Lectures on topological field theories

PRELIMINARY ROUGH DRAFT[1] – last revised 6-23-07, 10 am

Eric Sharpe

These lecture notes cover my talks at the June 2007 derived categories minicourse at the University of Utah. I begin by giving a very brief introduction to quantum field theory, quickly moving to nonlinear sigma models and the associated A and B model topological field theories, and the physical realization of derived categories. Later lectures will cover Landau-Ginzburg models, matrix factorization, gauged linear sigma models, and mirror symmetry.

[1]This is a preliminary rough draft. As such, it is bound to contain a number of unintentional errors. Also, as such, this document is not suitable for distribution outside the minicourse.

Contents

1 Path integrals and basic QFT

We're going to define quantum field theories by doing calculus on infinite-dimensional function spaces, or, more accurately, approximations thereof.

Let's begin by introducing an infinite-dimensional derivative.

1.1 Functional derivative

Let $x(t)$ be a function of one variable t. This could be the position of a point-particle along a line, as a function of time. We can define a function that depends upon $x(t)$ – such a quanitity is known as a *functional* of $x(t)$. For example,

$$S[x(t)] \;=\; \int dt \left(\frac{dx}{dt}\right)^2$$

We would like to define a derivative on the space of all functions $x(t)$. Such a derivative should vary the value of the function $x(t)$ at a single point, and not others. Let us denote such a derivative by $\delta/\delta x(t')$. From the property above, when $t \neq t'$, we need

$$\frac{\delta}{\delta x(t')} x(t) \;=\; 0$$

However, when $t = t'$, we need the derivative to be nonzero. In some sense, we'd want it to be equal to 1 at that point. More precisely, we define

$$\frac{\delta}{\delta x(t')} x(t) \;=\; \delta(t - t')$$

where $\delta(t - t')$ is the *Dirac delta function*, a distribution with the following properties:

$$\begin{aligned} \delta(t - t') \;&=\; 0 \text{ for } t \neq t' \\ \textstyle\int_{-\infty}^{\infty} dt' f(t') \delta(t - t') \;&=\; f(t) \end{aligned}$$

for any 'well-behaved' function $f(t)$. Furthermore, we expect this infinite-dimensional or "functional" derivative $\delta/\delta x(t')$ to have all the properties one would expect of an ordinary derivative acting on differentiable functions: there should be a product rule, a chain rule, derivatives should commute with one another, and so forth. For example:

$$\begin{aligned} \tfrac{\delta}{\delta x(t')} \left(x(t)\right)^2 \;&=\; 2x(t)\delta(t - t') \\ \tfrac{\delta}{\delta x(t')} \left(x(t) y(t)\right) \;&=\; \delta(t - t') y(t) \end{aligned}$$

A typical application of this derivative is in the 'calculus of variations,' where one is asked to find, for example, a function that minimizes some functional.

Here is one simple example: show that the shortest distance between two points is a straight line. For simplicity, let us assume the points lie in a fixed plane, and consider paths which can be represented as functions $y(x)$, where x, y are the obvious coordinates on the plane. For a given path $y(x)$ between two fixed endpoints, the length of the path is defined by

$$s \;=\; \int \left(1 \,+\, (dy/dx)^2\right)^{1/2} dx$$

To find the path that minimizes the arc length s, let us compute $\delta s / \delta y(x')$ and set it equal to zero. (Morally, this is just like first-semester calculus, where you find the extremum of a function $f(x)$ by solving $f'(x) = 0$.) Let us compute:

$$\begin{aligned}
\frac{\delta}{\delta y(x')}\, s &= \frac{\delta}{\delta y(x')} \int \left(1 \,+\, (dy/dx)^2\right)^{1/2} dx \\
&= \int dx (1/2) \left(1 \,+\, (y')^2\right)^{-1/2} (2)(y') \frac{d}{dx} \left(\frac{\delta y(x)}{\delta y(x')} \right) \\
&= \int dx \left(1 \,+\, (y')^2\right)^{-1/2} y' \frac{d}{dx} \delta(x - x')
\end{aligned}$$

Now, the expression above contains a derivative of a Dirac delta function. To make sense of this, we integrate by parts, and use the fact that the boundary terms will vanish so long as x' is not at the edges:

$$\begin{aligned}
\frac{\delta}{\delta y(x')}\, s &= -\int dx \frac{d}{dx} \left[\left(1 \,+\, (y')^2\right)^{-1/2} y' \right] \delta(x - x') \\
&= -\frac{d}{dx} \left[\left(1 \,+\, (y')^2\right)^{-1/2} y' \right] \\
&= -\left(1 \,+\, (y')^2\right)^{-3/2} y''
\end{aligned}$$

To find the path $y(x)$ that minimizes the arc length s, we find the function $y(x)$ for which $\delta s / \delta y(x') = 0$. From the expression above, we see that a sufficient condition is $y'' = 0$, which implies $y = Ax + B$ for some constants A and B. This is the equation of a straight line, so we have just verified, using functional derivatives, that the shortest distance between two points is a straight line.

A more common physics application is a rewriting of classical mechanics due to Hamilton, and known as Hamilton's least-action principle: the path taken by any object minimizes its action, where the action is a quantity constructed from the kinetic and potential energies. The differential equations one obtains from expanding $\delta S[y(x)] / \delta y(x') = 0$ are known as the *equations of motion.*

With that in mind, let us compute the functional derivative of the following functional:

$$S[x(t)] \;=\; \int_{-\infty}^{\infty} dt \left(\frac{m}{2} \left(\frac{dx}{dt} \right)^2 \;-\; V(x) \right)$$

From the discussion above,

$$\begin{aligned}
\frac{\delta}{\delta x(t')} S[x(t)] \;&=\; \frac{\delta}{\delta x(t')} \int_{-\infty}^{\infty} dt \left(\frac{m}{2} \left(\frac{dx}{dt} \right)^2 \;-\; V(x) \right) \\
&=\; \int_{-\infty}^{\infty} \left(m \frac{dx}{dt} \frac{d}{dt} \frac{\delta x(t)}{\delta x(t')} \;-\; V'(x) \frac{\delta x(t)}{\delta x(t')} \right) \\
&=\; \int_{-\infty}^{\infty} \left(m \frac{dx}{dt} \frac{d}{dt} \delta(t-t') \;-\; V'(x) \delta(t-t') \right) \\
&=\; \int_{-\infty}^{\infty} \left(-m \frac{d^2 x}{dt^2} \delta(t-t') \;-\; V'(x) \delta(t-t') \right) \\
&=\; -m \frac{d^2 x}{dt^2} \;-\; V'(x)
\end{aligned}$$

where in the next to last step we have integrated by parts and assumed the boundary terms vanished. In this case, the equations of motion are

$$m \frac{d^2 x}{dt^2} \;=\; -V'(x)$$

A reader who remembers a bit of classical mechanics might recall this expression. The quantity d^2x/dt^2 is the acceleration a of a particle, and the net force F on a particle in a potential $V(x)$ is given by $F = -V(x)$. So our equations of motion say $F = ma$, which is one of Newton's laws.

1.2 Functional integrals

Hamilton's least-action principle is a very elegant rephrasing of classical mechanics – no more mucking about with force diagrams, just minimize an action functional. But, you might ask, *why* should it describe classical mechanics?

When Richard Feynman was a graduate student, he asked himself the same question. He eventually showed, as part of his Ph.D. thesis, the Hamilton's least-action principle expresses the leading effect in an approximation to quantum mechanics. In order to do this, he defined the integrals that go hand-in-hand with the functional derivatives introduced in the last section. Known as *functional integrals* or *path integrals*, these integrals – formally over spaces of functions – can be used to define quantum mechanics and quantum field theory.

Given a classical mechanical system described by some action functional S, quantum mechanics is encoded in the functional integral denoted

$$\int [Dx] \exp(-S/\hbar)$$

where $\hbar$ is *Planck's constant*, which encodes the strength of quantum corrections. When $\hbar$ is small, quantum effects are suppressed. We have not yet tried to make sense of this expression, but we can already see – at least formally – how Hamilton's least-action principle is going to emerge. From the method of steepest descent described in section A.3, one would expect that for small $\hbar$, the dominant contribution to the path integral will come from functions $x(t)$ such that $\delta S/\delta x = 0$, and other contributions should be, comparatively, exponentially suppressed. But $\delta S/\delta x = 0$ is exactly the statement of Hamilton's least-action principle.

Now, how does one define a path integral, at least formally? Let us consider functions of a single variable t. Suppose the action is an integral over the interval $[a, b]$. Just as an ordinary integral can be defined as the limit of a sequence of Riemann sums, we can formally try to define the path integral as the limit of a series of approximations. Let N be an integer, and split the interval $[a, b]$ into N equal-size pieces. Define $t_k = a + k(b - a)/N$, and define $x_k = x(t_k)$. To try to integrate over a space of functions $x(t)$, let us integrate over the values of $x(t)$ at the t_k, and interpolate function values in between those points. In other words, approximate $\int [Dx]$ by

$$C^N \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_{N-1}$$

for some suitable normalizing constant C. Suppose the action functional S is given by

$$\int_a^b \frac{1}{2} \left(\frac{dx}{dt} \right)^2 dt$$

then we could approximate the functional by

$$\sum_{j=0}^{N-1} \frac{1}{2} \left(\frac{x_{j+1} - x_j}{(b-a)/N} \right)^2 \left(\frac{b-a}{N} \right)$$

In this form, we could then try to define

$$\int [Dx] \exp(-S/\hbar) = \lim_{N \to \infty} C^N \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_{N-1} \exp\left[-\frac{1}{\hbar} \sum_{j=0}^{N-1} \frac{1}{2} \left(\frac{x_{j+1} - x_j}{(b-a)/N} \right)^2 \left(\frac{b-a}{N} \right) \right]$$

If we try to pursue this program, then there are some problems one will run into. For example, it is far from clear that different approximation schemes for any given action functional S will result in the same path integral. In fact, there is indeed such an ambiguity in quantum mechanics – a given quantum-mechanical system is not uniquely specified by a classical mechanical system. In formal discussions, this boils down to the Stone-von Neumann theorem. This is not a limitation of path integrals, but rather reflects the reality that quantum mechanics is a more nearly correct description of the physics, so to be precise, one must specify the quantum mechanical system, for which classical mechanics is then just an approximation.

When phrased in this language, quantum field theory is very similar to quantum mechanics – the primary distinction is that the functions that one integrates over, depend upon more than a single independent variable. For example, let ϕ be a real-valued function on $\mathbf{R}^4$, which you can interpret as space-time[2]. One can still define action functionals; a typical example has the form

$$\int d^4x \frac{1}{2}\left[\partial_\mu \phi \partial^\mu \phi \, + \, m^2 \phi^2 \right]$$

where $\partial_\mu = \partial / \partial x^\mu$, and we are using Einstein's summation convention, in which one sums over repeated indices. The function ϕ is known as[3] a "field," and the quantity m is known as the "mass" of the field ϕ. One can define a functional derivative, just as before:

$$\frac{\delta}{\delta \phi(x')} \phi(x) \; = \; \delta^4(x - x')$$

where $\delta^4(x - x')$ denotes a product of four Dirac delta functions, one for each coordinate on $\mathbf{R}^4$. One can then imagine defining a functional integral

$$\int [D\phi] \exp(-S[\phi]/\hbar)$$

Without even attempting to define a functional integral rigorously, the reader might already object that there is something fishy going on here. After all, if we attempt to define the functional integral as the limit of a product of ordinary integrals, then we are going to end up summing over infinitely jagged, thoroughly non-differentiable, fields when we take a limit $N \rightarrow \infty$, but a typical action functional S involves derivatives of the fields, not to mention that one typically only wants to work with differentiable functions.

This difficulty manifests itself in other ways also – for example, it is ultimately the reason for the famous infinities of quantum field theory.

There are several ways to resolve this difficulty, known as *regularization schemes*. Each involves changing the definition of the fields ϕ in some form. For the purposes of these notes, let us work with the simplest possibility, a "momentum-cutoff regularization." This means

[2] Throughout these notes, we shall always assume all metrics are Riemannian, so as to avoid subtleties in dealing with Minkowski metrics. In the present case, we are assuming a flat Euclidean metric $g_{\mu\nu} = \delta_{\mu\nu}$.

[3] More generally, any real-, complex-, or even (as we shall see later) Grassmann-valued function or section of a bundle that one sums over in a functional integral is known as a field. This is the why this subject is known as quantum *field* theory. Note this notion of field has absolutely nothing at all to do with the notion of a field in mathematics.

the following. Since ϕ is a function on a vector space (specifically, $\mathbf{R}^4$), we can express it in terms of its Fourier transform:

$$\phi(x) \;=\; \frac{1}{(2\pi)^4} \int d^4p \exp(-ip\cdot x)\tilde{\phi}(p)$$

Instead of summing over all possible Fourier components, let us restrict to Fourier components whose momentum has magnitude bounded by some cutoff Λ:

$$\phi_\Lambda(x) \;=\; \frac{1}{(2\pi)^4} \int_{|p|\leq\Lambda} d^4p \exp(-ip\cdot x)\tilde{\phi}(p)$$

The path integral then sums over ϕ_Λ's, not ϕ's:

$$\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda]/\hbar)$$

By imposing such a momentum cutoff, we have forcibly removed arbitrarily jagged configurations from the path integral, and so removed one difficulty with defining the path integral. We could then try to define the path integral as a limit of a product of integrals, say in momentum space rather than on the original $\mathbf{R}^4$.

More generally, it is believed that a 'classical' path integral, over ordinary fields/functions, does not exist in general; only such 'regularized' path integrals are believed to exist. I say 'believed' because no one has succeeded in giving a rigorous definition of a path integral, but from a physics perspective, a choice of regularization, such as the momentum cutoff above, should be an important part of any rigorous definition ultimately developed.

1.3 Feynman diagrams

Typically, in a quantum field theory, one wishes to compute *correlation functions*. In statistics, a correlation function is something of the form

$$< f_1 \cdots f_n > = \sum_{events} (\text{probability of each event}) f_1 \cdots f_n$$

where the f_i are functions on the space of events. Here, we think of the factor $\exp(-S/\hbar)$ as being analogous to an unnormalized probability, and we define the correlation function of $f_1, \cdots, f_n$ to be

$$< f_1 \cdots f_n > = \frac{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda]/\hbar) f_1 \cdots f_n}{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda]/\hbar)}$$

where the f_i are real analytic functions of the ϕ_Λ.

To be more specific, let us consider the correlation function

$$< \phi_\Lambda(x_1) \cdots \phi_\Lambda(x_n) > = \frac{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda]) \phi_\Lambda(x_1) \cdots \phi_\Lambda(x_n)}{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda])}$$

where we have omitted $\hbar$ for simplicity. To compute this, we first rewrite the correlation function above as

$$\frac{\int [D\phi_\Lambda] \frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \exp(-S[\phi_\Lambda] + \int d^4x J(x)\phi_\Lambda(x)))\Big|_{J\equiv 0}}{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda])} \tag{1}$$

If the action functional S is quadratic in the ϕ_Λ's, *e.g.* if it has the form

$$\int d^4x \frac{1}{2}\left[\partial_\mu \phi \partial^\mu \phi \;+\; m^2\phi^2\right]$$

then we can solve this exactly, by completing the square in the exponent.

To see how completing the square works for such operators, let us instead consider completing the square for matrices. Consider the matrix equation

$$(1/2)v^T AV \;-\; J^T v$$

where A is a real, symmetric, invertible $N \times N$ matrix, and v, J are N-element column vectors. Write

$$v \;=\; v_0 \;+\; v'$$

where $v_0 = A^{-1}J$. Then,

$$\begin{aligned}
(1/2)v^T Av \;-\; J^T v \;&=\; (1/2)(v_0 + v')^T A(v_0 + v') \;-\; J^T(v_0 + v') \\
&=\; (1/2)v_0^T Av_0 \;+\; (1/2)(v')^T Av' \;+\; (v')^T Av_0 \;-\; J^T v_0 \;-\; J^T v' \\
&=\; (1/2)(v')^T Av' \;-\; (1/2)J^T A^{-1}J
\end{aligned}$$

After completing the square, v' and J are decoupled from one another.

In the present case, we have the expression

$$\int d^4x \left[\frac{1}{2\hbar}\phi_\Lambda\left(-\partial_\mu\partial^\mu \;+\; m^2\right)\phi_\Lambda \;-\; J\phi_\Lambda\right]$$

and we would like to do something similar – complete the square so as to decouple J from a translate of ϕ_Λ. Formally, if we identify v with ϕ_Λ and A with $(1/\hbar)(-\partial_\mu\partial^\mu \;+\; m^2)$, then we have immediately that

$$\begin{aligned}
&\int d^4x \left[\frac{1}{2\hbar}\phi_\Lambda\left(-\partial_\mu\partial^\mu \;+\; m^2\right)\phi_\Lambda \;-\; J\phi_\Lambda\right] \\
&=\; \int d^4x d^4y \left[-(1/2)J(x)A^{-1}(x,y)J(y) \;+\; \frac{1}{2\hbar}\phi'_\Lambda(-\partial_\mu\partial^\mu \;+\; m^2)\phi'_\Lambda\right]
\end{aligned}$$

where

$$\phi_\Lambda(x) \;=\; \int d^4y A^{-1}(x,y)J(y) \;+\; \phi'_\Lambda(x)$$

and $A^{-1}(x,y)$ is some function such that, for example,

$$(1/\hbar)(-\partial^x_\mu \partial^\mu_x \; + \; m^2) A^{-1}(x,y) \; = \; \delta^4_\Lambda(x-y)$$

where $\delta^4_\Lambda(x-y)$ is a Λ-regulated version of the Dirac delta function.

To make this concrete, we must find an explicit expression for A^{-1}, but using Fourier analysis, this is relatively straightforward. The ordinary Dirac delta function has a Fourier transform expression given by[4]

$$\delta^4(x-y) \; = \; \frac{1}{(2\pi)^4} \int d^4 p \exp(ip \cdot (x-y))$$

The Λ-regulated version should be given by

$$\delta^4_\Lambda(x-y) \; = \; \frac{1}{(2\pi)^4} \int_{|p| \leq \Lambda} d^4 p \exp(ip \cdot (x-y))$$

Formally, it should now be more or less clear that an explicit expression for A^{-1} is given by

$$A^{-1}(x,y) \; = \; \frac{\hbar}{(2\pi)^4} \int_{|p| \leq \Lambda} d^4 p \frac{\exp(ip \cdot (x-y))}{p^2 \; + \; m^2}$$

In particular, note $A^{-1}(x,y) = A^{-1}(y,x) = A^{-1}(x-y)$.

Let us pause to point out a technical issue that will be important later. The expression above for A^{-1} is sensible in no small part because there are no zero eigenvalues of A, which in particular is because we are implicitly working with L^2 normalizable functions. For example, the laplacian $\partial_\mu \partial^\mu$ has no normalizable zero modes on Euclidean space. In general, this will not be the case. For example, when discussing nonlinear sigma models, as we shall do shortly, there *will* be normalizable zero modes. To handle them, we must split them off from the rest of the path integral, treat them with a separate (ordinary) integral, and then do a path integral over remaining non-zero modes. We will return to this point later.

Let us now return to the computation of correlation functions in equation (1). We can now rewrite that expression as

$$\frac{\int [D\phi_\Lambda] \frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \exp\left(+(1/2) \int d^4 x d^4 y J(x) A^{-1}(x,y) J(y) \; - \; S[\phi'_\Lambda]\right)\Big|_{J \equiv 0}}{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda])} \tag{2}$$

If we assume that the path integral possesses a basic translation invariance, so that $[D\phi_\Lambda] = [D\phi_\Lambda]$, and pull the J's out of the $[D\phi'_\Lambda]$ path integral, then the expression above reduces to

$$\frac{\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \exp\left(+(1/2) \int d^4 x d^4 y J(x) A^{-1}(x,y) J(y)\right)\Big|_{J \equiv 0} \int [D\phi_\Lambda] \exp\left(- \; S[\phi_\Lambda]\right)}{\int [D\phi_\Lambda] \exp(-S[\phi_\Lambda])}$$

[4]Remember, the Dirac delta function is a distribution, so its Fourier transform is bound to look odd.

but this reduces immediately to

$$\frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \exp\left(+(1/2) \int d^4x d^4y J(x) A^{-1}(x,y) J(y) \right) \Bigg|_{J \equiv 0} \tag{3}$$

in which the path integral no longer appears explicitly.

So, by adding a source and completing the square, we have removed the path integral from expressions for correlation functions.

It is easy to check that, for example,

$$< \phi_\Lambda(x) \phi_\Lambda(y) > = \ A^{-1}(x,y)$$

$$\begin{aligned} &< \phi_\Lambda(x_1) \phi_\Lambda(x_2) \phi_\Lambda(x_3) \phi_\Lambda(x_4) > \\ &\quad = \ A^{-1}(x_1,x_2) A^{-1}(x_3,x_4) \ + \ A^{-1}(x_1,x_3) A^{-1}(x_2,x_4) \ + \ A^{-1}(x_1,x_4) A^{-1}(x_2,x_3) \end{aligned}$$

and similarly $< \phi_\Lambda(x_1) \cdots \phi_\Lambda(x_n) >= 0$ when n is odd.

