* (L. Brink, D. Friedan and A.M. Polyakov, eds.), World Sci.,Teaneck, NJ, 1990, pp. 389–433.

[190] P. Mayr, *Geometric construction of $N = 2$ gauge theories*, Fortschr. Phys. **47** (1999), 39–63.

[191] R. McLean, *Deformations of calibrated submanifolds*, Comm. Anal. Geom. **6** (1998), 705–747.

[192] N. D. Mermin and H. Wagner, *Absence of ferromagnetism or antiferromagnetism in one-dimensional or two-dimensional isotropic Heisenberg models*, Phys. Rev. Lett. **17** (1966), 1133–1136.

[193] J. Milnor, *Morse theory,* Annals of Math. Studies vol. 51, Princeton Univ. Press, Princeton, 1963.

[194] J. W. Milnor and J. D. Stasheff, *Characteristic classes*, Princeton Univ. Press, Princeton, 1974.

[195] R. Minasian and G. Moore, *K-theory and Ramond-Ramond charge*, J. High Energy Phys. (1997), no. 11, paper 2, 7 pp.

[196] D. R. Morrison and M. R. Plesser, *Summing the instantons: Quantum cohomology and mirror symmetry in toric varieties*, Nuclear Phys. B **440** (1995), 279–354.

[197] S. Mukai, *Duality between $D(X)$ and $D(\widehat{X})$ with its application to Picard sheaves*, Nagoya Math. Jour. **81** (1981), 153–175.

[198] S. Mukai, *On the moduli space of bundles on K3 surfaces,* I, in *Vector bundles on algebraic varieties* (Bombay, 1984), Tata Inst. Fund. Res. Stud. Math. vol. 11, Tata Inst. Fund. Res., Bombay, 1987, pp. 341–413.

[199] D. Mumford, *Towards an enumerative geometry of the moduli space of curves*, in *Arithmetic and geometry, vol. II*, (M. Artin, J. Tate, eds.) Progr. Math. vol. 36, Birkhäuser Boston, Boston, 1983, pp. 271–328.

[200] H. Nakajima, *Lectures on Hilbert schemes of points on surfaces*, Univ. Lect. Ser. vol. 18, Amer. Math. Soc., Providence, RI, 1999.

[201] K. S. Narain, M. H. Sarmadi and E. Witten, *A note on toroidal compactification of heterotic string theory*, Nuclear Phys. B **279** (1987), 369–379.

[202] A. Newlander and L. Nirenberg, *Complex analytic voordinates in almost-vomplex manifolds*, Ann. Math. **65** (1957), 391–404.

[203] T. Oda, *Convex bodies and algebraic geometry*, Springer–Verlag, Berlin, 1988.

[204] H. Ooguri, Y. Oz, and Z. Yin, *D-branes on Calabi-Yau spaces and their mirrors*, Nuclear Phys. B **477** (1996), pp. 407–430.

[205] H. Ooguri and C. Vafa, *Knot invariants and topological strings*, Nuclear Phys. B **577** (2000), 419–438.

[206] D. O. Orlov, *Equivalences of derived categories and K3 surfaces*, J. Math. Sci. **84** (1997), 1361–1381.

[207] R. Pandharipande, *Rational curves on hypersurfaces (after A. Givental)*, Astérisque **252** (1998), 307–340.

[208] R. Pandharipande, *Hodge integrals and degenerate contributions*, Comm. Math. Phys. **208** (1999), 489–506.

[209] M. E. Peskin, D. V. Schroeder, *An introduction to quantum field theory*, Perseus, Reading, MA, 1995.

[210] J. Polchinski, *Dirichlet branes and Ramond-Ramond charges*, Phys. Rev. Lett. **75** (1995), 4724–4727.

[211] J. Polchinski, *Lectures on D-branes*, in *Fields, strings and duality* (Boulder, CO, 1996) (C. Efthimiou and B. Greene, eds.), World Sci., River Edge, NJ, 1997, pp. 293–356.

[212] J. Polchinski, *String theory* I, Cambridge Univ. Press, Cambridge, 1998.

[213] J. Polchinski and Y. Cai,]textitConsistency of open superstring theories, Nuclear Phys. B **296** (1988), 91–128.

[214] J. Polchinski, S. Chaudhuri and C. V. Johnson, *Notes on D-branes*, preprint, 1996 (hep-th/9602052).

[215] A. Polishchuk and E. Zaslow, *Categorical mirror symmetry: The elliptic curve*, Adv. Theor. Math. Phys. **2** (1998), 443–470.

[216] D. Quillen *Determinants of Cauchy-Riemann operators over a Riemann surface*, Funkts. Anal. i Prilozn. **19** (1985), 37–41.

[217] D. Quillen, *Superconnections and the Chern character*, Topology **24** (1985), 89–95.

[218] D. B. Ray and I. M. Singer, *Analytic torsion for complex manifolds*, Ann. Math. **98** (1973), 154–177.