Rather than expand out all the derivatives in equation (3), there is a simple combinatorial trick to produce all the terms in these expressions, known as *Wick's theorem*. This says that the terms are obtained by summing up all possible 'contractions' of pairs in the correlation function. In other words, for every way of pairing up correlators, there is one term in the expansion, in which each such pair has been replaced by an A^{-1} factor. For example:

$$< \phi_\Lambda(x) \phi_\Lambda(y) > = < \phi_\Lambda(x) \ \phi_\Lambda(y) > = \ A^{-1}(x-y)$$

$$\begin{aligned} &< \phi_\Lambda(x_1) \phi_\Lambda(x_2) \phi_\Lambda(x_3) \phi_\Lambda(x_4) > \\ &= \ < \phi_\Lambda(x_1) \ \phi_\Lambda(x_2) \phi_\Lambda(x_3) \ \phi_\Lambda(x_4) > \ + \ < \phi_\Lambda(x_1) \phi_\Lambda(x_2) \ \phi_\Lambda(x_3) \ \ \phi_\Lambda(x_4) > \ + \ \cdots \\ &= \ A^{-1}(x_1,x_2) A^{-1}(x_3,x_4) \ + \ A^{-1}(x_1,x_4) A^{-1}(x_2,x_3) \ + \ A^{-1}(x_1,x_3) A^{-1}(x_2,x_4) \end{aligned}$$

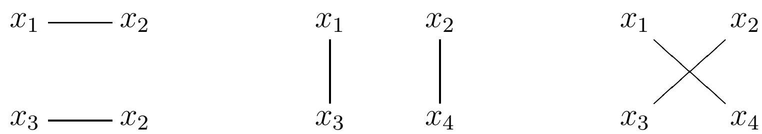
We can also associate Feynman diagrams. A Feynman diagram is just a graph in which the edges of the graph are in 1-1 correspondence with A^{-1}'s, which are known as "propagators." For example, the Feynman diagram for $< \phi_\Lambda(x) \phi_\Lambda(y) >$ is just a single line:

$$x \text{ ——— } y$$

The Feynman diagram for

$$< \phi_\Lambda(x_1) \phi_\Lambda(x_2) \phi_\Lambda(x_3) \phi_\Lambda(x_4) >$$

is a collection of three pairs of lines.



So far we have described a 'free' theory – one in which the action is purely quadratic (aside from the source term), so that the resulting Feynman diagrams expressing correlation functions graphically are nothing more than line segments connecting points. Next, we shall consider interactions.

So, suppose the action has the form

$$\frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\phi_\Lambda \left(-\partial_\mu\partial^\mu \;+\; m^2\right)\phi_\Lambda \;-\; V(\phi_\Lambda) \;-\; J\phi_\Lambda\right]$$

including the source J, for some real analytic function $V(\phi_\Lambda)$. For the rest of this lecture, we will work through the case $V(\phi_\Lambda) = \lambda\phi_\Lambda^4$.

Now, the action (for $J \equiv 0$) is no longer quadratic, so we cannot just complete the square as we did previously. Instead, utilizing the source term, the idea is to expand $V(\phi_\Lambda)$ in a Taylor series[5] and then convert each $\phi_\Lambda(x)^n$ term in the Taylor series to a $(\delta/\delta J(x))^n$. Thus, for example, in the case $V(\phi_\Lambda) = \lambda\phi_\Lambda^4$, the correlation function (1) becomes

$$\frac{\int[D\phi_\Lambda]\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_n)}\exp\left(+\frac{\lambda}{\hbar}\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\right)\exp(-S_0[\phi_\Lambda]+\int d^4x J(x)\phi_\Lambda(x)))\Big|_{J\equiv 0}}{\int[D\phi_\Lambda]\exp\left(+\frac{\lambda}{\hbar}\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\right)\exp\left(-S_0[\phi_\Lambda]+\int d^4x J(x)\phi_\Lambda(x)\right)\Big|_{J\equiv 0}} \tag{4}$$

where

$$S_0[\phi_\Lambda] \;=\; \frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\phi_\Lambda\left(-\partial_\mu\partial^\mu \;+\; m^2\right)\phi_\Lambda\right]$$

By turning the "potential" $V[\phi_\Lambda]$ into $\delta/\delta J$'s, we can again reduce to a quadratic action. Completing the square and proceeding as before, we find that the expression for the correlation function $< \phi_\Lambda(x_1)\cdots\phi_\Lambda(x_n) >$ in the theory with nonzero V (the "interacting" theory) is given by

$$\frac{\frac{\delta}{\delta J(x_1)}\cdots\frac{\delta}{\delta J(x_n)}\exp\left(+\frac{\lambda}{\hbar}\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\right)\exp\left(+(1/2)\int d^4x d^4y J(x)A^{-1}(x,y)J(y)\right)\Big|_{J\equiv 0}}{\exp\left(+\frac{\lambda}{\hbar}\int d^4x\left(\frac{\delta}{\delta J(x)}\right)^4\right)\exp\left(+(1/2)\int d^4x d^4y J(x)A^{-1}(x,y)J(y)\right)\Big|_{J\equiv 0}} \tag{5}$$

For example, it is now straightforward to compute that in this theory,

$$< \phi_\Lambda(x)\phi_\Lambda(y) > = \; A^{-1}(x-y) \;+\; (12)\frac{\lambda}{\hbar}\int d^2z A^{-1}(x-z)A^{-1}(y-z)A^{-1}(z-z) \;+\; \mathcal{O}(\lambda^2)$$

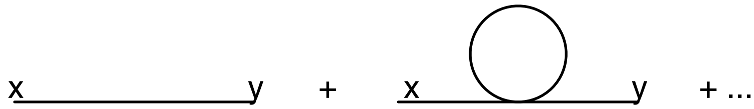


Figure 1: First few terms in two-point correlation function.

The corresponding Feynman diagrams are illustrated in figure 1.

The first graph, a straight line, shows the free propagator $A^{-1}(x-y)$. The second graph shows the effect of the $\lambda\phi^4$ interaction – a Feynman diagram with a 4-point vertex. If we had had a ϕ^3 interaction, for example, then the corresponding Feynman diagram would have had a 3-point vertex; a ϕ^n interaction leads to an n-point vertex.

If one only evaluated the numerator in equation (5), then one would also find contributions to $< \phi_\Lambda(x)\phi_\Lambda(y) >$ given by

$$3\frac{\lambda}{\hbar}A^{-1}(x-y)\int d^4z A^{-1}(z-z)A^{-1}(z-z)$$

which can be represented diagrammatially by the Feynman diagram shown below in figure 2

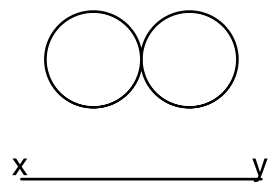


Figure 2: Vacuum diagram contribution to two-point correlation function.

The figure eight loop, in which particles propagate into one another but never to an external leg, is known as a *vacuum diagram*, and is interpreted physically as particle-antiparticle pairs spontaneously appearing out of the vacuum, interacting, and then annihilating one another back into nothingness.

However, although this term appears in the evaluation of the numerator of equation (5), it is cancelled out by a factor in the denominator. More generally, so long as one is careful to always include the denominators, there are never contributions to correlation functions from such vacuum diagrams.

[5] This is why we specified that V must be real analytic.

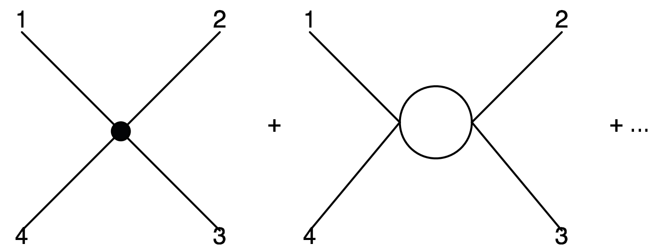


Figure 3: Some further Feynman diagrams appearing in the four-point correlation function.

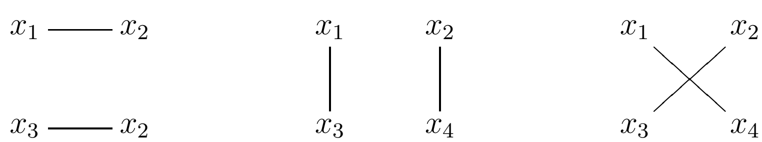
For another example, let us consider the four-point correlation function

$$< \phi_\Lambda(x_1)\phi_\Lambda(x_2)\phi_\Lambda(x_3)\phi_\Lambda(x_4) >$$

in this theory. One contribution to this correlation function will be the same as in the original theory without the ϕ^4 interaction, namely

$$A^{-1}(x_1, x_2)A^{-1}(x_3, x_4) \ + \ A^{-1}(x_1, x_3)A^{-1}(x_2, x_4) \ + \ A^{-1}(x_1, x_4)A^{-1}(x_2, x_3)$$

represented by the (disconnected) Feynman diagrams consisting of pairs of lines



However, there are now additional contributions. For example, there are also terms

$$\begin{aligned}
&\frac{\lambda}{\hbar}(4!) \int d^4 z A^{-1}(x_1 - z)A^{-1}(x_2 - z)A^{-1}(x_3 - z)A^{-1}(x_4 - z) \\
&\quad + \left(\frac{\lambda}{\hbar}\right)^2 (4!)^2 \int d^4 z_1 d^4 z_2 A^{-1}(x_1 - z_1)A^{-1}(x_4 - z_1)A^{-1}(x_2 - z_2)A^{-1}(x_3 - z_2)A^{-1}(z_1 - z_2)^2 \\
&\quad + \cdots
\end{aligned}$$

represented diagrammatically by the Feynman diagrams in figure 3.

The second term is one of a set of permutations – there is nothing special about the ordering of the labels on the outer legs, all possible orderings appear, but for simplicity we have only listed one of the terms and sketched one of the diagrams appearing.

WARNING: You might guess that the Feynman graph expansion above of correlation functions is a convergent series. In fact, it usually is not. Rather, it is usually an example of an *asymptotic* series[6]. In the conventions of the appendix, where asymptotic series of the form $\sum_n a_n/x^n$ are discussed, this is an asymptotic series in[7] λ^{-1}. As discussed in the appendix, these usually do not converge, but one can still get useful information from them. On the other hand, because they do not converge, it is possible for multiple functions to have the same asymptotic series – an asymptotic series expansion of a function loses a certain amount of information. That information is encoded in "nonperturbative effects," which are *not* uniquely determined by the perturbative expansion of the theory. This tension between perturbative and nonperturbative effects drives most basic discussions of quantum field theory – on the one hand, it makes getting exact answers very difficult, and on the other hand, makes nonperturbative contributions very interesting.

1.4 The renormalization group

The *renormalization group* or *renormalization group flow* is the name for the behavior of a quantum field theory as we lower the cutoff regulator Λ. (The term "renormalization group" itself is a historical artifact; there is no group present.) Given a theory defined at scale Λ, we can get a theory valid at scale $\Lambda - \delta\Lambda$ by performing the functional integral over degrees of freedom between Λ and $\Lambda - \delta\Lambda$, effectively "integrating out" those degrees of freedom. One finds that, even after taking into account the difference in cutoffs, the two theories will differ – there is a motion on an abstract space of quantum field theories. Also, this process of renormalization group flow loses information – it is possible for two theories that start different, to become the same after renormalization group flow. Equivalence classes identifying theories that become the same after renormalization group flow are known as *universality classes.*

To be specific, consider a theory of a single scalar ϕ_Λ with action

$$\frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\phi_\Lambda \left(-\partial_\mu \partial^\mu \, + \, m^2 \right) \phi_\Lambda \, - \, \lambda \phi_\Lambda^4 \right] \tag{6}$$

[6]The careful reader might wonder how this happened – this is arising from the Taylor series expansion of the exponential describing interactions, after all, so why isn't the final result also a Taylor series? Intuitively, the problem is that when $\lambda < 0$, the potential flips over and $\phi = 0$ is no longer a stable vacuum, so that the radius of convergence of any expansion in λ must be zero. A more thorough explanation involves carefully studying the numerical factors appearing in each Feynman diagram. We will not give such a thorough explanation here, but instead will state that analogous phenomena are relatively common. For example, one standard trick to obtain asymptotic series expansions is to integrate a Taylor series outside of its region of convergence – the resulting series is no longer a convergent series, but often can still be understood as an asymptotic series approximation.

[7]Expanding in λ gives us an expansion in the number of vertices appearing in Feynman diagrams. There is an alternative expansion parameter. It can be shown that powers of $\hbar$ count the number of loops appearing in the Feynman diagrams. Expanding in $\hbar$ also yields an asymptotic series expansion.

(Our analysis will closely follow [4][section 12.1].) Given a field $\phi_\Lambda(x)$, we can construct a new field $\phi_{\Lambda-\delta\Lambda}$ whose Fourier transform has the same components as ϕ_Λ for $|p| \leq \Lambda - \delta\Lambda$, but those higher-momentum components vanish. Define $\hat{\phi}$ to be a field whose Fourier transform has components that are nonzero only for $\Lambda - \delta\Lambda < |p| \leq \Lambda$, and zero outside that range, and inside that range, equal those of ϕ_Λ. Thus, we have

$$\phi_\Lambda(x) \;=\; \phi_{\Lambda-\delta\Lambda}(x) \;+\; \hat{\phi}(x)$$

What we want to do is to break the functional integral over ϕ_Λ into two pieces, one for $\phi_{\Lambda-\delta\Lambda}$ and one for $\hat{\phi}$, and then perform the functional integral over $\hat{\phi}$, which will give us an 'effective' quantum field theory for $\phi_{\Lambda-\delta\Lambda}$.

To that end, let us expand out the action (6) in the new variables. It is straightforward to check that the result is

$$\begin{aligned}
&\frac{1}{\hbar}\int d^4x \left[\frac{1}{2}(\phi_{\Lambda-\delta\Lambda} + \hat{\phi}) \left(-\partial_\mu \partial^\mu \;+\; m^2 \right) (\phi_{\Lambda-\delta\Lambda} + \hat{\phi}) \;-\; \lambda (\phi_{\Lambda-\delta\Lambda} + \hat{\phi})^4 \right] \\
&\quad = \;\frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\phi_{\Lambda-\delta\Lambda} \left(-\partial_\mu \partial^\mu \;+\; m^2 \right) \phi_{\Lambda-\delta\Lambda} \;-\; \lambda \phi_{\Lambda-\delta\Lambda}^4 \right] \\
&\qquad + \frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\hat{\phi} \left(-\partial_\mu \partial^\mu \;+\; m^2 \right) \hat{\phi} \;-\; \lambda \left(4(\phi_{\Lambda-\delta\Lambda})^3 \hat{\phi} \;+\; 6(\phi_{\Lambda-\delta\Lambda})^2 (\hat{\phi})^2 \;+\; 4\phi_{\Lambda-\delta\Lambda} (\hat{\phi})^3 \;+\; (\hat{\phi})^4 \right) \right]
\end{aligned}$$

where we have used the fact that

$$\int d^4x \phi_{\Lambda-\delta\Lambda} \left(-\partial_\mu \partial^\mu \;+\; m^2 \right) \hat{\phi} \;=\; \int d^4x \hat{\phi} \left(-\partial_\mu \partial^\mu \;+\; m^2 \right) \phi_{\Lambda-\delta\Lambda} \;=\; 0$$

because their Fourier transforms have nonoverlapping components.

Now, the next step is to – at least formally – integrate out the $\hat{\phi}$ degrees of freedom. As usual, we will not be able to do this exactly, but instead will merely be able to construct an asymptotic series expansion for the result. What we will find is that the mass m and coupling constant λ of the $\phi_{\Lambda-\delta\Lambda}$ theory will be shifted, among other things. Furthermore, they will change in a fashion that could not be compensated by rescaling the coordinates so as to effectively scale $\Lambda - \delta\Lambda$ back to Λ – changing the cutoff will change the theory, by more than just a motion in Λ space.

So, we need to construct a theory with only $\phi_{\Lambda-\delta\Lambda}$'s, that nevertheless reproduces correlation functions of the theory containing both $\phi_{\Lambda-\delta\Lambda}$'s as well as $\hat{\phi}$'s. To do this, let us compute some correlation functions in the theory with both $\phi_{\Lambda-\delta\Lambda}$'s and $\hat{\phi}$'s, and see what modifications will have to be made.

First, consider a two-point function $< \phi_{\Lambda-\delta\Lambda}(x_1)\phi_{\Lambda-\delta\Lambda}(x_2) >$. The first few Feynman diagram contributions to the correlation function, involving $\hat{\phi}$ propagators, are illustrated in figure 4.

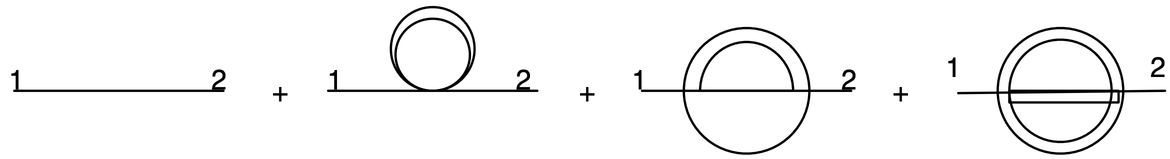


Figure 4: First few terms in two-point correlation function.

The first diagram in figure 4 just shows a single $\phi_{\Lambda-\delta\Lambda}$ propagator, propagating without interactions. In the second diagram, a single $\hat{\phi}^2\phi^2_{\Lambda-\delta\Lambda}$ interaction has been used. The third diagram involves two $\hat{\phi}\phi^3_{\Lambda-\delta\Lambda}$ interactions. The fourth diagram involves two $\hat{\phi}^3\phi_{\Lambda-\delta\Lambda}$ interactions.

Let us compute how the effect of one of these diagrams can be interpreted as a modification of the action for $\phi_{\Lambda-\delta\Lambda}$. Consider the second term in the figure above, involving the single $\hat{\phi}$ loop branching off of a $\phi_{\Lambda-\delta\Lambda}$ propagator. The $\phi_{\Lambda-\delta\Lambda}$ propagator is given by

$$A^{-1}(x-y) \;=\; \frac{\hbar}{(2\pi)^4}\int_{|p|\leq\Lambda-\delta\Lambda} d^4p \frac{\exp(ip\cdot(x-y))}{p^2\,+\,m^2}$$

and the $\hat{\phi}$ propagator is given by

$$B^{-1}(x-y) \;=\; \frac{\hbar}{(2\pi)^4}\int_{\Lambda-\delta\Lambda<|p|\leq\Lambda} d^4p \frac{\exp(ip\cdot(x-y))}{p^2\,+\,m^2}$$

Then the second term in the figure above is the term

$$\frac{\lambda}{\hbar}(2)\int d^4z A^{-1}(x_1-z)A^{-1}(x_2-z)B^{-1}(0)$$

in the asymptotic series expansion of the correlation function $<\phi_{\Lambda-\delta\Lambda}(x_1)\phi_{\Lambda-\delta\Lambda}(x_2)>$.

We are going to compute that contribution in detail, but before we do, let us first understand how to absorb this into a modification of the $\phi_{\Lambda-\delta\Lambda}$ action. The original $\phi_{\Lambda-\delta\Lambda}$ action, meaning, the terms not coupled to $\hat{\phi}$, are given by

$$\frac{1}{\hbar}\int d^4x \left[\frac{1}{2}\phi_{\Lambda-\delta\Lambda}\left(-\partial_\mu\partial^\mu\,+\,m^2\right)\phi_{\Lambda-\delta\Lambda}\,-\,\lambda\phi^4_{\Lambda-\delta\Lambda}\right]$$

If we were to shift m^2 by some amount, *i.e.* replace m^2 by $m^2+\delta m^2$, then the result would look like a new interaction term in the action, given by

$$\frac{1}{\hbar}\int d^4x \frac{1}{2}\phi^2_{\Lambda-\delta\Lambda}(\delta m^2)$$

and hence a contribution to the $<\phi_{\Lambda-\delta\Lambda}(x_1)\phi_{\Lambda-\delta\Lambda}(x_2)>$ correlation function given by

$$\frac{\delta m^2}{\hbar}\int d^4z A^{-1}(x_1-z)A^{-1}(x_2-z)$$

Thus, we can understand the effect of the $\hat{\phi}$ loop as a modification of the mass by an amount given by

$$\delta m^2 \; = 2\lambda B^{-1}(0)$$

In a moment we shall compute δm^2 in more detail, but first, let us understand the meaning of this contribution. The point here is that changing the value of the cutoff regulator Λ changes what we mean by m^2 – m^2 is not a true parameter but rather should be thought of as a function of Λ. As we integrate out the high-momentum degrees of freedom $\hat{\phi}$, we can understand their effects as a change in the "parameters" of the theory such as m^2.

This sort of effect is typical – parameters appearing in actions in quantum field theories are usually functions of cutoff scales.

Now, for completeness, let us derive an explicit expression for δm^2. Recall

$$\begin{aligned}
\delta m^2 &= 2\lambda B^{-1}(0) \\
&= \frac{2\lambda}{(2\pi)^4} \int_{\Lambda - \delta\Lambda < |p| \leq \Lambda} d^4 p \frac{1}{p^2 \, + \, m^2}
\end{aligned}$$

Since the integrand depends only upon the magnitude of p and not its direction, we can do the angular part separately; denote the integral over angular variables in Euclidean 4-space by Ω_4.

> In fact, an explicit expression for Ω_d, the integral of 1 over spherical angular variables in d dimensions, exists. It can be derived by rewriting
>
> $$\left(\int_{-\infty}^{\infty} dx \exp(-x^2) \right)^d \; = \; \pi^{d/2}$$
>
> as an integral over d-dimensional space. In general, the result is that
>
> $$\Omega_d \; = \; \frac{2\pi^{d/2}}{\Gamma(d/2)}$$
>
> where $\Gamma(x)$ is the "Gamma" function, a meromorphic function on the complex plane, obeying $\Gamma(x+1) = x\Gamma(x)$, $\Gamma(1/2) = \sqrt{\pi}$, and $\Gamma(n+1) = n!$ for n a positive integer. (See the appendix for more information on the gamma function.) We leave it as an exercise to verify that $\Omega_2 = 2\pi$ and $\Omega_3 = 4\pi$, as one would expect.

Using the fact that

$$\int dx \frac{x^3}{x^2 \, + \, a^2} \; = \; \frac{x^2}{2} \; - \; \frac{a^2}{2} \log\left(x^2 \, + \, a^2\right) \; + \; C$$

we find that

$$\delta m^2 \;=\; \frac{2\lambda}{(2\pi)^4}\Omega_4\left[\frac{1}{2}\left(\Lambda^2 \;-\; (\Lambda-\delta\Lambda)^2\right) \;-\; \frac{m^2}{2}\log\frac{\Lambda^2+m^2}{(\Lambda-\delta\Lambda)^2+m^2}\right]$$

Next, let us turn to the four-point function, and how the $\hat{\phi}$ contributions to correlation functions can be understood by modifying the $\phi_{\Lambda-\delta\Lambda}$ theory. A few interesting terms in the four-point function

$$< \phi_{\Lambda-\delta\Lambda}(x_1)\phi_{\Lambda-\delta\Lambda}(x_2)\phi_{\Lambda-\delta\Lambda}(x_3)\phi_{\Lambda-\delta\Lambda}(x_4) >$$

are illustrated in figure 5. A single line indicates a $\phi_{\Lambda-\delta\Lambda}$ propagator, whereas a double line indicates a $\hat{\phi}$ propagator.

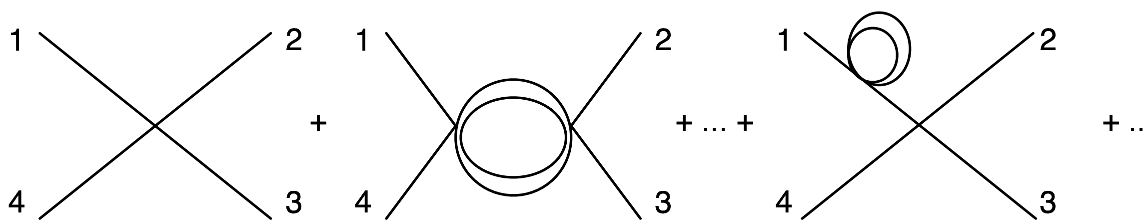


Figure 5: First few connected terms in four-point correlation function.

The first (left-most) diagram involves solely the $\lambda\phi^4_{\Lambda-\delta\Lambda}$ interaction. The second diagram is one of a set of three, involving two $\hat{\phi}^2\phi^2_{\Lambda-\delta\Lambda}$ interactions and a $\hat{\phi}$ running in an internal loop. There are three diagrams of that form, in which the labels on the four outer edges are permuted; we have only sketched one. The third diagram shown involves an effect we saw in the expansion of the two-point correlation function, in which a $\hat{\phi}$ loop runs off of one of the correlators. It involves a single $\hat{\phi}^2\phi^2_{\Lambda-\delta\lambda}$ interaction, as well as the $\lambda\phi^4_{\lambda-\delta\Lambda}$ interaction of the first diagram.

We have already implicitly taken into account the contribution of the third diagram, in our previous discussion of two-point correlation functions, where we saw that such a $\hat{\phi}$ loop branching off of a propagator had the effect of modifying the mass.

Here, let us examine the effect of the middle diagram and its permutations. We will argue that the effect of such diagrams is to shift the value of λ, just as the effect of the single $\hat{\phi}$ loop branching off of a propagator was to modify the mass.

The middle diagram itself corresponds to the following term in the four-point correlation function:

$$\frac{1}{2!}\left(\frac{\lambda}{\hbar}\right)^2(6)^2(2)^4\int d^4z_1\int d^4z_2 A^{-1}(x_1-z_1)A^{-1}(x_4-z_1)B^{-1}(z_1-z_2)^2A^{-1}(x_2-z_2)A^{-1}(x_3-z_2)$$

After some algebra, this can be rewritten as

$$\begin{aligned}
&\frac{1}{2!}\left(\frac{\lambda}{\hbar}\right)^2 (6)^2(2)^4\lambda^2\left(\frac{\hbar}{(2\pi)^4}\right)^4 \int_{|p|\leq\Lambda-\delta\Lambda} d^4p_1\cdots d^4p_4 \exp(ip_1x_1+ip_2x_4+ip_3x_2+ip_4x_3)\\
&\quad\cdot\delta^4(p_1+p_2+p_3+p_4)(p_1^2+m^2)^{-1}\cdots(p_4^2+m^2)^{-1}\\
&\quad\cdot\int_{\Lambda-\delta\Lambda<|p|\leq\Lambda} d^4p_5(p_5^2+m^2)((p_1+p_2-p_5)^2+m^2)^{-1}\Theta(p_1+p_2-p_5)
\end{aligned}$$

where

$$\Theta(p) \;=\; \begin{cases} 1 & \Lambda-\delta\Lambda<|p|\leq\Lambda \\ 0 & \text{else}\end{cases}$$

We can get a decent approximation to this integral by noting that the last factor will only receive contributions when $|p_5| \approx \Lambda$ and $|p_1+p_2-p_5| \approx \Lambda$, so that

$$\begin{aligned}
&\int_{\Lambda-\delta\Lambda<|p|\leq\Lambda} d^4p_5(p_5^2+m^2)((p_1+p_2-p_5)^2+m^2)^{-1}\Theta(p_1+p_2-p_5)\\
&\approx\;\int_{\Lambda-\delta\Lambda<|p|\leq\Lambda} d^4p_5(p_5^2+m^2)^{-2}\\
&=\;\Omega_4\int_{\Lambda-\delta\Lambda}^{\Lambda} r^3dr(r^2+m^2)^{-2}\\
&=\;\Omega_4\left[\frac{m^2}{2(\Lambda^2+m^2)} \;-\; \frac{m^2}{2((\Lambda-\delta\Lambda)^2+m^2)} \;+\; \frac{1}{2}\log\left(\frac{\Lambda^2+m^2}{(\Lambda-\delta\Lambda)^2+m^2}\right)\right]\\
&\approx\;\Omega_4\frac{\delta\Lambda}{\Lambda}
\end{aligned}$$

Now, by way of comparison, the diagram with no loops contributes

$$\begin{aligned}
&\left(\frac{\lambda}{\hbar}\right)(4!)\int d^4z A^{-1}(x_1-z)A^{-1}(x_2-z)A^{-1}(x_3-z)A^{-1}(x_4-z)\\
&\;=\;\left(\frac{\lambda}{\hbar}\right)(4!)\left(\frac{\hbar}{(2\pi)^4}\right)^4\int_{|p|\leq\Lambda-\delta\Lambda} d^4p_1\cdots d^4p_4\exp(ip_1x_1+\cdots+ip_4x_4)\\
&\quad\;\cdot(p_1^2+m^2)^{-1}\cdots(p_4^2+m^2)^{-1}\delta^4(p_1+p_2+p_3+p_4)
\end{aligned}$$

Comparing our results for the first and middle diagrams, we see that, to leading order[8], the effect of the internal $\hat{\phi}$ loop is to shift the value of λ to $\lambda+\delta\lambda$ where

$$\delta\lambda \;=\; 3\lambda^2\hbar\frac{1}{2!}\frac{1}{4!}(6)^2(2)^4\Omega_4\frac{\delta\Lambda}{\Lambda}$$

[8] To subleading order, another interesting bit of physics comes into play. Specifically, to subleading order, $\delta\lambda$ depends upon the inflowing momenta p_1 and p_2. This is not a mistake, but rather is another general feature of quantum field theory, one we shall not pursue here.

where the leading factor of 3 comes from the fact that there are a total of three diagrams of this form, relating by permuting the external legs.

So far, it appears naively that lowering the cutoff Λ has effectively changed the theory, by shifting the values[9] of m^2 and λ. The careful reader might object that the theory was defined by a triple m^2, λ, and Λ, and so it is not fair to compare theories with different cutoffs. Let us fix this issue by performing a rescaling that will restore the cutoff to its original value Λ. If we define

$$x' \;=\; \frac{\Lambda - \delta\Lambda}{\Lambda} x$$

then this will shift the dual momenta to

$$p' \;=\; \frac{\Lambda}{\Lambda - \delta\Lambda} p$$

and in particular after this rescaling, the new cutoff changes to

$$(\Lambda - \delta\Lambda) \frac{\Lambda}{\Lambda - \delta\Lambda} \;=\; \Lambda$$

When we perform that scaling in the action, we find that

$$m^2 \,+\, \delta m^2 \;\mapsto\; \left(\frac{\Lambda}{\Lambda - \delta\Lambda} \right)^4 \left(m^2 \,+\, \delta m^2 \right)$$

and similarly for $\lambda + \delta\lambda$. It is straightforward to check, order-by-order in $\delta\Lambda$, that these rescalings do not restore the original values of m^2 or λ. As a result, it really is the case that lowering the cutoff is changing the theory – rescaling to restore the original value of the cutoff, does not restore the original values of the parameters.

So far we have seen how lowering the cutoff Λ has effectively changed the theory – it effectively shifts m^2 and λ. It is natural in this context to define a function which is a derivative along this process. We define the *beta function* for λ to be

$$\beta \;\equiv\; \lim_{\delta\Lambda \rightarrow 0} \Lambda \frac{\partial}{\partial \delta\Lambda} \left(\lambda + \delta\lambda \right)$$

In the present case, to leading order, we compute that

$$\beta \;\approx\; 3\lambda^2 \hbar \frac{1}{2!} \frac{1}{4!} (6)^2 (2)^4 \Omega_4$$

(Note that although the beta function is defined invariantly, following the procedure outlined here the best we can compute is successive terms in an asymptotic series approximation to the beta function.)

[9] An even more attentive reader will note that in principle the scale of ϕ could also be changed – another possible quantum effect is to rescale the ϕ's, so that the action is written in terms of $Z\phi$ for some Z rather than just ϕ. This does happen in general, and plays an important role; but for simplicity and brevity, we are omitting this important effect.

This effect can also be described with a differential equation for the correlation functions, known as the renormalization group equation or Callan-Symanzik equation. This has the form

$$\left(-t\frac{\partial}{\partial t} \, + \, \beta\frac{\partial}{\partial \lambda} + \sum_{i=1}^{n} d_i \right) < \mathcal{O}_1 \cdots \mathcal{O}_n > = \, 0$$

in a theory in which $m = 0$. The variable t is a momentum scaling parameter, and the d_i's are the scaling dimensions[10] of the $\mathcal{O}_i$'s. In the special case that $\beta = 0$, this just says that the correlation function scales in the naive way; the beta function β acts as a sort of anomaly in scale invariance.

To summarize, as we lower the cutoff Λ, the quantum field theory changes – not just by a redefinition of Λ, but the 'parameters' of the theory are revealed to secretly be functions of Λ. This process of lowering the cutoff, integrating out high-momentum degrees of freedom, is a lossy, non-invertible process. In particular, it is possible for two quantum field theories that start different at some cutoff Λ, to become the same at a smaller cutoff after applying renormalization group flow. This trick is used in physical realizations of both derived categories and stacks; in the former case, this is how localization on quasi-isomorphisms is realized physically.

1.5 Fermions

So far we have described quantum field theory for scalar-valued functions. Next, we shall outline quantum field theory for fermions, in the special case of two dimensions[11].

In two dimensions, a fermion is a pair of Grassmann-valued sections $(\psi_+(x), \psi_-(x))$ of a bundle on the two-dimensional spacetime, $\sqrt{K}$ (K is the canonical bundle) in typical cases. To be Grassmann-valued means that multiplication anticommutes:

$$\psi(x)\psi(y) \; = \; -\psi(y)\psi(x)$$

Intuitively, to be Grassmann-valued means to behave as if multiplication were by the wedge product of differential geometry, though the technical details are a bit more complicated.

For Grassmann constants, there are notions of both derivatives and integrals. The derivative is defined in what should be a fairly obvious fashion. For example, if *theta* is a Grassmann

[10] The scaling dimensions need not be purely classical – they can receive quantum corrections. Briefly, the reason is that there is an additional possible parameter we have ignored. In addition to modifying m^2 and λ, we can also modify the scale of ϕ_Λ, by defining $\phi'_\Lambda = Z\phi_\Lambda$. Just as m^2 and λ are secretly functions of Λ, so too is Z. The freedom to rescale the ϕ_Λ's by a function ultimately allows the naive scaling dimension to be modified, via Z.

[11] Fermions also make sense in higher dimensions, but there are technicalities in their definition which are irrelevant for our purposes. We restrict to two dimensions for brevity, not because of any fundamental physical obstruction.

constant,

$$
\begin{aligned}
\frac{\partial}{\partial \theta}\theta &= 1 \\
\frac{\partial}{\partial \theta}(a \,+\, b\theta) &= b \text{ where } a, b \in \mathbf{C} \\
\frac{\partial}{\partial \theta_1}\theta_2\theta_1 &= -\theta_2 \frac{\partial}{\partial \theta_1}\theta_1 \;=\; -\theta_2
\end{aligned}
$$

and so forth. Note in particular that the derivative anticommutes with other Grassmann variables.

To properly define the integral requires more space than we have here. (One really ought to introduce supermanifolds, distinguish differential forms on the supermanifold from closely related but distinct integral forms, and then introduce the Berezinian to define the integral – see for example [33].) However, for our purposes here, and indeed for the purposes of most discussions of quantum field theory, a simplified approach suffices. We can define the Grassmann integral to be the same as the derivative:

$$
\int d\theta \;\equiv\; \frac{\partial}{\partial \theta}
$$

Thus, for example,

$$
\begin{aligned}
\int d\theta\, (a + b\theta) &= b \\
\int d\theta_1 d\theta_2\, (\theta_1\theta_2) &= \int d\theta_1 d\theta_2\, (-\theta_2\theta_1) \;=\; \int d\theta_1\, (-\theta_1) \;=\; -1 \\
&= -\int d\theta_2 d\theta_1\, (\theta_1\theta_2) \;=\; -\int d\theta_2 \theta_2 \;=\; -1
\end{aligned}
$$

For Grassmann-valued functions, and Grassmann-valued sections of bundles, there are corresponding notions of Grassmann functional derivatives and functional integrals, with the same anticommutivity properties. For example,

$$
\frac{\delta}{\delta\psi(x)}\frac{\delta}{\delta\psi(y)} \;=\; -\frac{\delta}{\delta\psi(y)}\frac{\delta}{\delta\psi(x)}
$$

A typical action for a two-dimensional fermion, not coupled to any additional bundles, is of the form

$$
\int d^2x \left[i\overline{\psi}_- \partial \psi_- \;+\; i\overline{\psi}_+ \overline{\partial} \psi_+ \right]
$$

where $\overline{\psi}$ denotes the complex conjugate. In principle, one can compute propagators, Feyman diagrams, *etc* more or less just as in the case of scalar field theory.

With an eye towards future developments, let us also take a moment to discuss zero modes. Since the propagator is the inverse of the operator appearing in the kinetic term, one must first factor out any zero eigenvalues of that operator. For the case of a scalar field on a Euclidean spacetime, there were no (normalizable) zero modes, hence there was nothing to factor out, but in principle one could expect a space $\mathcal{M}$ of bosonic zero modes, in which case the path integral measure for the scalar field would contain a factor

$$\int_{\mathcal{M}}$$

integrating over those zero modes separately. Similarly, if a worldsheet fermion ψ has zero modes, those must be factored out and handled separately. If we have coupled ψ_+ to some nontrivial bundle $\mathcal{E}$, say, with hermitian fiber metric $h_{\alpha\overline{\beta}}$, so that

$$\psi_+ \sim \Gamma_{C^\infty}\left(\sqrt{K} \otimes \mathcal{E}\right)$$

then the fermion kinetic term is

$$\int d^2 x i h_{\alpha\overline{\beta}} \overline{\psi}_+^{\overline{\beta}} D_{\overline{z}} \psi_+^\alpha$$

where $D_{\overline{z}}$ is an antiholomorphic covariant derivative. If $\psi^{i_1}, \cdots, \psi^{i_n}$ are a set of zero modes, then the path integral measure contains a Grassmann-integral factor

$$\int d\psi^{i_1} \cdots d\psi^{i_n} \omega_{i_1 \cdots i_n}$$

where ω is a nowhere-zero holomorphic section of $\Lambda^n \mathcal{E}$.

1.6 Exercises

1. Fermat's principle of optics states that light travels in a path for which the quantity $\int n(x,y,z) ds$ is a minimum, where ds is the infinitesimal arc length and n is the "index of refraction." Restrict to paths in a plane for simplicity, and (in polar coordinates) suppose that $n(r,\theta) = r^k$ for some integer k. Show that when $k = -1$, a light ray can travel in a circle about the origin.

2. Consider the action for a field ϕ with action

$$S = \int d^3 x \left(\partial_\mu \phi \partial^\mu \phi - g\phi^3 - \lambda \phi^4 + \gamma \phi^6 \right)$$

   Use Hamilton's least-action principle to find the equations of motion of ϕ.

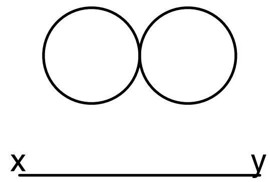
3. Using the identity

$$\left(\int_{-\infty}^{\infty} dx \exp\left(-x^2\right) \right)^d = \pi^{d/2}$$

   show that the integral of 1 over spherical angular variables in d dimensions is

$$\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$$

4. In evaluating the two-point correlation function $< \phi_\Lambda(x)\phi_\Lambda(y) >$, we encountered a vacuum diagram, represented graphically by the Feynman diagram



but argued that this contribution in the numerator of equation (5) was cancelled out by a contribution from the denominator of that expression. Check that claim – show that a contribution from the denominator cancels out this term in the numerator. (Hint: Use the Taylor expansion

$$(1 + x)^{-1} = 1 - x + x^2 - x^3 + \cdots$$

when expanding out the denominator in equation (5).) More generally, show that all such vacuum diagrams always cancel out, and so never contribute to correlation functions.

5. We extensively discussed a theory of a single real-valued function $\phi(x)$ with a $\lambda\phi^4$ interaction. Now, change the $\lambda\phi^4$ interaction to a $g\phi^3$ interaction. What are the first few terms in the asymptotic series expansion for the correlation function

$$< \phi_\Lambda(x_1)\phi_\Lambda(x_2)\phi_\Lambda(x_3) >$$

? (Write out the terms both in terms of A^{-1}'s as well as Feynman diagrams.) (Hint: a ϕ^n interaction should generate n-point vertices in Feynman diagrams.) Argue here that vacuum diagram contributions again cancel out. If you are feeling ambitious, also estimate the leading order term in the beta function for the coupling g.

6. Suppose we replace the $\lambda\phi^4$ by $g\exp(a\phi)$ for constant a, g. Show that the sum of all vacuum diagrams through first order in g is given by

$$\frac{g}{\hbar}\int d^4x \exp\left(\frac{a^2}{2}A^{-1}(x,x)\right)$$

Hint: you can either expand out the terms and find the series expansion, or, you can use an infinite-dimensional analogue of the identity

$$\exp\left(a\frac{d}{dx}\right) f(x) = f(x+a)$$

namely,

$$\exp\left(a\frac{\delta}{\delta J(x)}\right) f[J(y)] = f\left[J(y) + a\delta(y-x)\right]$$

Asymptotic series problems

7. For $x > 1$, consider the series
$$\sum_{n=0}^{\infty} \frac{(-)^n}{x^{n+1}}$$
Determine if this series is asymptotic to the function
$$\frac{1}{1 + x}$$

8. Show that a series $\sum \frac{A_n}{x^n}$ is asymptotic to $f(x)$ if and only if it is also asymptotic to $f(x)+\exp(-x)$. (In QFT, this is the statement that Feynman diagrams do not capture all the information of correlation functions, they miss nonperturbative effects, which appear as terms proportional to $\exp(-kx)$ for some k.) In particular, an asymptotic series expansion of a function does not uniquely determine the function.

9. The modified Bessel function $K_0(x)$ has the integral representation
$$K_0(x) \;=\; \int_0^\infty \exp(-x\cosh t)dt$$
Use the method of steepest descent, applied to the integral above, to derive the leading term in an asymptotic expansion of $K_0(x)$, valid for large x.

2 Nonlinear sigma models

2.1 Physical nonlinear sigma models

2.1.1 Definition

A *nonlinear sigma model* is a two-dimensional quantum field theory describing maps from the two-dimensional spacetime or *worldsheet* Σ into some Riemannian manifold X (with a complex structure and hermitian metric) known as the *target space*. We will also assume the hermitian metric is Kähler – we will not need that fact to define the nonlinear sigma model, but, we will need it to insure that the model has "supersymmetry," which we shall discuss later.

The action for a nonlinear sigma model is given by
$$\frac{1}{\alpha'}\int_\Sigma d^2z\left(\frac{1}{2}g_{\mu\nu}\partial\phi^\mu\overline{\partial}\phi^\nu \;+\; \frac{i}{2}B_{\mu\nu}\partial\phi^\mu\overline{\partial}\phi^\nu \;+\; ig_{i\overline{\jmath}}\psi_-^{\overline{\jmath}}D_z\psi_-^i \;+\; ig_{i\overline{\jmath}}\psi_+^{\overline{\jmath}}D_{\overline{z}}\psi_+^i \;+\; R_{i\overline{\jmath}k\overline{l}}\psi_+^i\psi_+^{\overline{\jmath}}\psi_-^k\psi_-^{\overline{l}}\right)$$
The α' is a constant, nomenclature determined by convention, that (as we shall see later) is related to the coupling constant in the theory. The ϕ fields are maps from the worldsheet

Σ into the target space X, expressed in local holomorphic coordinates $i, \overline{\jmath}$ or Euclidean coordinates μ, ν. Both the ϕ and ψ fields are implicitly assumed to be regulated; but for convenience of notation, we have omitted the Λ subscripts, and will assume them implicitly henceforth. The metric $g_{\mu\nu}$ is a metric on the target space X, pulled back to the worldsheet. As such, it is a function of the ϕ fields. $R_{i\overline{\jmath}k\overline{l}}$ is the pullback of the Riemann curvature (4-)tensor on X, determined by the metric $g_{\mu\nu}$, and expressed in components. $B_{\mu\nu}(\phi)$ is an antisymmetric closed 2-form on the target space, pulled back to the worldsheet.

The $\psi_{\pm}^{i,\overline{\imath}}$ are Grassmann-valued sections of the following bundles:

$$\begin{array}{ll} \psi_{+}^{i} \in \Gamma_{C^\infty}\left(K_\Sigma^{1/2} \otimes \phi^* T^{1,0}X\right) & \psi_{-}^{i} \in \Gamma_{C^\infty}\left(\overline{K}_\Sigma^{1/2} \otimes (\phi^* T^{0,1}X)^\vee\right) \\ \psi_{+}^{\overline{\imath}} \in \Gamma_{C^\infty}\left(K_\Sigma^{1/2} \otimes (\phi^* T^{1,0}X)^\vee\right) & \psi_{-}^{\overline{\imath}} \in \Gamma_{C^\infty}\left(\overline{K}_\Sigma^{1/2} \otimes \phi^* T^{0,1}X\right) \end{array}$$

The line bundles K_Σ and $\overline{K}_\Sigma$ are the holomorphic and antiholomorphic canonical bundles on Σ. (Technically, to define the nonlinear sigma model, we must pick a particular spin structure – a set of choices of square-root bundles $K^{1/2}$ – and different choices give slightly different nonlinear sigma models.)