[219] M. Reid, *Decomposition of toric morphisms*, in *Arithmetic and Geometry, vol. II* (M. Artin and J. Tate, eds.), Progr. Math. vol. 36, Birkhäuser, Boston, 1983, pp. 395–418.

[220] M. Roček and E. Verlinde, *Duality, quotients, and currents*, Nuclear Phys. B **373** (1992), 630–646.

[221] W.-D. Ruan, *Lagrangian tori fibration of toric Calabi-Yau Manifold III: Symplectic topological SYZ mirror construction for general quintics*, preprint, 1999 (math.DG/9909126), to appear in J. Differential Geom.

[222] Y. Ruan and G. Tian, *A mathematical theory of quantum cohomology*, J. Differential Geom. **42** (1995), 259–367.

[223] N. Sakai and I. Senda, *Vacuum energies of string compactified on torus*, Prog. Theor. Phys. **75** (1986), 692–705. [Erratum: ibid. **77** (1986), 773.]

[224] R. Schoen and J. Wolfson, *Minimizing volume among Lagrangian submanifolds*, in *Differential Equations: La Pietra 1996* (Florence) (M. Giaquinta, J. Shatah and S. R. S. Varadhan, eds.), Proc. Sympos. Pure Math. vol. 65, Amer. Math. Soc., Providence, RI, 1999, pp. 181–199.

[225] N. Seiberg, *Naturalness versus supersymmetric nonrenormalization theorems*, Phys. Lett. B **318** (1993), 469–475.

[226] P. Seidel, *Lagrangian two-spheres can be symplectically knotted*, J. Differential Geom. **52** (1999), 145–171.

[227] P. Seidel, *Graded Lagrangian submanifolds*, Bull. Soc. Math. France. **128** (2000), 103–146.

[228] P. Seidel, *Vanishing cycles and mutation*, in *European Congress of Mathematics, vol. II* (Barcelona, 2000), Progr. Math. vol. 202, Birkhäuser, Boston, 2001, pp. 65–85; *More about vanishing cycles and mutation*, in *Symplectic geometry and mirror symmetry* (Seoul, 2000) (K. Fukaya et al, eds.), World Sci., River Edge, NJ, 2001, pp. 429–465.

[229] P. Seidel and R. P. Thomas, *Braid group actions on derived categories of sheaves*, Duke Math. J. **108** (2001), 37–108.

[230] A. Sen, *Stable non-BPS bound states of BPS D-branes*, J. High Energy Phys. (1998), no. 8, paper 10, 25 pp.;

[231] A. Sen, *Tachyon condensation on the brane antibrane system*, J. High Energy Phys. (1998), no. 8, paper 12, 6pp.

[232] A. Shapere and C. Vafa, *BPS structure of Argyres-Douglas superconformal theories*, preprint, 1999 (hep-th/9910182).

[233] K. Smoczyk, *A canonical way to deform a Lagrangian submanifold*, preprint, 1996 (dg-ga/9605005).

[234] A. Strominger, S.-T. Yau, and E. Zaslow, *Mirror Symmetry is T-Duality*, Nuclear Phys. B **479** (1996), 243–259.

[235] G. Tian, *Smoothness of the universal deformation space of compact Calabi-Yau manifolds and its Petersson-Weil metric*, in *Mathematical Aspects of String Theory* (San Diego, CA, 1986) (S.-T. Yau, ed.), Adv. Ser. Math. Phys. vol. 1, World Sci., Singapore, 1987, pp. 629–646.

[236] R. P. Thomas, *A holomorphic Casson invariant for Calabi-Yau 3-folds, and bundles on K3 fibrations*, J. Differential Geom. **54** (2000), 367–438.

[237] R. P. Thomas, *An obstructed bundle on a Calabi-Yau 3-fold*, Adv. Theor. Math. Phys. **3** (1999), 567–576.

[238] R. P. Thomas, *Derived categories for the working mathematician*, in *Winter school on mirror symmetry, vector bundles and Lagrangian submanifolds* (Cambridge, MA, 1999) (C. Vafa and S.-T. Yau, eds.), AMS/IP Stud. Adv. Math. vol. 23, Amer. Math. Soc., Providence, RI, 2001, pp. 349–361.

[239] R. P. Thomas, *Moment maps, monodromy and mirror manifolds*, in *Symplectic geometry and mirror symmetry* (Seoul, 2000) (K. Fukaya et al, eds.), World Sci., River Edge, NJ, 2001, pp. 467–498.

[240] R. P. Thomas and S.-T. Yau, *Special Lagrangians, stable bundles and mean curvature flow*, Comm. Anal. Geom. **10** (2002), 1053–1085.

[241] A. N. Todorov, *The Weil-Petersson geometry of the moduli space of* $\mathrm{SU}(n \geq 3)$ *(Calabi-Yau) manifolds*, I, Comm. Math. Phys. **126** (1989), 325–346.