Finally, the D's in $D_z\psi_{-}^{i}$ and $D_{\overline{z}}\psi_{+}^{i}$ are 'covariant derivatives.' In general, if ψ is a smooth Grassmann-valued section of some bundle $\mathcal{E}$, then the covariant derivative $D\psi$ is defined so as to be a smooth Grassmann-valued section of $T\Sigma \otimes \mathcal{E}$. An ordinary derivative $\partial\psi$ will not have this property – across coordinate patches, the derivative will act on the transition functions and so general inhomogeneous terms. To fix that, we add terms to the ordinary derivatives which also transform inhomogeneously across coordinate patches, balanced so that the inhomogeneities all cancel out. In the present case,

$$D_z\psi_{-}^{i} \;=\; \partial\psi_{-}^{i} \;+\; \partial\phi^j \Gamma^i_{jm}\psi_{-}^{m} \;+\; \frac{1}{2}\omega_z \psi_{-}^{i}$$

The Γ^i_{jm} is a connection on a principle bundle associated to TX, and is defined in terms of the metric $G_{\mu\nu}$ on X by

$$\Gamma^{\mu}_{\nu\lambda} \;=\; \frac{1}{2} g^{\mu\rho}\left(\partial_\lambda g_{\rho\nu} \;+\; \partial_\nu g_{\rho\lambda} \;-\; \partial_\rho g_{\nu\lambda}\right)$$

It is known as the 'Christoffel connection,' and is well-known in differential geometry. On a Kähler manifold, defined by a hermitian metric with the property that $\partial_i g_{j\overline{k}} = \partial_j g_{i\overline{k}}$, the Christoffel connection has the property that only components with all holomorphic or all antiholomorphic indices are nonzero (exercise: check!). The Christoffel connection term is present because ψ_{-}^{i} couples to $\phi^* TX$. Finally, ω_z is a connection on a principal bundle associated to K_Σ, and it is present because ψ_{-}^{i} couples to $\sqrt{K_\Sigma}$.

For those acquainted with Gromov-Witten theory, we should point out that we are *not* summing over worldsheet metrics / complex structures in this quantum field theory. Instead, we are working implicitly with a fixed choice of metric on Σ, which has mostly been suppressed in writing the action above. It is also possible to consider a different theory in which one sums over choices of metrics on Σ. The resulting quantum field theory is much more complicated to describe than the quantum field theory above, and leads after twisting to what is known as a topological *string* theory, whereas the theory above will lead to topological *field* theories. Thus, in the theory above, invariants built from certain topological twists will be closely analogous to, but not quite the same as, Gromov-Witten invariants.

Let us take a moment to rewrite the purely bosonic part of the action

$$
\begin{aligned}
& \frac{1}{2} g_{\mu\nu} \partial\phi^\mu \overline{\partial}\phi^\nu \,+\, \frac{i}{2} B_{\mu\nu} \partial\phi^\mu \overline{\partial}\phi^\nu \\
&= \frac{1}{2} g_{i\overline{\jmath}} \left(\partial\phi^i \overline{\partial}\phi^{\overline{\jmath}} \,+\, \overline{\partial}\phi^i \partial\phi^{\overline{\jmath}} \right) \,+\, \frac{i}{2} B_{i\overline{\jmath}} \left(\partial\phi^i \overline{\partial}\phi^{\overline{\jmath}} \,-\, \overline{\partial}\phi^i \partial\phi^{\overline{\jmath}} \right) \\
&= g_{i\overline{\jmath}} \overline{\partial}\phi^i \partial\phi^{\overline{\jmath}} \,+\, \frac{1}{2} \left(g_{i\overline{\jmath}} \,+\, i B_{i\overline{\jmath}} \right) \left(\partial\phi^i \overline{\partial}\phi^{\overline{\jmath}} \,-\, \overline{\partial}\phi^i \partial\phi^{\overline{\jmath}} \right) \\
&= g_{i\overline{\jmath}} \overline{\partial}\phi^i \partial\phi^{\overline{\jmath}} \,+\, \frac{1}{2} \phi^* \left(\omega \,+\, iB \right)
\end{aligned}
$$

where ω is the Kähler form on the target space. The first term amounts to $|\overline{\partial}\phi^i|^2$, and tells us that holomorphic maps will be zero modes, for which the kinetic term vanishes. The second term is a topological term, determined by the topological class of the map ϕ.

2.1.2 Perturbation theory

How do we make sense of this as a quantum field theory? Previously, we've discussed models in which the kinetic term for the scalars was of the form $\partial_\mu \phi \partial^\mu \phi$, but here, we have a term of that form with a multiplicative factor containing a function of ϕ, *i.e.*, here the kinetic term is more nearly of the form $f(\phi)\partial_\mu \phi \partial^\mu \phi$.

In order to do perturbative quantum field theory, we must first split off the zero modes of the Laplacian. For example, constant maps into X are examples of zero modes, so we must split them off in the path integral and then do perturbation theory around each fixed zero mode. More generally, holomorphic maps $\Sigma \to X$ are further examples of zero modes that should be split off. (We can see this from the rewritten form of the bosonic terms above – the first term amounts to $|\overline{\partial}\phi^i|^2$, and will vanish identically for holomorphic maps; the second term is determined by the topological class of ϕ, and does not contribute to the dynamics.) In other words, we can write

$$
\int [D\phi] \,=\, \sum_d \int_{\mathcal{M}_d} \int [D\phi]'
$$

where d is the degree of the map, $\mathcal{M}_d$ is a moduli space of such maps of fixed degree, and $\int [D\phi]'$ denotes a path integral over nonzero eigenfunctions of the Laplacian. In the case of degree 0 maps, these are constant maps $\Sigma \to X$, so $\mathcal{M}_0 = X$. To do perturbative quantum field theory, for any one fixed constant map ϕ_0, we expand[12] $\phi = \phi_0 + \phi'$, and then replace quantities such as $g_{\mu\nu}(\phi)$ with a Taylor-series expansion

$$g_{\mu\nu}(\phi_0) \; + \; \cdots$$

Furthermore, with clever choices of local coordinates[13], we can arrange for $g_{\mu\nu}(\phi_0)$ to be $\delta_{\mu\nu}$, and also simultaneously set the first derivatives to zero. Thus, expanding around a fixed zero mode ϕ_0, we can write

$$g_{i\bar{\jmath}}\overline{\partial}\phi^i \partial \phi^{\bar{\jmath}} \; = \; \delta_{i\bar{\jmath}}\overline{\partial}\phi^i \partial \phi^{\bar{\jmath}} \; + \; \text{(infinite series of interaction terms)}$$

This begins to make it possible to do perturbative quantum field theory, in the sense described earlier in these notes. In particular, for each point on the moduli space $\coprod_d \mathcal{M}_d$, we get a series of Feynman diagrams, determined by the infinite tower of interactions that arose from the Taylor series for the metric $g_{i\bar{\jmath}}$. Each Feynman diagram is defined at a specific point of the bosonic moduli space $\coprod_d \mathcal{M}_d$; we then integrate the sum of Feynman diagrams over the moduli space to get the correlation function.

There is a closely analogous story for the fermions. To make sense of propagators there, again we must first split off the zero modes, which will be either holomorphic (for ψ_+, since the kinetic term is proportional to $D_{\overline{z}}\psi_+$) or antiholomorphic sections (for ψ_-, since the kinetic term is proportional to $D_z\psi_-$) of the bundles. Instead of an ordinary integral, we will have a Grassmann integral.

Putting this together, we see that formally correlation functions can be expressed as

$$< \mathcal{O}_1 \cdots \mathcal{O}_n > = \sum_d \int_{\mathcal{M}_d} \int d\psi_0 \cdots d\psi_0 \frac{\int [D\phi]' \int [D\psi]' \exp(-S/\hbar) \mathcal{O}_1 \cdots \mathcal{O}_n}{\int [D\phi]' \int [D\psi]' \exp(-S/\hbar)}$$

When defining correlation functions previously in, say, $\lambda\phi^4$ theory, we did not have any $\int_{\mathcal{M}}$ over moduli spaces of bosonic zero modes because there were no bosonic zero modes. The expression above is a more general expression for correlation functions than that considered previously, that takes into account the possibility of zero modes.

If Σ is $\mathbf{R}^2$, then, it is possible to compute propagators and Feynman diagrams in exactly the same way described earlier. If Σ is a compact Riemann surface, then, constructing propagators becomes much more difficult, though large-momentum features such as renormalization will behave in the same fashion as on $\mathbf{R}^2$.

Let us also take a moment to discuss coupling constants in the nonlinear sigma model. When the metric is scaled up to be very large, so that all features of the space are a nearly

[12] How can we 'add' two maps into a space X? Remember, we are working in local coordinates everywhere, so this addition is actually taking place in the image of the coordinate charts, which is a vector space. Of course, we also need to assume that subset of the vector space has certain obvious connectivity properties for the addition to be well-defined.

[13] Such as Riemann normal coordinates [1].

infinite distance away from one another, the nonlinear sigma model should be weakly coupled, intuitively, whereas when the metric is scaled down, all features of the space are close to one another and should be easily visible, so intuitively the nonlinear sigma model should be strongly coupled. That intuition is correct, and we can make it precise as follows. Fix one particular metric $g^0_{\mu\nu}$, and then for any other metric $g_{\mu\nu}$ related by an overall scale, write $g_{\mu\nu} = r^2 g^0_{\mu\nu}$ for some constant r. Then the bosonic kinetic term in the action has the form

$$\frac{1}{\alpha'} \int d^2x \left[\frac{1}{2} g_{\mu\nu} \partial \phi^\mu \overline{\partial} \phi^\nu \right] \; = \; \frac{r^2}{\alpha'} \int d^2x \left[\frac{1}{2} g^0_{\mu\nu} \partial \phi^\mu \overline{\partial} \phi^\nu \right]$$

Expand the map ϕ^μ about a constant map ϕ_0:

$$\phi^\mu \; = \; \phi^\mu_0 \; + \; \delta \phi^\mu$$

then the action above can be written

$$\frac{r^2}{\alpha'} \int d^2x \left[\frac{1}{2} g^0_{\mu\nu}(\phi_0) \partial(\delta\phi^\mu) \overline{\partial}(\delta\phi^\nu) \; + \; \frac{1}{2} (\partial_\rho g^0_{\mu\nu})(\phi_0) (\delta\phi^\rho) \partial(\delta\phi^\mu) \overline{\partial}(\delta\phi^\nu) \; + \; \cdots \right]$$

Finally, we need to rescale the fields so that the basic kinetic term, the first term in the Taylor expansion, which will define propagators, has coefficient 1. Define

$$\tilde{\phi}^\mu \; = \; \frac{r}{\sqrt{\alpha'}} \delta \phi^\mu$$

Then the action can be rewritten in the final form

$$\int d^2x \left[\frac{1}{2} g^0_{\mu\nu}(\phi_0) \partial \tilde{\phi}^\mu \overline{\partial} \tilde{\phi}^\nu \; + \; \frac{\sqrt{\alpha'}}{r} \frac{1}{2} (\partial_\rho g_{\mu\nu})(\phi_0) \tilde{\phi}^\rho \partial \tilde{\phi}^\mu \overline{\partial} \tilde{\phi}^\nu \; + \; \cdots \right]$$

We see that all the interaction terms, arising from the Taylor expansion of the metric, are multiplied by positive powers of $\sqrt{\alpha'}/r$. Thus, when r is large, the interaction terms are suppressed, and the nonlinear sigma model is nearly a free field theory – weakly coupled, as our intuition suggested, and similarly, when r is small, interactions are magnified and the theory is strongly coupled.

2.1.3 Scale invariance

The nonlinear sigma model has a classical scale invariance. If we rescale the coordinates on Σ as $z \mapsto \lambda z$, then the action is invariant if we rescale the fermions as $\psi \mapsto \lambda^{-1/2} \psi$.

Quantum-mechanically, that scale invariance can be spoiled by quantum corrections. Recall from our earlier discussion of beta functions that the beta function acts as a sort of anomaly in scale-invariance: if the beta function is nonzero, then the correlation functions will not scale in the fashion one would expect.

In the present case, the interactions determined by the metric have a 2-tensor beta function $\beta_{\mu\nu}$, which to leading order is given by the Ricci tensor of the Riemannian metric on X:

$$\beta_{\mu\nu} \propto R_{\mu\nu}$$

(In the current theory, with fermions, it can be shown that perturbatively the only further contribution to the beta function arises at four-loop order.) Vanishing of the beta function is a necessary condition for a conformal field theory. This is part of the reason why Calabi-Yau manifolds are of interest to physicists: after all, a necessary condition for a complex Riemannian manifold X to be Calabi-Yau is that $R_{\mu\nu} = 0$. (It can be shown [2, 3] that on a Calabi-Yau, one can make $\beta_{\mu\nu} = 0$ to all orders, not just leading order, by making small adjustments to the Calabi-Yau metric.)

2.1.4 Supersymmetry

This nonlinear sigma model also possesses another property, known as *supersymmetry*. Supersymmetry is a symmetry that exchanges bosons ϕ with fermions ψ. Because of the complex structures on both Σ and X, and the compatible metrics, there are a total of four different supersymmetry transformations possible, which are naturally grouped into two pairs and called "(2,2) supersymmetry." The (infinitesimal) supersymmetry transformations are given by

$$\begin{aligned}
\delta\phi^i &= i\alpha_- \psi_+^i + i\alpha_+ \psi_-^i \\
\delta\phi^{\bar{\imath}} &= i\tilde{\alpha}_- \psi_+^{\bar{\imath}} + i\tilde{\alpha}_+ \psi_-^{\bar{\imath}} \\
\delta\psi_+^i &= -\tilde{\alpha}_- \partial\phi^i - i\alpha_+ \psi_-^j \Gamma^i_{jm} \psi_+^m \\
\delta\psi_+^{\bar{\imath}} &= -\alpha_- \partial\phi^{\bar{\imath}} - i\tilde{\alpha}_+ \psi_-^{\bar{\jmath}} \Gamma^{\bar{\imath}}_{\bar{\jmath}\bar{m}} \psi_+^{\bar{m}} \\
\delta\psi_-^i &= -\tilde{\alpha}_+ \overline{\partial}\phi^i - i\alpha_- \psi_+^j \Gamma^i_{jm} \psi_-^m \\
\delta\psi_-^{\bar{\imath}} &= -\alpha_+ \overline{\partial}\phi^{\bar{\imath}} - i\tilde{\alpha}_- \psi_+^{\bar{\jmath}} \Gamma^{\bar{\imath}}_{\bar{\jmath}\bar{m}} \psi_-^{\bar{m}}
\end{aligned}$$

The "(2,2)" refers to the fact that there are four supersymmetry transformation parameters, $\alpha_\pm$, $\tilde{\alpha}_\pm$, one pair of each chirality. These supersymmetry transformation parameters are Grassmann-valued C^∞ sections of $K^{-1/2}$, $\overline{K}^{-1/2}$.

2.2 The A, B model topological field theories

There are several different ways to obtain topological field theories. One of the most commonly described is to 'twist' an ordinary supersymmetric quantum field theory, by changing which bundles various fields couple to. In the case of the nonlinear sigma model described in the last section, given the fermions $\psi_\pm^{i,\bar{\imath}}$, we can replace $K_\Sigma^{1/2}$ with either 1 or K_Σ. There is a constraint – we must require that each term in the action remain a $(1,1)$ form, so if in ψ_+^i we replace $K_\Sigma^{1/2}$ with 1, then correspondingly in $\psi_+^{\overline{\imath}}$ we must replace $K_\Sigma^{1/2}$ with K_Σ.

As a result, there are four different choices of 'twist' one can make. Two of those choices are physically equivalent to the other two, leaving only two interesting possibilities. Those two possibilities have been labelled the A and B models.

A very readable reference on the A and B model on Riemann surfaces without boundary (the 'closed string' A, B models) is [6]. This section will closely follow that reference.

For the most part, we will be interested in the B model, since that is the model to which derived categories are relevant. However, we shall also outline A model computations, as they provide some insight into why the B model works the way it does.

2.2.1 The closed string A model

In this section we will outline the A model on a worldsheet Σ without boundary – the 'closed string' A model.

As outlined above, the A model is nearly identical to the nonlinear sigma model – the action functional has the same form, for example. The difference is that we interpret the fermions ψ as coupling to slightly different bundles – which means the covariant derivatives $D\psi$ are defined slightly differently. In this model, the fermions couple to bundles as follows:

$$
\begin{array}{cc}
\psi_+^i (\equiv \chi^i) \,\in\, \Gamma_{C^\infty}\left(\phi^* T^{1,0}X\right) & \psi_-^i (\equiv \psi_{\overline{z}}^i) \,\in\, \Gamma_{C^\infty}\left(\overline{K}_\Sigma \otimes (\phi^* T^{0,1}X)^\vee\right) \\
\psi_+^{\overline{\imath}} (\equiv \psi_z^{\overline{\imath}}) \,\in\, \Gamma_{C^\infty}\left(K_\Sigma \otimes (\phi^* T^{1,0}X)^\vee\right) & \psi_-^{\overline{\imath}} (\equiv \chi^{\overline{\imath}}) \,\in\, \Gamma_{C^\infty}\left(\phi^* T^{0,1}X\right)
\end{array}
$$

We have also introduced new notation for the fermions, following the conventions of [6].

The original supersymmetry transformation parameters coupled to $K_\Sigma^{-1/2}$, $\overline{K}_\Sigma^{-1/2}$, but after this twist, half of them now couple to 1, *i.e.* are Grassmann-valued scalars. Furthermore, it can be shown that that subset is nilpotent. In physics, any nilpotent symmetry with Grassmann-valued scalar parameters is called a *BRST symmetry*, in honor of Becci-Rouet-Stora-Taylor who first introduced such a symmetry into quantum field theory.

In the present case, the BRST transformation parameters are α_- and $\tilde{\alpha}_+$, which we shall rename α and $\tilde{\alpha}$ for the rest of this section. The BRST transformations of the fields (a subset of the supersymmetry transformations) are then given as follows:

$$
\begin{aligned}
\delta \phi^i &= i\alpha\chi^i \\
\delta \phi^{\overline{\imath}} &= i\tilde{\alpha}\chi^{\overline{\imath}} \\
\delta \chi^i &= 0
\end{aligned}
$$

$$
\begin{aligned}
\delta \chi^{\overline{\imath}} &= 0 \\
\delta \psi_z^{\overline{\imath}} &= -\alpha \partial \phi^{\overline{\imath}} - i\tilde{\alpha} \chi^{\overline{\jmath}} \Gamma^{\overline{\imath}}_{\overline{\jmath}\overline{m}} \psi_z^{\overline{m}} \\
\delta \psi_{\overline{z}}^{i} &= -\tilde{\alpha} \overline{\partial} \phi^{i} - i\alpha \chi^{j} \Gamma^{i}_{jm} \psi_{\overline{z}}^{m}
\end{aligned}
$$

BRST transformations with $\alpha = \tilde{\alpha}$ are sometimes denoted with a Grassmann-valued operator Q, defined by $\delta - = \alpha Q -$, so for example, $Q\phi^i = i\chi^i$.

Now, as we have described it so far, this quantum field theory, the A model, is nearly identical to the original nonlinear sigma model. The difference between the two is somewhat subtle – just a change in the definitions of the fermions. That modification changes the covariant derivatives in the action, but otherwise the form of the action functional is unchanged.

However, in the A model, there is a special subsector of the theory, known as the *topological subsector*, in which we can make very strong statements. The topological subsector consists of correlation functions of BRST-invariant combinations of fields. Clearly anything built from products of χ's will be BRST-invariant, but we can also multiply by certain functions of ϕ in special circumstances. Let[14]

$$
b(\phi)_{i_1 \cdots i_p \overline{\imath}_1 \cdots \overline{\imath}_q} \chi^{i_1} \cdots \chi^{i_p} \chi^{\overline{\imath}_1} \cdots \chi^{\overline{\imath}_q}
$$

denote a very general product of fields that should include all (worldsheet scalar) BRST invariant field products as a subset.

In fact, more can be done: it can be shown there is a one-to-one correspondence between BRST-closed field products and d-closed differential forms on X. In this isomorphism, we identify χ^i with dz^i, $\chi^{\overline{\imath}}$ with $d\overline{z}^{\overline{\imath}}$, and modulo factors of i, we identify Q with the exterior derivative d. The Grassmann property of the χ's maps to the antisymmetric property of the wedge product.

> A little more generally, on any given component of the moduli space[15] of bosonic zero modes, the BRST-closed field products should be identified with differential forms on that component. In general, the moduli space of bosonic

[14] An attentive reader might note the following problem with the expression above. In correlation functions computed previously, when two fields approached one another, the correlation function typically diverged as some function of the cutoff Λ. Therefore, placing fields on top of one another, with zero separation, should surely not be well-defined. This observation is correct, and there is a fix. The fix is to define a *composite operator*, in which such divergences have been removed. Now, in general, defining composite operators is complicated, but QFT's in two dimensions are a special case, and all such divergences can be removed through a very mild process of 'normal ordering.' We shall not define normal ordering here, but will simply state that all products of fields, evaluated at the same point on Σ, are implicitly normal ordered in these lectures.

[15] In fact, the classical moduli spaces are not compact. Physically this leads to "IR divergences" which have to be regulated, and so one is led to pick a compactification of the moduli spaces. There is nowadays a well-established theory of how one does this, which we shall not attempt to cover here.

zero modes will be a disjoint union

$$\coprod_d \mathcal{M}_d$$

and each BRST-invariant field product $\mathcal{O}_i$ will define a differential form of fixed degree on each $\mathcal{M}_d$. For $d = 0$, $\mathcal{M}_d = X$, so in particular we get a differential form on X. However, there is a more or less standard way to derive the differential forms on other components from the differential form on X, so to describe the forms on all components, it suffices to describe the differential form on X.

Moreover, physically it can be shown that any correlation function of a Q-exact product of fields vanishes:

$$< Q\{\cdots\} > = \ 0$$

(We will not explain the reason for this here, but will simply state it as a fact. It is relatively straightforward to derive, but involves a few ideas we do not have the space here to describe.)

As a result, in the topological subsector, we are only interested in Q-closed field products modulo Q-exact ones, because if one of the correlators is Q-exact, then since the others are Q-closed, the product of all correlators can be written as a Q-exact expression, and the correlation function must vanish:

$$< \mathcal{O}_1 \cdots \mathcal{O}_n (QV) > = < Q\{\mathcal{O}_1 \cdots \mathcal{O}_n V\} > = \ 0$$

where $Q\mathcal{O}_i = 0$ by assumption.

Now, let us outline how to calculate a correlation function of Q-closed field products $\mathcal{O}_i$. In principle, from our general remarks on nonlinear sigma models, the correlation function should have the form

$$< \mathcal{O}_1 \cdots \mathcal{O}_n > = \ \sum_d \int_{\mathcal{M}_d} < \mathcal{O}_1 \cdots \mathcal{O}_n >_{d, x \in \mathcal{M}_d}$$

It can be shown that, within the topological subsector of this particular theory, there are (usually) no net Feynman diagram contributions. Roughly, diagrams with bosons ϕ running in loops cancel out against diagrams with fermions ψ running in loops, a typical property of supersymmetric theories. More precisely, correlation functions in the topological subsector are independent of α', and since, as we showed earlier, the coupling constant in the theory is α'/r^2 for some scale r, α' independence usually means there are no Feynman diagram corrections. The α'-independence of correlation functions in the topological subsector follows from the fact that the action for the nonlinear sigma model can be written in the form

$$\frac{1}{\alpha'} \int d^2x \, (Q \cdot V) \ + \ \frac{1}{\alpha'} \int d^2x \phi^* \, (\omega \ + \ iB)$$

where

$$V \;=\; g_{i\bar{\jmath}} \left(\psi_z^{\bar{\jmath}} \overline{\partial} \phi^i \;+\; \partial \phi^{\bar{\jmath}} \psi_{\bar{z}}^i \right)$$

As a result, if we formally take a derivative of a correlation function with respect to α', the result is to insert into the action a Q-exact expression plus a topological term. The term with the Q-exact expression vanishes, and the other term is topological. Thus, correlation functions are independent of α', except insofar as α' also multiplies the topological term in the action. (A very similar argument will also hold in the B-twisted nonlinear sigma model, which is also independent of α'.)

Thus, correlation functions can be computed in free field theory, and boil down to working with zero modes.

If we identify each correlator $\mathcal{O}_i$ with some differential form ω_i, then each correlator

$$< \mathcal{O}_1 \cdots \mathcal{O}_n >_{d, x \in \mathcal{M}_d}$$

can be identified with a wedge product of differential forms on $\mathcal{M}_d$, evaluated at a point $x \in \mathcal{M}_d$.

Now, let us finish evaluating these correlation functions. In addition to the bosonic zero modes, encapsulated in the moduli space components $\mathcal{M}_d$, there are also fermionic zero modes, and corresponding Grassmann integral contributions to the path integral. In the case in which there are no $\psi_z^{\bar{\imath}}$ or $\psi_{\bar{z}}^i$ zero modes, only χ^i, $\chi^{\bar{\imath}}$ zero modes, then it can be shown that the number of such zero modes, for maps of any fixed degree d is the same as the dimension of the moduli space component $\mathcal{M}_d$, and so after performing those Grassmann integrals we have

$$\int_{\mathcal{M}_d} < \mathcal{O}_1 \cdots \mathcal{O}_n >_{d, x \in \mathcal{M}_d} = \int_{\mathcal{M}_d} \omega_1 \wedge \cdots \wedge \omega_n \;=\; \int_{\mathcal{M}_d} (\text{top-form})$$

where the ω_i are differential forms associated to the $\mathcal{O}_i$. In the more general case, in order to absorb the Grassmann integrals arising from $\psi_z^{\bar{\imath}}$ or $\psi_{\bar{z}}^i$ zero modes, we must utilize the interaction[16] term

$$\frac{1}{\alpha'} \int_\Sigma d^2 x \left(R_{i\bar{\jmath}k\bar{l}} \psi_+^i \psi_+^{\bar{\jmath}} \psi_-^k \psi_-^{\bar{l}} \right)$$

[16]The careful reader will wonder why this is allowed, since we have previously argued that correlation functions should be α'-independent, save for the topological term $\int \phi^* (\omega + iB)$. The correct answer is slightly subtle – a careful analysis of the $\psi_z^{\bar{\imath}}$, $\psi_{\bar{z}}^i$ zero modes, more careful than we have space for here, shows that their Grassmann integrals in the path integral measure are weighted by factors of $\sqrt{\alpha'}$. (A quick way to see why this is reasonable is to compare dimensions – these fields are worldsheet vectors, not scalars, so the Grassmann integrals must be multiplied by something with dimensions of length to make everything consistent, and $\sqrt{\alpha'}$ is a natural choice. This is only an intuitive justification, not a solid argument.) In any event, the factors of α' in the path integral measure cancel out the factors of α' arising from the interaction, so the final result is that correlation functions are, indeed, independent of α', as originally claimed.

We bring down one copy of this interaction for each pair of $\psi_z^{\bar{\imath}}$, $\psi_{\bar{z}}^{i}$ zero modes, and the result can be interpreted mathematically as modifying the correlation function computation by inserting the top Chern class of the obstruction bundle [5]:

$$\int_{\mathcal{M}_d} < \mathcal{O}_1 \cdots \mathcal{O}_n >_{d, x \in \mathcal{M}_d} = \int_{\mathcal{M}_d} \omega_1 \wedge \cdots \wedge \omega_n \wedge c_{top} \, (\mathrm{Obs})$$

As a sample computation, let us outline correlation functions in the A model on $\mathbf{P}^n$. To simplify matters, we shall simply state as a fact that the compactified moduli spaces are given by

$$\mathcal{M}_d \; = \; \mathbf{P}^{(n+1)(d+1)-1}$$

Let $\mathcal{O}$ be a BRST-closed field product corresponding to a $(1,1)$ form on $\mathbf{P}^n$ that generates its classical cohomology ring. In this particular case, there are no $\psi_{\bar{z}}^{i}$ or $\psi_z^{\bar{\imath}}$ zero modes, only χ zero modes. A given correlation function $< \mathcal{O}^k >$ will be nonzero in a sector d if and only if it defines a top-form, *i.e.* the dimension of $\mathcal{M}_d$ is k. Define q to be a constant such that in a sector of maps of degree d,

$$q^d \; = \; \exp(\int \phi^* (\omega + iB))$$

Then we find that

$$< \mathcal{O}^k > = \left\{ \begin{array}{ll} q^d & k = (n+1)(d+1) - 1 \\ 0 & \mathrm{else} \end{array} \right.$$

Put another way,

$$\begin{array}{rcl} < \mathcal{O}^n > & = & 1 \\ < \mathcal{O}^n \mathcal{O}^{n+1} > & = & q \\ < \mathcal{O}^n \mathcal{O}^{d(n+1)} > & = & q^d \end{array}$$

Notice that we can describe these correlation functions more compactly: given that $< \mathcal{O}^n >= 1$, we can recover the other correlation functions formally by identifying $\mathcal{O}^{n+1} \sim q$. Such a relationship is known in quantum field theory as an *operator product*, and one way of describing the correlation functions above is in terms of the *operator product ring*. This ring consists of products of $\mathcal{O}$ with itself modulo the relation above, *i.e.* $\mathbf{C}[\mathcal{O}]/(\mathcal{O}^{n+1} - q)$. That ring is a deformation of the classical cohomology ring of $\mathbf{P}^n$, namely $\mathbf{C}[x]/(x^{n+1})$. For that reason, we refer to the new ring as a *quantum cohomology ring*. We shall not discuss quantum cohomology rings further, but thought it might be illuminating to outline how they arise physically.

Although we have formulated the A model as a topological twist of a nonlinear sigma model with a fixed complex and Kähler structure, in fact the A model can be shown to be independent of the choice of complex structure – it only depends upon the choice of Kähler structure. Briefly, this is because the action can be written in the form $Q \cdot V$ with all information about the complex structure buried in V – so deforming the complex structure will merely add Q-exact terms, which will leave correlation functions in the topological subsector invariant. We will see that for the B model, the opposite is true – the B model is independent of Kähler structure, but depends upon the choice of complex structure.

2.2.2 The closed string B model

In this section we will outline the B model on a worldsheet Σ without boundary – the 'closed string' B model. Adding a boundary complicates matters, but we will definitely discuss this later, because it is the 'open string' B model or, B model on a Riemann surface with boundary, that is the physical realization of derived categories.

In the closed string B model, we perform the opposite topological twist to the A model. Here, we take the fermions ψ to couple to the following bundles:

$$\begin{array}{ll} \psi_+^i \in \Gamma_{C^\infty}\left(K \otimes \phi^* T^{1,0}X\right) & \psi_-^i \in \Gamma_{C^\infty}\left(\overline{K} \otimes (\phi^* T^{0,1}X)^\vee\right) \\ \psi_+^{\overline{\imath}} \in \Gamma_{C^\infty}\left((\phi^* T^{1,0}X)^\vee\right) & \psi_-^{\overline{\imath}} \in \Gamma_{C^\infty}\left(\phi^* T^{0,1}X\right) \end{array}$$

It is convenient to define[17]

$$\begin{aligned} \eta^{\overline{\imath}} &= \psi_+^{\overline{\imath}} + \psi_-^{\overline{\imath}} \\ \theta_i &= g_{i\overline{\imath}}\left(\psi_+^{\overline{\imath}} - \psi_-^{\overline{\imath}}\right) \\ \rho_z^i &= \psi_+^i \\ \rho_{\overline{z}}^i &= \psi_-^i \end{aligned}$$

In these variables, the action can be written

$$\frac{1}{\alpha'}\int_\Sigma d^2x\left[g_{\mu\nu}\partial\phi^\mu\overline{\partial}\phi^\nu + i g_{i\overline{\jmath}}\eta^{\overline{\jmath}}\left(D_z\rho_{\overline{z}}^i + D_{\overline{z}}\rho_z^i\right) + i\theta_i\left(D_{\overline{z}}\rho_z^i - D_z\rho_{\overline{z}}^i\right) + R_{i\overline{\imath}j\overline{\jmath}}\rho_z^i\rho_{\overline{z}}^j\eta^{\overline{\imath}}\theta_k g^{k\overline{\jmath}}\right]$$

The supersymmetry transformation parameters that are scalars after the twist are $\tilde{\alpha}_+$ and $\tilde{\alpha}_-$. For simplicity, assume $\tilde{\alpha}_+ = \tilde{\alpha}_- = \alpha$. The corresponding BRST transformations (a subset of the supersymmetry transformations) are given by

$$\begin{aligned} \delta\phi^i &= 0 \\ \delta\phi^{\overline{\imath}} &= i\alpha\eta^{\overline{\imath}} \end{aligned}$$

[17] A careful reader will object that $\psi_+^{\overline{\imath}}$ and $\psi_-^{\overline{\imath}}$ couple to different bundles, so as we have defined them, it does not make sense to simply add them in the definitions of $\eta^{\overline{\imath}}$ or θ_i. However, there is a trivial fix. We can use the target-space metric to dualize; and also, to be consistent, we should have defined one of the fields with lower indices instead of upper indices, and used the metric to relate them. Thus, for example, a more nearly technically accurate definition of $\eta^{\overline{\imath}}$ is as $\psi_+^{\overline{\imath}} + g^{\overline{\imath}j}\psi_{-j}$, where $\psi_+^{\overline{\imath}}$ and ψ_{-j} are defined to be sections of the same bundles. Because the metric is always present, it is a standard convention to omit it from discussions, and leave the reader to fill it in (as we have done here) as needed.

$$\begin{aligned}
\delta \eta^{\bar{\imath}} &= 0 \\
\delta \theta_i &= 0 \\
\delta \rho^i &= -\alpha d\phi^i
\end{aligned}$$

Here, the correlators in the topological subsector are of the form

$$b(\phi)^{j_1 \cdots j_q}_{\bar{\imath}_1 \cdots \bar{\imath}_p} \eta^{\bar{\imath}_1} \cdots \eta^{\bar{\imath}_p} \theta_{j_1} \cdots \theta_{j_q}$$

and as before, we are only interested in Q closed products modulo Q exact products, for the same reasons. Here, the BRST operator Q can be identified with the operator $\overline{\partial}$, $\eta^{\bar{\imath}}$ can be identified with $d\overline{z}^{\bar{\imath}}$, and θ_i can be identified with $\partial/\partial z^i$. The cohomology of the BRST operator can then be identified with the sheaf chomology group

$$H^p\left(X, \Lambda^q T^{1,0} X\right)$$

There is a consistency condition that one needs to make sense of the B model, arising from the fermion zero modes. Recall from our discussion of basic QFT, that for each set of fermion zero modes, there must exist a nowhere-zero section of the bundle to which the zero modes couple. In this case, if we restrict to just ψ_+, say, in degree $d = 0$, then there are as many ψ_+ zero modes as the dimension of X, and we need a nowhere-zero section of $\Lambda^{top} T^{1,0} X$. In fact, to be compatible with the BRST operator, that nowhere-zero section must be holomorphic (else the BRST symmetry is broken across coordinate patches). A holomorphic nowhere-zero section of a line bundle is a trivialization, so this implies that X must be Calabi-Yau. For this reason, one usually says that the B model is only well-defined on Calabi-Yau's.

A couple of comments are in order at this juncture.

- First, the discussion above ignored the ψ_- zero modes. When one takes them into account, the Calabi-Yau condition can be very slightly weakened: one only needs a trivialization of $K_X^{\otimes 2}$, not K_X. Unfortunately, that does not buy you very much, because if a complex Kähler manifold X has the property that $K_X^{\otimes 2} \cong \mathcal{O}_X$ but $K_X \not\cong \mathcal{O}_X$, then X is not simply-connected, and it has a double cover which is Calabi-Yau.
- There is an analogous story in the A model, but there, the effects of the ψ_+ and ψ_- cancel out. One requires a trivialization of $K_X \otimes K_X^{-1} \cong \mathcal{O}_X$ instead of $K_X^{\otimes 2}$, but a trivialization of $\mathcal{O}_X$ always exists, so there is no constraint in the A model. Furthermore, that trivialization can be chosen to be constant, and so guaranteed to commute with the A model's BRST operator, which is d rather than $\overline{\partial}$.

For a more detailed discussion of these subtleties, see for example [7].

It can be shown, in much the same way as for the A model, that the B model is independent of Kähler structure on X, just as the A model was independent of complex structure, because Kähler deformations add Q-exact terms to the action. See [6] for details.

Furthermore, the B model is independent of α', for the same reasons as the A model, and so correlation functions reduce to semi-classical zero-mode manipulations.

From the same general analysis as the A model, one would then expect that correlation functions should have the general form

$$< \mathcal{O}_1 \cdots \mathcal{O}_n > = \sum_d \int_{\mathcal{M}_d} \omega_1 \wedge \cdots \wedge \omega_n$$

where $\omega_i \in H^p(\mathcal{M}_d, \Lambda^q T\mathcal{M}_d)$, and the integrals on the right side might occasionally have $\wedge c_\top$ factors or something analogous. However, the Kähler-moduli-independence implies that this result should simplify considerably. Because we can make the metric as large as we want, we should be able to suppress instanton effects to an arbitrary large degree. In other words, there should only be a contribution from $d = 0$ sectors, and nothing else:

$$< \mathcal{O}_1 \cdots \mathcal{O}_n > = \int_X \omega_1 \wedge \cdots \wedge \omega_n$$

Put another way, the open string B model is purely classical – it does not get any worldsheet instanton corrections.

2.3 More abstract discussions of topological field theories

Shortly after topological field theories were introduced to the physics literature, as special quantum field theories in which there existed a subset of correlation functions (defining the the *topological subsector*) which were independent of the metric on the space on which the theory is defined, there arose in the mathematics literature certain mathematical translations of properties of the topological subsector.

The first such example, to the author's knowledge, was work of Atiyah [23] in which a topological field theory was defined as a functor from the cobordism category to the category of Hilbert spaces. For example, for the A, B models above, in this language, we think of associating BRST-cohomology of field products to one-dimensional boundaries of Riemann surfaces, then think of the Riemann surface with boundary as describing how the system evolves. Since the theory is independent of worldsheet metric, only the topology of the Riemann surface with boundary is pertinent.

Now, the detailed structure of the topological subsector of the A, B models can be somewhat richer than that description implies, and so there have been further refinements. The most recent version of these refinements, to our knowledge, is due to Costello [24, 25], who has a description of both open and closed string topological conformal field theories, which includes most (if not all) of the algebraic structures present in the general case, including L_∞ and A_∞ algebras.

Thus, at this point in time, there seems to be a detailed purely mathematical understanding of the topological subsector of topological field theories, and it is possible to map from the topological subsector of a given topological field theory to one of these more categorical descriptions.

However, the reverse construction is not known. Given, say, a functor from the cobordism category to the category of Hilbert spaces, ala Atiyah, or even a later refinement, it is not yet known in general how to build a quantum field theory containing a topological subsector whose properties are described by the categorical framework.

Thus, at this point in time, there is a difference between what physicists mean by a topological field theory (a quantum field theory containing a topological subsector) and what many mathematicians mean by a topological field theory (as exemplified by Atiyah's functor on the cobordism category). The topological subsector of a quantum field theory nearly always represents only a miniscule fraction of the correlation functions in the full quantum field theory – the quantum field theory contains a great deal of information not described in the topological subsector.

2.4 Exercises

1. In the discussion of the nonlinear sigma model, supersymmetry transformations of the fields were listed. Check that the action is invariant under those transformations, in the special case of a constant metric. (The more general case requires much more work but is not especially more illuminating.)

2. In the discussion of the closed string A model, the BRST transformations of the fields were listed. Check that the BRST transformations are a subset of the susy transformations of the nonlinear sigma model, and that they are nilpotent.

3. In the case of the closed string A model, show that in a $d = 0$ sector (*i.e.* constant maps), that there are no $\psi_z^{\bar{\imath}}$ or $\psi_{\bar{z}}^{i}$ zero modes, and that there are as many χ zero modes as the dimension of X, in the case that the worldsheet $\Sigma = \mathbf{P}^1$.

4. In the case of the closed string B model, we have argued in general terms that the quantum field theory does not get any worldsheet instanton corrections. Let us examine a particular case more closely. Suppose our target space X is the total space of the bundle $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \rightarrow \mathbf{P}^1$, and consider a worldsheet instanton sector describing maps into the base $\mathbf{P}^1$, of degree d, so that $\phi^* T^{1,0} X = \mathcal{O}(2d) \oplus \mathcal{O}(-d) \oplus \mathcal{O}(-d)$. In

this case, the moduli space $\mathcal{M}_d$ has dimension $2d+1$. Show that when the worldsheet $\Sigma = \mathbf{P}^1$, the number of η and θ zero modes is $2(d+1)$, for $d > 0$, which is wrong to match the dimension of the moduli space, hence it is impossible for physics to generate a top-form on the moduli space. (For the A model, on the other hand, the analogous computation gives $2d+1$ zero modes (check!), exactly right to match the dimension of the moduli space.) This argument unfortunately does not work for more general normal bundles; for other cases, the reason why the B model gets no quantum corrections from those sectors is a more subtle issue. See [7] for details.

3 The open string B model and derived categories

3.1 Boundary conditions and Chan-Paton factors

There are two important differences between the open string B model and the closed string B model, *i.e.*, the B model on Riemann surfaces with and without boundary, respectively. The first difference is that we must specify boundary conditions on the open strings. Those boundary conditions will tie the ends of the open string to submanifolds of the space, so that, at low energy, the open string only propagates along the submanifold. The second difference is that there are additional degrees of freedom located solely along the boundary. These extra degrees of freedom, known as *Chan-Paton factors*, are specified by a smooth principal bundle with connection living over the distinguished submanifold.

The choice of submanifold, together with a bundle on it, is to a first approximation what is meant by a *D-brane*. The submanifold is sometimes known as the *worldvolume* of the D-brane, and the bundle over that submanifold is sometimes known as the *gauge bundle*.

In the case of the open string B model, both the submanifold and the bundle are constrained: the submanifold must be a complex submanifold, and the bundle must be equivalent[18] to a holomorphic vector bundle.

3.1.1 Boundary conditions

Let us first discuss boundary conditions, without adding any extra degrees of freedom on the boundary. After we add those extra degrees of freedom, we shall see the boundary conditions change, but first let us work through boundary conditions in the simpler case.

We need to impose some sort of boundary conditions on the open string fields in order to make the theory well-defined on a Riemann surface with boundary. The two basic boundary

[18]Modulo a choice of hermitian fiber metric.

conditions that we work with are Neumann boundary conditions and Dirichlet boundary conditions. Neumann boundary conditions are given by

$$\partial \phi^\mu \;=\; \overline{\partial} \phi^\mu$$

and Dirichlet boundary conditions are given by

$$\partial \phi^\mu \;=\; -\overline{\partial} \phi^\mu$$

These two boundary conditions may look indistinguishable, but there is an important physical distinction between them. First, let us express these conditions in real coordinates rather than complex ones. Write $z = x + iy$, and work on the upper half-plane. Then, $\partial + \overline{\partial} = \partial_x$ and $\partial - \overline{\partial} = -i\partial_y$. Thus, we see that Neumann boundary conditions imply

$$\partial_y \phi^\mu = 0$$

or equivalently $\partial_n \phi^\mu \;=\; 0$ along the boundary (n for normal), whereas Dirichlet boundary conditions imply

$$\partial_x \phi^\mu \;=\; 0$$

or equivalently $\partial_t \phi^\mu = 0$ along the boundary (t for tangent). These boundary conditions still sound very similar. To understand precisely how they differ physically, let us temporarily switch from Euclidean metrics and indices to Lorentzian ones. (Lorentzian metrics are more relevant than Euclidean metrics physically, but can be more subtle to handle properly, so for most of these lectures we work with Euclidean metrics instead.)