[242] A. N. Tyurin, *Nonabelian analogues of Abel's theorem*, Izv. Ross. Akad. Nauk Ser. Mat. **65** (2001), 133–196; transl. in Izv. Math. **65** (2001), 123-180.

[243] K. Uhlenbeck and S.-T. Yau, *On the existence of Hermitian–Yang–Mills connections in stable vector bundles*, Comm. Pure Appl. Math. **39** (1986), 257–S293.

[244] C. Vafa, *Topological Landau-Ginzburg models*, Mod. Phys. Lett. A **6** (1991), 337–346.

[245] C. Vafa, *Topological mirrors and quantum rings*, in *Essays on mirror manifolds* (S.-T. Yau, ed.), Internat. Press, Hong Kong, 1992, pp. 96–119.

[246] C. Vafa, *Extending Mirror Conjecture to Calabi-Yau with Bundles*, Comm. Contemp. Math. **1** (1999), 65–70.

[247] C. Vafa and N. P. Warner, *Catastrophes and the classification of conformal theories*, Phys. Lett. B **218** (1989), 51–58.

[248] I. Vainsencher, *Enumeration of* n*-fold tangent hyperplanes to a surface*, J. Alebraic. Geom. **4** (1995), 503–526.

[249] C. Voisin, *A mathematical proof of a formula of Aspinwall and Morrison*, Comp. Math. **104** (1996), 135–151.

[250] C. Weibel, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. vol. 38, Cambridge Unive. Press, Cambridge, 1994.

[251] S. Weinberg, *The quantum theory of fields. Vol. 1: Foundations*, Cambridge Univ. Pr., Cambridge, 1995.

[252] J. Wess and J. Bagger, *Supersymmetry and supergravity*, Princeton Univ. Press, 1983.

[253] J. Wess and B. Zumino, *A Lagrangian model invariant under supergauge transformations*, Phys. Lett. B **49** (1974), 52–54.

[254] K. G. Wilson and J. B. Kogut, *The renormalization group and the epsilon expansion*, Phys. Rept. C **12** (1974), 75–200.

[255] K. G. Wilson, *The renormalization group: Critical phenomena and the Kondo problem*, Rev. Mod. Phys. **47** (1975), 773–840.

[256] E. Witten, *A supersymmetric form of the nonlinear* σ *model in two dimensions*, Phys. Rev. D **16** (1977), 2991–2994.

[257] E. Witten, *Instantons, the quark model, and the* $1/N$ *expansion*, Nuclear Phys. B **149** (1979), 285–320.

[258] E. Witten, *Dynamical breaking of supersymmetry*, Nuclear Phys. B **188** (1981), 513–554.

[259] E. Witten, *Constraints on supersymmetry breaking*, Nuclear Phys. B **202** (1982), 253–316.

[260] E. Witten, *Supersymmetry and Morse theory*, J. Differential Geom. **17** (1982), 661–692.

[261] E. Witten, *Topological sigma models*, Comm. Math. Phys. **118** (1988), 411–449.

[262] E. Witten, *On the structure of the topological phase of two-dimensional gravity*, Nuclear Phys. B **340** (1990), 281–332.

[263] E. Witten, *Two-dimensional gravity and intersection theory on moduli space*, in *Survey in Differential Geometry* (Cambridge, 1990), Lehigh Univ., Bethlehem, PA, 1991, pp. 243–310.

[264] E. Witten, *Mirror manifolds and topological field theory*, in *Essays on mirror manifolds* (S.-T. Yau, ed.), Internat. Press, Hong Kong, 1992, pp. 121–160.

[265] E. Witten, *Phases of* $N = 2$ *theories in two dimensions*, Nuclear Phys. B **403** (1993), 159–222.

[266] E. Witten, *The Verlinde algebra and the cohomology of the Grassmannian*, in *Geometry, topology, and physics* (Cambridge, 1993) (S.-T. Yau, ed.), Conf. Proc. Lecture Notes Geom. Topology vol. IV, Internat. Press, Cambridge, MA, 1995, pp. 357–422.

[267] E. Witten, *Chern-Simons gauge theory as a string theory*, in *The Floer memorial volume* (H. Hofer et al, eds.), Progr. Math. vol. 133, Birkhäuser, Boston, MA, 1995, pp. 637–678.

[268] E. Witten, *Bound states of strings and p-branes*, Nuclear Phys. B **460** (1996), 335–350.

[269] E. Witten, *D-branes and* K*-theory*, J. High Energy Phys. (1998), no. 12, paper 19, 41 pp.

[270] S.-T. Yau, *On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation*, I, Comm. Pure Appl. Math. **31** (1978), 339–411.

[271] S.-T. Yau and E. Zaslow, *BPS states, string duality, and nodal curves on K3*, Nuclear Phys. B **471** (1996), 503–512.

[272] B. Zumino, *Supersymmetry and Kähler manifolds*, Phys. Lett. B **87** (1979), 203–206.

Index