If our worldsheet is an infinite strip, say, describing an interval propagating in time, then the tangent direction to the boundary can be identified with the time axis, and the normal to the boundary can be identified with the spacelike direction. Then, Dirichlet boundary conditions force the ϕ^μ to be independent of time. If we say that some coordinate ϕ^1, say, has Dirichlet boundary conditions, then that means that the boundary of the string will lie on the hyperplane $\phi^1 = 0$.

Thus, Dirichlet boundary conditions force the endpoints of the string to lie along some submanifold of the target space X, whereas Neumann boundary conditions allow the endpoints to propagate freely. By imposing Dirichlet boundary conditions on some of the fields ϕ^μ, we force the ends of the string to lie on some submanifold S, and in so doing, we obtain strings describing D-branes.

Earlier we said that a D-brane is roughly a choice of submanifold S together with some extra degrees of freedom we have not yet specified. Now we see where the first part of that data arises – as a specification of Dirichlet boundary conditions on some of the fields. (In fact, the "D" in the name D-brane comes from Dirichlet.)

An aside on nomenclature: Sometimes we specify the submanifold S by saying that the D-brane is *wrapped* on S. Also, one sometimes uses the notation Dp-brane to indicate a

D-brane with p spatial dimensions, though in these notes we will avoid that notation so as to sidestep confusion involving Euclidean versus Lorentzian metrics. Also, in these notes we will use the terms "D-brane" and "brane" interchangeably, though in general the terms are *not* interchangeable – there exist examples of branes which are not D-branes. We will not see any such in these notes, though.

From the supersymmetry transformations, compatible boundary conditions on the worldsheet fermions are, for Neumann boundary conditions

$$\psi_+^\mu = \psi_-^\mu$$

and for Dirichlet boundary conditions,

$$\psi_+^\mu = -\psi_-^\mu$$

Note that in general one could choose combinations of Dirichlet and Neumann boundary conditions that are not compatible with the complex structure – for $z = x + iy$, one could impose Neumann boundary conditions on the x direction and Dirichlet boundary condition on the y direction, for example. So, in general D-branes will wrap real submanifolds, not necessarily complex submanifolds.

Now, let us apply these boundary conditions to the B model. Because ψ_+^i and $\psi_+^{\bar{\imath}}$ couple to different bundles, one a worldsheet vector the other a worldsheet scalar, and the ψ_-'s are symmetric, to be consistent the boundary conditions must preserve the complex structure. Thus, the submanifold of the Calabi-Yau X specified by the boundary conditions, cannot be any submanifold, but rather must be a *complex* submanifold.

It should be trivial to check that for Neumann boundary conditions, $\theta_i = 0$, whereas for Dirichlet boundary conditions, $\eta^{\bar{\imath}} = 0$.

Now, in the closed string B model, we interpreted $\eta^{\bar{\imath}}$ as a $d\overline{z}^{\bar{\imath}}$ on the target space X, whereas the θ_i was interpreted as a vector field, coupling to TX. Here, if S is the submanifold specified by D-brane boundary conditions, then in the open string B model we would like to interpret $\eta^{\bar{\imath}}$ as $d\overline{z}^{\bar{\imath}}$ on the submanifold S, and θ_i as vectors in $TX|_S$ coupling to the normal bundle $\mathcal{N}_{S/X}$.

The careful reader will note that we need to be slightly careful about the interpretations above – we run the risk of needing to assume that $TX|_S$ splits holomorphically as $TS \oplus \mathcal{N}_{S/X}$, if we state the boundary conditions poorly. In general, although $TX|_S$ will so split as a smooth bundle, holomorphically it will not split. For example, consider conic curves in the projective plane $\mathbf{P}^2$. These curves are topologically $\mathbf{P}^1$'s, but embedded nontrivially. Here, $TC = \mathcal{O}(2)$, $T\mathbf{P}^2|_C = \mathcal{O}(3) \oplus \mathcal{O}(3)$, and $\mathcal{N}_{S/X} = \mathcal{O}(4)$. As C^∞ bundles on $\mathbf{P}^1$, $\mathcal{O}(3) \oplus \mathcal{O}(3) \cong$

$\mathcal{O}(2) \oplus \mathcal{O}(4)$, but they are distinct as holomorphic bundles. In general, we merely have an extension

$$0 \longrightarrow TS \longrightarrow TX|_S \longrightarrow \mathcal{N}_{S/X} \longrightarrow 0$$

In the case of the closed string B model, massless spectra were given by BRST-invariant field products corresponding to elements of the cohomology group $H^{\cdot}(X, \Lambda^{\cdot} T^{1,0} X)$. Here, the analogous computation gives BRST-invariant field products corresponding to elements of $H^{\cdot}(S, \Lambda^{\cdot} \mathcal{N}_{S/X})$. We shall return to this matter later, and will see how ultimately this gives rise to Ext groups between sheaves.

There is an additional consistency check of this choice of boundary conditions that we should perform. For Neumann boundary conditions, the string endpoints are allowed to propagate freely on the worldvolume of the D-brane, and so one would expect the equations of motion that one derives form the classical action to behave well.

So, let us turn to the nonlinear sigma model, on a worldsheet Σ that now has a boundary, and compute the equations of motion of the ϕ fields. From the bosonic kinetic term, proportional to

$$\int_{\Sigma} d^2 z g_{\mu\nu} \partial_\alpha \phi^\mu \partial^\alpha \phi^\nu$$

(where the α are worldsheet Euclidean indices, and μ, ν are target-space Euclidean indices), we see that when we compute the equations of motion for ϕ, we will find a term which we will have to integrate by parts:

$$\begin{aligned} \frac{\delta S}{\delta \phi^\lambda(z')} &= \cdots + \int_{\Sigma} d^2 z g_{\mu\lambda} \left(\partial_\alpha \phi^\mu \right) \left(\partial^\alpha \delta^2(z - z') \right) \\ &= \cdots + \int_{\Sigma} d^2 z \partial^\alpha \left(g_{\mu\lambda} \left(\partial_\alpha \phi^\mu \right) \delta^2(z - z') \right) - \int_{\Sigma} d^2 z g_{\mu\lambda} \left(\partial_\alpha \partial^\alpha \phi^\mu \right) \delta^2(z - z') \end{aligned}$$

The second term above will lead to the usual equations of motion. The first term, however, given by

$$\int_{\Sigma} d^2 z \partial^\alpha \left(g_{\mu\lambda} \left(\partial_\alpha \phi^\mu \right) \delta^2(z - z') \right) = \int_{\partial\Sigma} g_{\mu\lambda} \left(\partial_n \phi^\mu \right) \delta^2(z - z')$$

will be a problem when z' lies on the boundary of Σ. In order for the equations of motion to be valid everywhere, we need to impose the boundary condition $\partial_n \phi^\mu = 0$. So, as advertised, we have recovered Neumann boundary conditions as a constraint for well-behaved equations of motion for ϕ fields propagating freely along the worldvolume.

3.1.2 Chan-Paton factors

In addition to specifying boundary conditions on the bulk fields, we can also optionally add a boundary interaction, which couples to a connection on some principal bundle. For the case of a connection on a principal $U(1)$ bundle, this interaction on the boundary of the worldsheet Σ has the form

$$\int_{\partial\Sigma} \left[A_\mu d\phi^\mu \; - \; 2i F_{\mu\nu} \psi_+^\mu \psi_-^\nu \right] \tag{7}$$

along the submanifold specified by the boundary conditions (*i.e.* for Neumann boundary conditions). This boundary interaction is known as *Chan-Paton factors*, and the bundle to which the connection A_μ couples is known as the *Chan-Paton bundle*. It is straightforward to check that the interaction above is supersymmetric for B-model-compatible boundary conditions so long as the curvature tensor

$$F_{\mu\nu} \; = \; \partial_\mu A_\nu \; - \; \partial_\nu A_\mu$$

is of type $(1,1)$, *i.e.* $F_{ij} = F_{\overline{\imath}\overline{\jmath}} = 0$.

Now, the presence of such an interaction modifies the boundary conditions for the open string. The reason for this goes back to our argument that Neumann boundary conditions leads to well-behaved equations of motion: in the present case, the presence of Chan-Paton factors modifies the boundary contributions, and so should modify our conclusions. (This was first discussed in [26], and our discussion here will follow their conventions.)

Begin with the relevant bosonic part of the action, with all factors added:

$$\frac{1}{\alpha'} \int_\Sigma d^2 z \left[\frac{1}{2} g_{\mu\nu} \partial \phi^\mu \overline{\partial} \phi^\nu \right] \; - \; \frac{1}{4\alpha'} \int_{\partial\Sigma} A_\mu d\phi^\mu \; = \; \frac{1}{8\alpha'} \int_\Sigma d^2 z \left[g_{\mu\nu} \partial^\alpha \phi^\mu \partial_\alpha \phi^\nu \right] \; - \; \frac{1}{4\alpha'} \int_{\partial\Sigma} A_\mu d\phi^\mu$$

(Factors in front of the Chan-Paton factor were chosen for later convenience.) When we compute the equations of motion, we get the boundary terms

$$\begin{aligned}
\frac{2}{8\alpha'} \int_\Sigma d^2 z \left[\partial_\alpha \left(g_{\mu\lambda} \partial^\alpha \phi^\mu \delta^2(z-z') \right) \right] \; - \; \frac{1}{4\alpha'} \int_{\partial\Sigma} \partial_\lambda A_\mu \delta^2(z-z') d\phi^\mu \; - \; \frac{1}{4\alpha'} \int_{\partial\Sigma} A_\lambda d \left(\delta^2(z-z') \right) \\
= \; \frac{1}{4\alpha'} \int_{\partial\Sigma} dx \left[g_{\mu\lambda} \partial_n \phi^\mu \delta^2(z-z') \right] \; - \; \frac{1}{4\alpha'} \int_{\partial\Sigma} F_{\lambda\mu} \delta^2(z-z') d\phi^\mu \\
= \; \frac{1}{4\alpha'} \int_{\partial\Sigma} dx \delta^2(z-z') \left[g_{\mu\lambda} \partial_n \phi^\mu \; - \; F_{\lambda\mu} \partial_t \phi^\mu \right]
\end{aligned}$$

so now we see that for the equations of motion to be well-behaved at the boundaries, we must require

$$g_{\mu\lambda} \partial_n \phi^\mu \; = \; F_{\lambda\mu} \partial_t \phi^\mu$$

in place of the original Neumann boundary conditions (to which this reduces when $F = 0$). From supersymmetry, the B-model fermions should then obey

$$\theta_i \; = \; F_{i\overline{\jmath}} \eta^{\overline{\jmath}}$$

in place of the original Neumann boundary conditions $\theta_i = 0$. (As usual, we have to be slightly careful about the interpretation of that constraint when $TX|_S$ does not split holomorphically as $TS \oplus \mathcal{N}_{S/X}$.)

3.1.3 Notes on the open string A model

In most of these notes, we are concerned with the B model, since that is where derived categories enter physics. However, for completeness, let us briefly mention a few facts concerning the open string A model.

In the A model, it is ψ_+^i and $\psi_-^{\bar{\imath}}$ that are both worldsheet scalars, so consistent boundary conditions will have to relate ψ_+^i to $\psi_-^{\bar{\imath}}$ – which means that the boundary conditions do not respect the complex structure on X. There is an analogous modification to the conditions on the curvatures of the Chan-Paton factors.

With more work along these lines, one can show that the A model 'typically' couples to D-branes on Lagrangian submanifolds of X together with flat vector bundles over those submanifolds.

There are a few special exceptions to that rule – it is also possible for the open string A model to describe D-branes on coisotropic submanifolds with non-flat vector bundles. See [38] for more information.

3.1.4 Further complications

Very briefly, let us mention in passing two further complications. They play an important role, but understanding that role is beyond the scope of these lectures.

One complication is the "Freed-Witten anomaly" [30]. This says two things. First, a D-brane can not consistently wrap a submanifold unless the normal bundle admits a Spin^c structure (or unless one makes $H \neq 0$, but that would lead to a discussion of Hitchin's generalized complex geometry, beyond the scope of these lectures). In the B model, all submanifolds are complex submanifolds of complex manifolds, and for these the normal bundle will always admit a Spin^c structure, so there is no obstruction. Second, the Freed-Witten anomaly tells us that if the normal bundle admits a Spin^c structure but not a Spin structure, then the Chan-Paton factors must be 'twisted.' This, unfortunately, *is* a problem for us, For example, the tangent bundle of $\mathbf{P}^2$ admits a Spin^c structure but not a Spin structure, and for submanifolds of Calabi-Yau's, the normal bundle admits a Spin structure if and only if the submanifold does. Thus, if we wrap a D-brane on a $\mathbf{P}^2$ inside a Calabi-Yau, then the Chan-Paton factors must be suitably twisted. In practice, what this means is that the D-brane corresponding to the sheaf $i_* \mathcal{E}$, where $i : S \hookrightarrow X$ and $\mathcal{E}$ a holomorphic vector bundle on S, will have Chan-Paton factors coupling to $\mathcal{E} \otimes \sqrt{K_S}$ instead of $\mathcal{E}$, where K_S is the canonical bundle of S [31]. In general, $\sqrt{K_S}$ will not be an honest bundle, merely a 'twisted' bundle, but we shall not attempt to explain what that means in detail here. This difference, although it plays a crucial role in many technical discussions of derived categories in physics [31], will not appear in these lectures, so we will not belabor it further.

Another complication that emerges in general is the open string analogue of the Calabi-Yau condition in the B model. Just as the structure of the Grassmann zero mode integrals in the path integral implies that the closed string B model can only be defined on spaces X with $K_X^{\otimes 2} \cong \mathcal{O}$, there is an analogous condition in the open string B model which naively appears to significantly restrict which D-branes one can consistently stretch an open string between [31]. In the simplest cases, for trivial Chan-Paton factors, the constraint is that an open string can stretch between two D-branes wrapped on submanifolds S, T if

$$\Lambda^{top}\mathcal{N}_{S\cap T/S} \otimes \Lambda^{top}\mathcal{N}_{S\cap T/T} \;\cong\; \mathcal{O}$$

This condition also plays an important role in more detailed analyses [31]; however, the fact that the Riemann surface has boundary complicates the question of whether this is truly a constraint. A complete discussion of this matter does not yet exist.

3.2 Relevance of sheaves

So far, we have described D-branes in the open string B model as, a complex submanifold $i : S \hookrightarrow X$ together with a holomorphic vector bundle $\mathcal{E} \to S$. Such a pair can be naturally described as the coherent sheaf $i_* \mathcal{E}$.

To make that identification useful, we need to be able to compute physical quantities in terms of sheaves, and in the next section we shall see how that can be done.

Given that sheaves of the form $i_* \mathcal{E}$ are relevant, a good question to ask is, what about other coherent sheaves? Do they play any role, or are they physically irrelevant? To this there are two answers:

- For a more general coherent sheaf, by replacing it with a projective resolution we will be oble to construct a corresponding physical system consisting of D-branes and antibranes, see section 3.4. Assuming that worldsheet renormalization group flow in the B model is the same thing as localization on quasi-isomorphisms, then this gives us one way to associate physics to more general coherent sheaves. Alas, it is not terribly satisfying, as it does not directly address the question of how to interpret more general coherent sheaves as individual D-branes.
- A direct attack on this question was conducted in [32]. There, it was argued that coherent sheaves of the form of structure sheaves of nonreduced subschemes of X, and related sheaves, could be directly interpreted as D-branes with nonvanishing 'Higgs field vevs,' which was checked by comparing massless spectra to Ext groups. That still doesn't give an interpretation of all sheaves, but fills in a significant gap, and gives some idea of what a general case might look like.

3.3 Massless spectra = Ext groups

Historically, sheaves were first introduced to the study of D-branes by Harvey and Moore in [27], as a throwaway observation. The reason that this description is of interest to physicists is that sheaves can be used to do physics computations.

For one example, massless spectra can be computed using sheaf theory. The states of the open string B model, on an infinite strip with the boundaries associated to two distinct D-branes / sheaves, should be some sort of cohomology, dependent upon an ordered pair of sheaves. A natural guess is that they should be computed by Ext groups between the sheaves on either side.

Physicists discovered experimentally after [27] that this natural guess seemed to be correct, as Ext groups do indeed compute spectra in a number of examples. However, understanding *why* it should be correct took longer.

Some basic cases were implicit in [29]. In that paper, the open string B model for open strings connecting D-branes wrapped on the entire space was described. In this case, the open string B model states – states effectively on the boundary of an infinite strip in the far future or past – are of the form

$$b(\phi)^{\alpha\beta}_{\bar{\imath}_1\cdots\bar{\imath}_n}\eta^{\bar{\imath}_1}\cdots\eta^{\bar{\imath}_n}$$

(utilizing a Dolbeault-type realization of such groups) where α, β are Chan-Paton indices, coupling to holomorphic vector bundles associated with either boundary. The BRST cohomology of these states is in one-to-one correspondence with elements of the sheaf cohomology groups

$$H^n\left(X, \mathcal{E}^\vee \otimes \mathcal{F}\right)$$

Understanding how Ext groups arise for D-branes wrapped on submanifolds took a much longer time. It was first worked out in [31]. To illustrate how this works, let us specialize to the case of two sets of D-branes both wrapped on the same submanifold $i : S \hookrightarrow X$, both with the same holomorphic vector bundle $\mathcal{E}$, which to simplify things even further, we shall assume is a line bundle. The physical spectrum computation realizes the spectral sequence[19]

$$H^p\left(S, \mathcal{E}^\vee \otimes \mathcal{E} \otimes \mathcal{N}_{S/X}\right) \Longrightarrow \mathrm{Ext}^{p+q}(i_*\mathcal{E}, i_*\mathcal{E})$$

Suppose, for example, we wish to describe an element of $\mathrm{Ext}^1(i_*\mathcal{E}, i_*\mathcal{E})$. On the face of it, one would guess that there are contributions from field products of the forms

$$b^{i\alpha\beta}\theta_i, \quad b'^{\alpha\beta}_{\bar{\imath}}\eta^{\bar{\imath}}$$

[19] For those who wish to understand how this spectral sequence is derived, it can be derived as a specialization of the local-to-global spectral sequence.

The trick is, making sense of BRST cohomology. If the $TX|_S$ splits holomorphically as $TS \oplus \mathcal{N}_{S/X}$, and the Chan-Paton bundle $\mathcal{E}$ is trivial, then the BRST operator acts just as $\overline{\partial}$, and the states above span

$$H^0(S, \mathcal{E}^\vee \otimes \mathcal{E} \otimes \mathcal{N}_{S/X}) \oplus H^1(S, \mathcal{E}^\vee \otimes \mathcal{E})$$

which is noncanonically isomorphic to $Ext^1(i_*\mathcal{E}, i_*\mathcal{E})$, since it can be shown the spectral sequence degenerates. If $TX|_S$ does not split holomorphically, and $\mathcal{E}$ is not trivial, then we have to work harder. In this case, we cannot make sense of θ_i globally on S, both partly because $TX|_S$ does not split holomorphically, so we cannot make sense of θ_i by itself, and also because the boundary conditions are modified, so that $\theta_i = F_{i\overline{\jmath}}\eta^{\overline{\jmath}}$. We can lift $\mathcal{N}_{S/X}$ to $TX|_S$ in order to make sense of the expression "$b^{i\alpha\beta}\theta_i$," albeit at the cost of breaking the holomorphic structure, so that it is no longer BRST closed by itself. If we then demand that after we apply the boundary condition $\theta_i = F_{i\overline{\jmath}}\eta^{\overline{\jmath}}$ that the state be closed, then in effect we have the composition

$$H^0(S, \mathcal{E}^\vee \otimes \mathcal{E} \otimes \mathcal{N}_{S/X}) \longrightarrow H^0(S, \mathcal{E}^\vee \otimes \mathcal{E} \otimes TX|_S) \longrightarrow H^1(S, \mathcal{E}^\vee \otimes \mathcal{E})$$

which is precisely [31] the relevant differential of the spectral sequence.

3.4 Brane/antibrane configurations and derived categories

Finally, we are ready to begin to understand how derived categories appear physically in the open string B model. For any object of a derived category $D^b(X)$ over a smooth Calabi-Yau X, *i.e.* a complex of sheaves, the general idea is to map that complex of sheaves to a set of alternating D-branes and anti-D-branes (corresponding to the sheaves) with 'tachyons' between them corresponding to the maps in the complex.

Before trying to make this precise, let us take a moment to explain what an 'anti-D-brane' and a 'tachyon' are. An *anti-D-brane* is specified by the same data as a D-brane, in fact in the B model is nearly identical to a D-brane, except that physically anti-D-branes and D-branes try to annihilate one another. Their worldsheet specification is identical – same boundary conditions, same Chan-Paton factors. They are distinguished from one another by the addition of a certain operator to the boundary of an open string between a section corresponding to a brane and another section corresponding to an antibrane. That operator shifts a certain '$U(1)$' charge, which modifies the physical interpretation.

In any event, a physical theory containing both branes and antibranes is not stable. A *tachyon* is a field[20] that pops up whenever there are simultaneously branes and anti-branes

[20] A field in the effective theory on the target space. The tachyon is not a field on the worldsheet, though it can be represented on the worldsheet by a BRST-closed field combination corresponding to an element of $\mathrm{Ext}^0_X(\mathcal{S}_1, \mathcal{S}_2)$.

present. (More generally, the appearance of tachyons is correlated with the appearance of instability, and the fact that branes and anti-branes try to annihilate one another makes a theory containing both very unstable.) Technically, if one has branes and antibranes both wrapped on the same submanifold S of spacetime, then the tachyon that appears is a map between their respective holomorphic vector bundles.

Now, let us clean up the proposed dictionary above slightly. As remarked earlier, we do not know how to associate D-branes to general sheaves, only to certain sheaves. However, in the derived category that is not a problem. For any object of a derived category $D^b(X)$ over a smooth Calabi-Yau X, pick a representative in the same equivalence class that is a complex of locally-free sheaves, *i.e.* holomorphic vector bundles on X. Then, that complex we can certainly represent physically – with branes and antibranes wrapped on all of the space X, with holomorphic vector bundles on their worldvolumes.

This proposal was originally made in [34], and later repeated in [35].

However, this proposal, as stated so far, still leaves a lot of questions unanswered. For example: how does the condition for a complex arise physically in terms of properties of tachyons?

This question was not answered until the publication of [36]. For simplicity, let us again assume that the holomorphic vector bundles on the D-brane and anti-brane worldvolumes are line bundles, so as to avoid subtleties in more general cases. A tachyon is represented on the open string worldsheet by a BRST-closed field product corresponding to an element of $\mathrm{Ext}^0_X(\mathcal{E}, \mathcal{F})$, when both the branes and anti-branes are wrapped on all of X, where $\mathcal{E}$ and $\mathcal{F}$ are the holomorphic vector bundles on each worldvolume. Such a BRST-closed field product is represented by $\varphi^{\alpha\beta}$, where α, β are indices coupling to $\mathcal{E}$, $\mathcal{F}$, and with no other factors of $\eta^{\bar{\imath}}$ or θ_i. To describe open strings in the presence of a nonzero tachyon, we add the term

$$\int_{\partial\Sigma} G \cdot \varphi^{\alpha\beta} \;\propto\; \int_{\partial\Sigma} \left(\psi_+^i \,+\, \psi_-^i \right) \partial_i \varphi^{\alpha\beta}$$

to the boundary action, where G represents the non-BRST half of the supersymmetry transformations.

> In fact, we need to do a bit more, though this was not mentioned in [36]. The complete boundary action describing a nonzero tachyon background also includes a bosonic potential term of the form $\varphi^*\varphi$ (see for example [39][section 5.1.2], [40][section 4], [41][section 2], or [42]; there is a short discussion also in section 4.2.2).

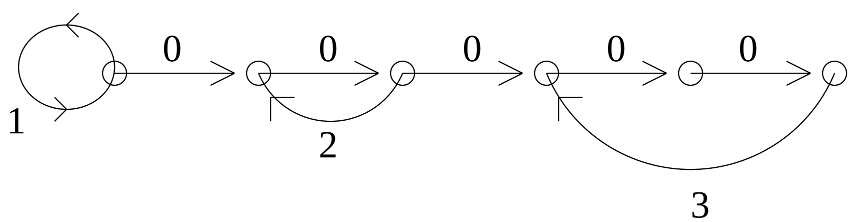
The addition of the term above to the boundary action has the effect of modifying the BRST operator. In the closed string B model, as discussed previously, the BRST operator can be interpreted simply as $\overline{\partial}$. Here, however, the BRST operator is modified[21] to become $\overline{\partial} + \varphi$.

[21] One way to see this is to evaluate the correlation function $< Q \cdot V >$ for any V, and then commute the BRST operator Q past the new boundary terms. See [36, 37] for more information.

A necessary condition to have supersymmetry along the boundary in the twisted theory is that $Q^2 = 0$, which now implies that φ must be holomorphic ($\overline{\partial}\varphi = 0$), and also, if we have several different branes and antibranes with tachyons between them, that composition of successive tachyons must vanish:

$$\varphi_1^{\alpha\beta}\varphi_2^{\beta\gamma} \;=\; 0$$

In this fashion, we find the condition to have a complex.

> In fact, we need not stop there. We can also consider open strings in the presence of other Ext elements. Elements of Ext^1 merely deform the sheaves; elements of Ext^n for $n > 1$ lead to the "generalized complexes" of [43]. An example of such a generalized complex is given by
>
> 
>
> in which the Ext degrees corresponding to each map have been labelled. As above, the BRST operator Q is deformed, and the condition $Q^2 = 0$ implies a constraint among the various maps which is exactly the condition for a generalized complex in the sense of [43]. These generalized complexes can be used to build "enhanced triangulated categories" as well as an enhancement of ordinary derived categories of sheaves. See [37, 44, 45, 46] for more information.

So far we have described how complexes can be understood physically – as sets of branes and antibranes with tachyons corresponding to the maps between them. Now, how do we realize, for example, localization on quasi-isomorphisms?

The answer is that localization on quasi-isomorphisms is believed to be realized by worldsheet renormalization group flow. To see this, let us consider a simple example: a single D-brane wrapped on a divisor D, versus a brane and anti-brane with bundles $\mathcal{O}$, $\mathcal{O}(-D)$ and a tachyon realizing the complex

$$\mathcal{O}(-D) \;\longrightarrow\; \mathcal{O}$$

These are equivalent to one another as objects of the derived category $D^b(X)$. Physically, however, they are significantly different. The D-brane wrapped on a divisor D is described by a conformally-invariant theory, whereas in the open string describing the brane/anti-brane system, the boundary terms describing a nonzero tachyon break scale-invariance of the worldsheet theory, classically. Thus, in this simple example we see that two open strings describing equivalent objects of the derived category are described by very different physics: one scale-invariant, the other not. Now, the renormalization group can product a scale-invariant theory from one that is not, so the correct claim to make is that different representatives of the same object of the derived category should be realized physically by different open string worldsheets in the same universality class of worldsheet renormalization group flow.

As discussed earlier, it is not possible in general to try to follow renormalization group flow exactly. The best we can usually do is to compute an asymptotic series expansion to a tangent vector to the flow. Instead, to check this conjecture, we work through large numbers of computations, and if we can find no contradictions, then eventually we believe that it is *probably* (though not provably) correct.

> There is a closely analogous problem in the realization of (smooth Deligne-Mumford) stacks in physics. Stacks look locally like global quotients by finite groups, but, such a local description cannot be used to built a QFT. Stacks also have a global description as quotients $[X/G]$ by groups G that need be neither finite nor effectively-acting, and that description can be translated into physics. Unfortunately, when G is finite, the corresponding physical theory is scale-invariant, whereas when G is not finite, the corresponding physical theory is not scale-invariant. Since a given stack can have multiple presentations of the form $[X/G]$ for G's both finite and non-finite, we again have a potential presentation-dependence problem, which we conjecture is resolved by renormalization group flow – stacks classify universality classes of physical theories. Unfortunately, unlike the case of derived categories, there are several basic consistency checks that appear to be violated, making the question of whether it makes sense to talk about strings propagating on stacks, somewhat problematic for several years. However, those issues have been resolved and the conjecture has by now been checked in a large number of different ways. See [12, 13, 14, 15, 16, 17] for more details.

3.5 Ext groups from tachyons

Previously we have discussed how Ext groups arise in direct computations involving D-branes wrapped on curves. It is also possible to derive Ext groups, or more properly $\mathbf{R}$Hom's, for general objects in the derived category, using the brane / anti-brane / tachyon language.

To illustrate the method, let us work through an example. Let us calculate $\mathrm{Ext}^*_{\mathbf{C}}\left(\mathcal{O}_D, \mathcal{O}\right)$ where D is some divisor on $\mathbf{C}$. The torsion sheaf $\mathcal{O}_D$ has a two-step resolution:

$$0 \longrightarrow \mathcal{O}(-D) \longrightarrow \mathcal{O} \longrightarrow \mathcal{O}_D \longrightarrow 0$$

so it is easy to check that the answer should be given by

$$\dim \mathrm{Ext}^n_{\mathbf{C}}\left(\mathcal{O}_D, \mathcal{O}\right) \;=\; \begin{cases} 0 & n=0 \\ 1 & n=1 \end{cases}$$

We shall check our physical picture by verifying, through a physical calculation, the Ext groups just described. (We will find that the physical calculation is too messy for practical use, but working through the details should illuminate the ideas.)

On one end of the open string, put Chan-Paton factors corresponding to the bundle $\mathcal{O}(-D)\oplus \mathcal{O}$, and on the other end of the open string, put Chan-Paton factors corresponding to $\mathcal{O}$. We shall also give a vev to the tachyon corresponding to the holomorphic map $\phi : \mathcal{O}(-D) \rightarrow \mathcal{O}$, on the corresponding end of the open string. As a result of this tachyon vev, the BRST operator is now deformed to the form $Q_{BRST} = \overline{\partial} + \phi^{\vee}$. (We write $\phi^{\vee}$ instead of ϕ to emphasize the fact that ϕ acts solely within the dualized half of the Chan-Paton factors.)

Proceeding as before, boundary BRST-invariant field combinations are of the general form

$$b^{\alpha\beta}_{\overline{\imath}_1 \cdots \overline{\imath}_n} \eta^{\overline{\imath}_1} \cdots \eta^{\overline{\imath}_n}$$

where α, β are Chan-Paton factors coupling to $\mathcal{O}(-D) \oplus \mathcal{O}$ and $\mathcal{O}$, respectively. Since we will need to work with the detailed components, let us write

$$(b^{\alpha\beta}_{\overline{\imath}_1 \cdots \overline{\imath}_n}) \: = \: \left[\begin{array}{c} b_{0 \overline{\imath}_1 \cdots \overline{\imath}_n} \\ b_{1 \overline{\imath}_1 \cdots \overline{\imath}_n} \end{array} \right]$$

where b_0 is associated with $\mathcal{O}(-D)^{\vee} \otimes \mathcal{O}$ and b_1 i s associated with $\mathcal{O}^{\vee} \otimes \mathcal{O}$.

Now, let us compute BRST cohomology. We need to be careful to keep track of degrees properly – for example, degree zero states are *not* of the form

$$\left[\begin{array}{c} b_0 \\ b_1 \end{array} \right]$$

since the tachyon ϕ forces b_0 and b_1 to have different charges. Instead, the only degree zero state is b_1, and the condition for this state to be BRST closed is that $\overline{\partial} b_1 = 0$ and $\phi^{\vee} b_1 = 0$. The only holomorphic function on $\mathbf{C}$ that is annihilated by multiplication by x (assuming D is the divisor $\{ x = 0 \}$, without loss of generality) is the zero function, hence the space of degree zero states is zero-dimensional, exactly as desired.

The degree one states are of the form

$$b_0 \: + \: b_{1 \overline{\imath}} \eta^{\overline{\imath}}$$

The condition for these states to be BRST-closed is that

$$\begin{aligned} \overline{\partial} b_0 \: &= \: - \phi^{\vee} \left(b_{1 \overline{\imath}} \eta^{\overline{\imath}} \right) && (8) \\ \overline{\partial} \left(b_{1 \overline{\imath}} d \overline{z}^{\overline{\imath}} \right) \: &= \: 0 && (9) \end{aligned}$$

and BRST-exact states are of the form

$$\begin{aligned} b_0 \: &= \: \phi^{\vee} a \\ b_{1 \overline{\imath}} d \overline{z}^{\overline{\imath}} \: &= \: \overline{\partial} a \end{aligned}$$

for some a. Condition (9) means that $b_{1\bar{i}}\eta^{\bar{i}}$ is an element of $H^1(\mathcal{O})$. Condition (8) means that if we define $b_0' = b_0 \mod \mathrm{im}\, \phi^\vee$, then $\overline{\partial} b_0' = 0$, and more to the point,

$$b_0' \in H^0\left(D, \mathcal{O}(-D)^\vee|_D \otimes \mathcal{O}|_D\right) = \mathrm{Ext}^1\left(\mathcal{O}_D, \mathcal{O}\right)$$

(Technically b_0' can be interpreted as defining a form on D because we are modding out the image of an element of $H^1(\mathcal{O})$.) Conversely, given an element of

$$\mathrm{Ext}^1\left(\mathcal{O}_D, \mathcal{O}\right) = H^0\left(D, \mathcal{O}(-D)^\vee|_D \otimes \mathcal{O}|_D\right)$$

we can define b_0 and b_1, using the long exact sequence

$$\cdots \longrightarrow H^0(\mathcal{O}) \longrightarrow H^0(\mathcal{O}(D)) \longrightarrow H^0(D, \mathcal{O}(D)|_D) \xrightarrow{\delta} H^1(\mathcal{O}) \longrightarrow \cdots$$

The element b_1 is the image under δ, and b_0 is the lift to an element of $\mathcal{A}^{0,0}(\mathcal{O}(D))$.

This is clearly not very efficient, but the fact that it can be done at all is a good consistency check of our assumption that renormalization group flow realizes localization on quasi-isomorphisms.

3.6 Correlation functions

One thing that one would certainly like to be able to do is compute correlation functions in the open string B model. In particular, correlation functions on a disk with various Ext group elements inserted at points on the boundary of the disk.

Unfortunately, direct computations are often very difficult, but, we can apply ideas from the realization of derived categories in physics to simplify the computations.

3.6.1 Direct computation attempt

Let us briefly outline an example to illustrate the difficulty of direct computation of correlation functions in the open string B model.

In particular, let us consider an example involving D-branes wrapped on an obstructed $\mathbf{P}^1$ inside a Calabi-Yau threefold, with normal bundle $\mathcal{O} \oplus \mathcal{O}(-2)$. The structure of the normal bundle suggests that the $\mathbf{P}^1$ should come in a one-parameter family (since $h^0(\mathbf{P}^1, \mathcal{N}) = 1$), but, the normal bundle represents only a linearization of the complex structure, and omits information. For example, an nth order obstruction is realized by a complex structure described as follows. Let (x, y_1, y_2), (w, z_1, z_2) be two coordinate patches on the total space of the normal bundle over $\mathbf{P}^1$, then the (deformation obstructed) complex structure is defined by the transition functions

$$\begin{aligned} w &= x^{-1} \\ z_1 &= x^2 y_1 + x y_2^n \\ z_2 &= y_2 \end{aligned}$$

The fact that deformations are obstructed should be represented physically[22] by a nonzero correlation function on a disk involving operators corresponding to elements of $\mathrm{Ext}^1_X(\mathcal{S}, \mathcal{S})$ where $\mathcal{S}$ is a sheaf describing a D-brane wrapped on the obstructed $\mathbf{P}^1$.

For simplicity, let us assume that the vector bundle on the worldvolume of the D-brane is a trivial line bundle. (We can also assume, without loss of generality, that $TX|_{\mathbf{P}^1}$ splits holomorphically as $T\mathbf{P}^1 \oplus \mathcal{N}$.) Then the BRST-closed field product corresponding to an element of $\mathrm{Ext}^1_X(\mathcal{S}, \mathcal{S})$ is of the form $b(\phi)^{i\alpha\beta}\theta_i$. For a particular order of obstruction[23], the correlation function

$$< \left(b^{i\alpha\beta}\theta_i \right)^3 >$$

should be nonzero.

Previously, we have been able to compute correlation functions in topological field theories by little more than counting zero modes. However, that does not work here. In the present case, since we are working in the B model, we need consider only constant maps into the target space, so zero-mode counting reduces to counting the ranks of bundles and applying boundary conditions. In particular, there are 2 θ_i zero modes (corresponding to the rank of the normal bundle to the $\mathbf{P}^1$), and one $\eta^{\bar{\imath}}$ zero mode (since the rank of $T\mathbf{P}^1$ is one). On the face of it, to get a nonvanishing correlation function, we would need the correlation function to contain one η and two θ's, but our correlation function above contains three θ's, so naively one would think that our correlation function should always vanish.

The subtlety we are forgetting is the freedom to use the worldsheet bulk interaction term

$$\frac{1}{\alpha'} \int_\Sigma R_{i\bar{\jmath}k\bar{l}} \psi_+^i \psi_+^{\bar{\jmath}} \psi_-^k \psi_-^{\bar{l}}$$

By performing various Wick contractions, (legitimate here so long as the final result is independent of α'), one can derive an expression for the correlation function involving the Riemann curvature R. In principle, the information about the obstruction in the complex structure should be encoded in the (hermitian, Ricci-flat) metric on the Calabi-Yau; however, we do not know of an explicit way to relate them, so we cannot confirm independently that the result really is correctly reproducing the order of the obstruction.

[22] We shall simply state this as a fact. This will be more or less obvious to physicists reading these notes, but, to explain to others, would require more time than we wish to devote to the point.

[23] For higher orders, the correlation function should involve "descendants" of the $b(\phi)\theta_i$'s, so for simplicity we shall only consider this case.

Clearly, direct computations are cumbersome at best.

In the next section, we shall briefly outline a more efficient method to achieve the same ends.

3.6.2 Aspinwall-Katz methods

A more efficient method to compute correlation functions [47] is to take advantage of the fundamental assumption underlying the description of derived categories in physics: namely, that localization on quasi-isomorphisms is accomplished via renormalization group flow. In particular, correlation functions in topological subsectors of topological field theories are, in principle, independent of representative of universality class, so if a computation is very difficult in one presentation, then, pick a different presentation of the same universality class and compute there instead. In the present case, since computations directly involving branes wrapped on submanifolds are different, we shall work instead with brane/antibrane/tachyon combinations on the entire space. One should get equivalent results, but hopefully with less effort.

Open string correlation functions on a disk, with all operators located on the edge of the disk, have a natural A_∞ structure. Roughly, the n-point correlation function

$$< \mathcal{O}_1 \cdots \mathcal{O}_n >$$

can be understood as $< \mathcal{O}_1, m_{n-1}(\mathcal{O}_2, \cdots, \mathcal{O}_n) >$ where $<,>$ denotes an inner product (the composition of m_2 and a trace that exists for Calabi-Yau's), and m_n is the n-point multiplication. Not any A_∞ structure can be allowed – since the $\mathcal{O}_i$ are inserted on the edge of a disk, clearly there must exist a cyclic symmetry. For brevity, we will not attempt to list either necessary or sufficient conditions here.

Given some A_∞ algebra A, we can take the cohomology of m_1 (the 1-point multiplication) to get an $H^\cdot(A)$. It can be shown that one can define an A_∞ structure on $H^\cdot(A)$ such that there is an A_∞ morphism $f : H^\cdot(A) \longrightarrow A$ with f_1 equal to a fixed embedding $H^\cdot(A) \hookrightarrow A$ defined by taking representatives, and such that $m_1 = 0$ on $H^\cdot(A)$.

Briefly, the idea behind the method of [47] is to take A to be the differential graded algebra of bundle-valued forms on X with multiplication defined by the wedge product, then the A_∞ structure one can derive on $H^\cdot(A)$ defines the disk correlation functions in the open string B model.

3.7 What is the D-brane category?

So far, we have discussed the D-branes in the open string B model as described by $D^b(X)$, together with higher (Massey) products defined by correlation functions on a disk and defining an A_∞ structure.

Alternatively, and perhaps more efficiently, we can think of the D-brane category as a dg-category, a dg-enhancement of the derived category. There are several (quasi-equivalent) dg-enhancements; for example, one could take the category of complexes of vector bundles on X with the hom complexes being the Dolbeault complexes of hom's. In this language, the derived category $D^b(X)$ with its A_∞ structure can be thought of as the minimal model of the corresponding dg-category, and on general principles [56] these two pictures are quasi-equivalent. So, in principle, one can work with either one.

3.8 Exercises

1. Check that the nonlinear sigma model still possesses supersymmetry on a Riemann surface with boundary, when one imposes either Dirichlet or Neumann boundary conditions in the B model. For simplicity, take the metric to be constant, and also assume the Chan-Paton factors are trivial.

2. Check that the boundary interaction (7) describing the Chan-Paton factors is invariant under supersymmetry when the boundary conditions are B-model-compatible and the connection has curvature of type $(1,1)$, *i.e.* $F_{ij} = F_{\overline{\imath}\overline{\jmath}} = 0$. For simplicity, assume the metric is constant and that the fields on $\partial\Sigma$ obey Neumann boundary conditions $\partial\phi^\mu = \overline{\partial}\phi^\mu$, $\psi_+^\mu = \psi_-^\mu$.

3. We mentioned that non-reduced subschemes correspond physically to D-branes with "nilpotent Higgs fields." In this problem, we shall outline an example of what that means. Consider open strings on $X = \mathbf{C}^2$ connecting two sets of D-branes each supported at a point, the origin of $\mathbf{C}^2$. Each D-brane will have a trivial rank 2 vector bundle. However, one of the D-branes will have a nontrivial Higgs field. A Higgs field is a section of $\mathcal{N}_{S/X} \otimes \mathrm{End}\, \mathcal{E}$, which in this simple case means a pair of 2×2 matrices. Take that Higgs field to be defined by the two matrices

$$\Phi_x = \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array} \right], \quad \Phi_y = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right]$$

(Mathematically, these Higgs fields define a deformation of the ring action on the module describing the rank 2 locally-free sheaf; the resulting (deformed) module is that of a nonreduced scheme.) This Higgs field deforms the BRST operator to

$$Q = \overline{\partial} + \Phi_x \theta_x + \Phi_y \theta_y$$

where the θ's are the θ's of the B model, namely Grassmann-valued fields on $S =$ point, and the Φ's act by matrix multiplication on the left, say. (In this case, the $\overline{\partial}$ is irrelevant, since S is a point – everything will be constant, this problem is an exercise in matrix multiplication.) Now, "states of degree zero" have the form

$$\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$$

for complex numbers a, b, c, d, "states of degree one" have the form

$$\left[\begin{array}{cc} a_x & b_x \\ c_x & d_x \end{array}\right]\theta_x \; + \; \left[\begin{array}{cc} a_y & b_y \\ c_y & d_y \end{array}\right]\theta_y$$

and "states of degree two" have the form

$$\left[\begin{array}{cc} e & f \\ g & h \end{array}\right]\theta_x\theta_y$$

Compute the BRST cohomology of these states (hint: nothing more than an exercise in computing kernels and images of matrix multiplication). Show that the BRST cohomology of these states at degree n matches the Ext groups

$$\mathrm{Ext}^n_{\mathbf{C}^2}\left(D_x, \mathcal{O}_0^2\right)$$

where $\mathcal{O}_0$ is the structure sheaf of the origin of $\mathbf{C}^2$, and D_x is the structure sheaf of a nonreduced subscheme of $\mathbf{C}^2$ defined by the ideal (x^2, y). Use the locally-free resolution

$$0 \longrightarrow \mathcal{O}_{\mathbf{C}^2} \stackrel{\left[\begin{array}{c} -y \\ x^2 \end{array}\right]}{\longrightarrow} \mathcal{O}_{\mathbf{C}^2}^2 \stackrel{[x^2,y]}{\longrightarrow} \mathcal{O}_{\mathbf{C}^2} \longrightarrow D_x \longrightarrow 0$$

NOTE that in effect, we have turned the problem of computing Ext groups between structure sheaves of nonreduced subschemes, which sounds horribly nasty, into a simple exercise in linear algebra. More information on this can be found in [37, 32].

4 Landau-Ginzburg models

4.1 Closed strings

A *Landau-Ginzburg model* is a nonlinear sigma model with a (supersymmetrized) potential added. To specify a Landau-Ginzburg model, one must specify both a complex Riemannian manifold as well as a holomorphic function over that Riemannian manifold.

The most general Landau-Ginzburg model (over a space) that one can write down has the following action:

$$\frac{1}{\alpha'}\int_\Sigma d^2 z \Big(\frac{1}{2} g_{\mu\nu}\partial\phi^\mu \overline{\partial}\phi^\nu \; + \; \frac{i}{2} B_{\mu\nu}\partial\phi^\mu\overline{\partial}\phi^\nu \; + \; \frac{i}{2} g_{\mu\nu}\psi_-^\mu D_z \psi_-^\nu \; + \; \frac{i}{2} g_{\mu\nu}\psi_+^\mu D_{\overline{z}}\psi_+^\nu \; + \; R_{i\overline{j}k\overline{l}}\psi_+^i\psi_+^{\overline{j}}\psi_-^k\psi_-^{\overline{l}}$$
$$- g^{i\overline{j}}\partial_i W \partial_{\overline{j}}\overline{W} \; + \; i\psi_+^i\psi_-^j D_i\partial_j W \; + \; i\psi_+^{\overline{\imath}}\psi_-^{\overline{\jmath}} D_{\overline{\imath}}\partial_{\overline{\jmath}}\overline{W}\Big)$$

where W is a holomorphic function over the target space X, known as the *superpotential*, and

$$D_i \partial_j W \;=\; \partial_i \partial_j W \;-\; \Gamma^k_{ij} \partial_k W$$

The bosonic potential, which is of the form $\sum_i |\partial_i W|^2$, is sometimes known as a *F-term*. The action possesses the supersymmetry transformations:

$$\begin{array}{rcl}
\delta \phi^i & = & i\alpha_- \psi_+^i \;+\; i\alpha_+ \psi_-^i \\
\delta \phi^{\overline{\imath}} & = & i\tilde{\alpha}_- \psi_+^{\overline{\imath}} \;+\; i\tilde{\alpha}_+ \psi_-^{\overline{\imath}} \\
\delta \psi_+^i & = & -\tilde{\alpha}_- \partial \phi^i \;-\; i\alpha_+ \psi_-^j \Gamma^i_{jm} \psi_+^m \;+\; \alpha_+ g^{i\overline{\jmath}} \partial_{\overline{\jmath}} \overline{W} \\
\delta \psi_+^{\overline{\imath}} & = & -\alpha_- \partial \phi^{\overline{\imath}} \;-\; i\tilde{\alpha}_+ \psi_-^{\overline{\jmath}} \Gamma^{\overline{\imath}}_{\overline{\jmath}\overline{m}} \psi_+^{\overline{m}} \;+\; \tilde{\alpha}_+ g^{\overline{\imath}j} \partial_j W \\
\delta \psi_-^i & = & -\tilde{\alpha}_+ \overline{\partial} \phi^i \;-\; i\alpha_- \psi_+^j \Gamma^i_{jm} \psi_-^m \;-\; \alpha_- g^{i\overline{\jmath}} \partial_{\overline{\jmath}} \overline{W} \\
\delta \psi_-^{\overline{\imath}} & = & -\alpha_+ \overline{\partial} \phi^{\overline{\imath}} \;-\; i\tilde{\alpha}_- \psi_+^{\overline{\jmath}} \Gamma^{\overline{\imath}}_{\overline{\jmath}\overline{m}} \psi_-^{\overline{m}} \;-\; \tilde{\alpha}_- g^{\overline{\imath}j} \partial_j W
\end{array}$$

The ordinary nonlinear sigma model is classically scale-invariant, but notice that the action above is not scale-invariant even classically when $W \neq 0$. This means that Landau-Ginzburg models are not themselves conformal field theories. However, we can use them to define conformal field theories by applying renormalization group flow – the endpoint of renormalization group flow is a (possibly trivial) conformal field theory.

In the physics literature, the name "Landau-Ginzburg model" is sometimes reserved for the special case that the target space $X = \mathbf{C}^n$ for some n. Other cases are then called "hybrid Landau-Ginzburg" models.

In principle, the A and B twists of nonlinear sigma models extend to these nonlinear sigma models with superpotential. We shall focus on the case of a B twist, as that is what is most directly pertinent to derived categories. The B twist for the theory with superpotential is defined by taking the fermions to be the sections of the same bundles as before. The scalar supersymmetry transformation parameters are $\tilde{\alpha}_+$ and $\tilde{\alpha}_-$, as before. Taking $\tilde{\alpha}_+ = \tilde{\alpha}_- = \alpha$ as before, the BRST transformations are now

$$\begin{array}{rcl}
\delta \phi^i & = & 0 \\
\delta \phi^{\overline{\imath}} & = & i\alpha \eta^{\overline{\imath}} \\
\delta \eta^{\overline{\imath}} & = & 0 \\
\delta \theta_i & = & 2\alpha \partial_i W \\
\delta \rho^i & = & -\alpha d\phi^i
\end{array}$$

These are almost the same as for the ordinary B model, except that θ_i is no longer BRST-invariant.

The action for the nonlinear sigma model with a superpotential is not itself BRST exact, unlike the ordinary B model; however, it is still independent of the metric on the worldsheet

Σ, in the sense that rescaling the worldsheet metric is equivalent to adding BRST-exact terms to the action. (See [48] for more information.) Under such a rescaling $z \mapsto \lambda z$, the superpotential-dependent terms in the action become

$$\frac{1}{\alpha'} \int_{\Sigma} d^2 z \left(-\lambda^2 g^{i\overline{\jmath}} \partial_i W \partial_{\overline{\jmath}} \overline{W} \; + \; i \psi_+^i \psi_-^j D_i \partial_j W \; + \; i \lambda^2 \psi_+^{\overline{\imath}} \psi_-^{\overline{\jmath}} D_{\overline{\imath}} \partial_{\overline{\jmath}} \overline{W} \right)$$

in the B twisted theory (where we have used the fact that $\psi_{\pm}^i$ are worldsheet vectors and $\psi_{\pm}^{\overline{\imath}}$ are worldsheet scalars). Since λ is arbitrary, we can take a large λ limit from which we see that contributions to correlation functions will arise from fields ϕ such that $\partial W = 0$, which (since W is holomorphic), is equivalent to configurations such that $dW = 0$.

Now, let us evaluate a correlation function

$$< F_1(\phi) \cdots F_k(\phi) >$$

in this theory, where the F_i are analytic functions of the ϕ. As usual in the B model, there will be no Feynman diagram corrections. Furthermore, also as usual in the B model, there will be no worldsheet instanton contributions from maps of degree $d \neq 0$, and as already argued, the degree $d = 0$ maps will be clustered around solutions of $dW = 0$.

For the moment, for simplicity, let us assume that the zeroes of dW are isolated points, and let us further assume that the target space $X = \mathbf{C}^n$, *i.e.* that we are considering a Landau-Ginzburg model in the classical sense.

The path integral reduces to ordinary and Grassmann integrals over bosonic and fermionic zero modes, and for large λ reduces to the product of an integral over bosonic zero modes (for the nonlinear sigma model on the target $X = \mathbf{C}^n$)

$$\int_X d\phi \exp \left(- \frac{A}{\alpha'} \sum_i | \lambda \partial_i W |^2 \right)$$

and an integral over the fermionic zero modes

$$\int \prod (\sqrt{\alpha'} d\psi_+^i)(\sqrt{\alpha'} d\psi_-^i) \exp \left(i \frac{1}{\alpha'} \int_{\Sigma} \psi_+^i \psi_-^j \partial_i \partial_j W \right) \int \prod d\psi_+^{\overline{\imath}} d\psi_-^{\overline{\imath}} \exp \left(i \frac{A}{\alpha'} \lambda^2 \psi_+^{\overline{\imath}} \psi_-^{\overline{\jmath}} \partial_{\overline{\imath}} \partial_{\overline{\jmath}} \overline{W} \right)$$

where A is the area of the worldsheet Σ. The $\psi_{\pm}^{\overline{\imath}}$ zero modes are holomorphic sections of the pullback of the tangent bundle of X, which is trivial, so there are dim X such zero modes. The $\psi_{\pm}^i$ zero modes are holomorphic sections of K tensored with the pullback of the tangent bundle of X. (Since only the $\psi_{\pm}^{\overline{\imath}}$ zero modes are constants, only in the second factor can the integral over the worldsheet Σ be done trivially to give a factor of the worldsheet area A.)

Now, let us evaluate the bosonic factor

$$\int_X d\phi \exp \left(- \frac{A}{\alpha'} \sum_i | \lambda \partial_i W |^2 \right)$$

We will argue in a moment that the method of steepest descent will give an *exact* answer for this integral, not just the leading order term in an asymptotic series, because we can make λ arbitrarily large. To see this, expand

$$\phi \, = \, \phi_0 \, + \, \delta\phi$$

where ϕ_0 is a constant map that solves $dW = 0$, and $\delta\phi$ is a perturbation to another (constant) map. Then we see that the potential term in the action can be expanded as

$$\frac{1}{\alpha'} \int_\Sigma d^2 z \left[-\lambda^2 g^{i\overline{\jmath}} (\partial_k \partial_i W)(\phi_0) (\partial_{\overline{k}} \partial_{\overline{\jmath}} \overline{W})(\phi_0) (\delta\phi^k)(\delta\phi^{\overline{k}}) \, + \, \mathcal{O}((\delta\phi)^3) \right]$$

Now, define

$$\tilde{\phi} \, = \, \frac{\lambda}{\sqrt{\alpha'}} \delta\phi$$

We see that the expansion of the potential term in the action now has the form

$$\int_\Sigma d^2 z \left[-g^{i\overline{\jmath}} (\partial_k \partial_i W)(\phi_0) (\partial_{\overline{k}} \partial_{\overline{i}} \overline{W})(\phi_0) \tilde{\phi}^k \tilde{\phi}^{\overline{k}} \, + \, \frac{\sqrt{\alpha'}}{\lambda} \mathcal{O}((\tilde{\phi})^3) \right]$$

This is closely analogous to the procedure we followed for defining perturbation theory for ordinary nonlinear sigma models, except here we absorb the ratio $\lambda/\sqrt{\alpha'}$ into $\delta\phi$, instead of $r/\sqrt{\alpha'}$. In any event, we see that the higher-order terms past the quadratic are suppressed by factors proportional to λ^{-1}, and since we can make λ arbitrarily large without changing correlation functions in the topological subsector of the theory, that means that the higher-order terms are suppressed, and the method of steepest descent will give an exact answer for this integral, not just the leading term in an approximation.

In any event, given the general analysis above, we can evaluate the bosonic zero mode integral

$$\int_X d\phi \exp\left(-\frac{A}{\alpha'} \sum_i |\lambda \partial_i W|^2 \right)$$

This is simply a multivariable Gaussian. Define

$$H \, = \, \det\left(\partial_i \partial_j W \right), \quad \overline{H} \, = \, \det\left(\partial_{\overline{\imath}} \partial_{\overline{\jmath}} \overline{W} \right)$$

then we can read off, more or less immediately, that

$$\int_X d\phi \exp\left(-\frac{A}{\alpha'} \sum_i |\lambda \partial_i W|^2 \right) \, = \, \sum_{dW=0} \pi^n \left(\frac{\alpha'}{A} \right)^n \lambda^{-2n} H^{-1} \overline{H}^{-1}$$

where the Hessians H, $\overline{H}$ are to be evaluated at each solution of $dW = 0$.

In passing, we should also mention that for a more general case, for a more general X in a sector of maps of degree d not necessarily zero, the bosonic contribution to the path integral would have the form

$$\int_{\mathcal{M}_d} \exp\left(-\frac{A}{\alpha'}\lambda^2|\nabla\tilde{W}|^2\right)$$

where $\mathcal{M}_d$ is a (compactified) moduli space of maps into the target space X and $\tilde{W}$ is a function on $\mathcal{M}_d$ induced by the holomorphic function W on X. This is irrelevant for the B-twisted theory, but, is very relevant for the A-twisted theory.

Next, let us evaluate the factor corresponding to fermionic zero modes:

$$\int \prod(\sqrt{\alpha'}d\psi_+^i)(\sqrt{\alpha'}d\psi_-^i)\exp\left(i\frac{1}{\alpha'}\int_\Sigma \psi_+^i\psi_-^j\partial_i\partial_j W\right)\int\prod d\psi_+^{\bar{\imath}}d\psi_-^{\bar{\imath}}\exp\left(i\frac{A}{\alpha'}\lambda^2\psi_+^{\bar{\imath}}\psi_-^{\bar{\jmath}}\partial_{\bar{\imath}}\partial_{\bar{\jmath}}\overline{W}\right)$$

The $\psi_\pm^{\bar{\imath}}$ zero modes are holomorphic sections of the pullback of the tangent bundle of X, which is trivial, so there are dim X such zero modes. The $\psi_\pm^i$ zero modes are holomorphic sections of K tensored with the pullback of the tangent bundle of X. The normalizing factors of $\sqrt{\alpha'}$ were outlined in an earlier footnote, and have the effect of maintaining the correct dimensions. Given that information, we can immediately evaluate[24]

$$\begin{aligned}\int\prod d\psi_+^{\bar{\imath}}d\psi_-^{\bar{\imath}}\exp\left(i\frac{A}{\alpha'}\lambda^2\psi_+^{\bar{\imath}}\psi_-^{\bar{\jmath}}\partial_{\bar{\imath}}\partial_{\bar{\jmath}}\overline{W}\right) &= \frac{1}{n!}\left(i\lambda^2\frac{A}{\alpha'}\right)^n\overline{H} \\ \int\prod(\sqrt{\alpha'}d\psi_+^i)(\sqrt{\alpha'}d\psi_-^i)\exp\left(i\frac{1}{\alpha'}\int_\Sigma\psi_+^i\psi_-^j\partial_i\partial_j W\right) &= \frac{1}{(ng)!}\left(\frac{i}{\alpha'}\right)^{ng}(\alpha')^{ng}H^g\end{aligned}$$

where g is the genus of the worldsheet Σ.

Putting this all together, we can finally evaluate correlation functions.

$$\begin{aligned}< F_1\cdots F_k > &= \sum_{dW=0}\frac{1}{n!(ng)!}\pi^n\left(\frac{\alpha'}{A}\right)^n\lambda^{-2n}H^{-1}\overline{H}^{-1}\left(i\lambda^2\frac{A}{\alpha'}\right)^n\overline{H}(i)^{ng}H^gF_1\cdots F_k \\ &= \sum_{dW=0}\frac{1}{n!(ng)!}\pi^n(i)^{n(g+1)}H^{g-1}F_1\cdots F_k\end{aligned}$$

where the F_i are analytic functions of the ϕ^i.

Note, as a consistency check, that the expression for the correlation function is independent of the worldsheet area A, the scaling factor λ, and also α', exactly as one would hope for this topological field theory, which is supposed to be independent of the worldsheet metric, α', and λ.

[24] Up to signs which depend upon the ordering of the Grassmann integrals. We shall not attempt to give a rigorous justification of the signs, as in any event any sign can be absorbed into a trivial rescaling of the original path integral.

We usually reabsorb factors of i, π, and $n!$ into the definition of the path integral, so as to write the result more cleanly as

$$< F_1 \cdots F_k > = \sum_{dW=0} \frac{F_1 \cdots F_k}{H^{1-g}}$$

As an example, let us compute correlation functions in the following Landau-Ginzburg model. Consider a Landau-Ginzburg theory defined as a nonlinear sigma model on $\mathbf{C}^n$ with superpotential

$$W(y_1, \cdots, y_n) \;=\; \exp(y_1) \;+\; \exp(y_2) \;+\; \cdots \;+\; \exp(y_n) \;+\; q \exp(-y_1 - y_2 - \cdots - y_n)$$

Define the quantity $\Pi = \exp(-y_1 - \cdots - y_n)$, then the classical vacua (solutions of $dW = 0$) are defined by

$$\exp(y_1) \;=\; \exp(y_2) \;=\; \cdots \;=\; \exp(y_n) \;=\; q\Pi$$

which implies that $(\exp(y_i))^{n+1} = q$ for all i. Now, genus zero correlation functions are given by

$$< F_1 \cdots F_k > = \sum_{dW=0} \frac{F_1 \cdots F_k}{H}$$

where $H = \det(\partial_i \partial_j W)$, and the F_j are analytic functions in the chiral superfields y_i. In the present case, it is straightforward to compute that $H = (n+1) q^n \Pi^n$. Since for vacua $\exp(y_1) = \cdots = \exp(y_n)$, let x denote any of the $\exp(y_i)$, then we have that

$$< x^m > = \sum_{dW=0} \frac{x^m}{(n+1) x^n}$$

using the fact that $x = q\Pi$ for vacua, and where the sum runs over x such that $x^{n+1} = q$, *i.e.* $(n+1)$th roots of q. This expression can only be nonvanishing when $m - n$ is divisible by $n+1$, thus the only nonvanishing correlation functions are

$$< x^n >, \quad < x^{2n+1} >, \quad < x^{3n+2} >, \quad \cdots$$

In particular, we find

$$\begin{aligned} < x^n > \;&=\; 1 \\ < x^{2n+1} > \;&=\; q \\ < x^{n+d(n+1)} > \;&=\; q^d \end{aligned}$$

The attentive reader will recognize that these are the same correlation function as we obtained in a different-looking theory, namely in the A-twisted nonlinear sigma model on $\mathbf{P}^n$. There, we also discussed how the correlation functions are encoded in the *quantum cohomology relation* $\omega^{n+1} = q$, which also appears to match the relation $x^{n+1} = q$ above for classical vacua in this Landau-Ginzburg model. However, we arrived at these correlation functions in two very different ways: here, they are classical correlation functions, all in a sector of maps of zero degree, whereas in the A model on $\mathbf{P}^n$, the source of these correlation functions was corrections from maps of nonzero degree – so-called "nonperturbative" effects.

This matching is not an accident – later we shall see that the A model on $\mathbf{P}^n$ and the particular Landau-Ginzburg model considered here are "mirror" to one another, which means the two quantum field theories are secretly equivalent to one another, via a duality that exchanges perturbative and nonperturbative effects.

4.2 Open strings and matrix factorization

4.2.1 The Warner problem

We saw previously that for a nonlinear sigma model on a Riemann surface with boundary, supersymmetry still holds even taking into account the boundary terms. For Landau-Ginzburg models, on the other hand, matters are not so simple, and the boundary terms that one picks up under supersymmetry transformations do not vanish by themselves.

To be specific, consider the Landau-Ginzburg model defined in the previous subsection, and assume that there are no Chan-Paton factors on the boundary to modify the boundary conditions. Under a supersymmetry transformation, one picks up the following total derivative terms:

$$\frac{1}{\alpha'}\int_{\Sigma} d^2z \left[\partial\left(-\frac{i}{2}\alpha_- \partial_{\bar{\jmath}} \overline{W} \psi_-^{\bar{\jmath}} \right) \, + \, \overline{\partial}\left(\frac{i}{2}\alpha_+ \partial_{\bar{\jmath}} \overline{W} \psi_+^{\bar{\jmath}} \right) \, + \, \partial\left(-\frac{i}{2}\tilde{\alpha}_- \partial_i W \psi_-^i \right) \, + \, \overline{\partial}\left(\frac{i}{2}\tilde{\alpha}_+ \partial_i W \psi_+^i \right) \right] \tag{10}$$

If we take Σ to be the upper half-plane for simplicity, so that

$$\int_{\Sigma} d^2z \partial \; = \; \frac{1}{2i}\int_{\partial\Sigma} dx, \quad \int_{\Sigma} d^2z \overline{\partial} \; = \; -\frac{1}{2i}\int_{\partial\Sigma} dx$$

then we see the total derivative terms above become

$$\begin{aligned} \frac{1}{\alpha'}\frac{1}{2i}\int_{\partial\Sigma} dx \left[-\frac{i}{2}\alpha \partial_{\bar{\imath}} \overline{W} \psi_-^{\bar{\imath}} \, - \, \frac{i}{2}\alpha \partial_{\bar{\imath}} \overline{W} \psi_+^{\bar{\imath}} \right] \; &= \; -\frac{1}{\alpha'}\frac{1}{4}\int_{\partial\Sigma} dx \left[\alpha \partial_{\bar{\imath}} \overline{W} \left(\psi_+^{\bar{\imath}} \, + \, \psi_-^{\bar{\imath}} \right) \right] \\ \frac{1}{\alpha'}\frac{1}{2i}\int_{\partial\Sigma} dx \left[-\frac{i}{2}\tilde{\alpha} \partial_i W \psi_-^i \, - \, \frac{i}{2}\tilde{\alpha} \partial_i W \psi_+^i \right] \; &= \; -\frac{1}{\alpha'}\frac{1}{4}\int_{\partial\Sigma} dx \left[\tilde{\alpha} \partial_i W \left(\psi_+^i \, + \, \psi_-^i \right) \right] \end{aligned}$$

where we have defined $\alpha = \alpha_- = \alpha_+$, $\tilde{\alpha} = \tilde{\alpha}_+ = \tilde{\alpha}_-$, using an identity that exists for both Dirichlet and Neumann boundary conditions.

In the special case of Dirichlet boundary conditions, $\psi_+^\mu = -\psi_-^\mu$, so we see the terms above cancel out. However, for Neumann boundary conditions, $\psi_+^\mu = +\psi_-^\mu$, and so the terms above do not cancel out.

Thus, supersymmetry fails to be a symmetry of Landau-Ginzburg models defined on Riemann surfaces with boundary, with Neumann boundary conditions, unless possibly we add something else to the boundary to cancel out the undesired variations. This is known as the *Warner problem* [49], named after the individual who first discovered it.

4.2.2 Matrix factorization

In order to resolve the Warner problem, we are going to need for the open string to connect a D-brane/antibrane pair, with a nonvanishing pair of tachyons between them. The supersymmetry variations of the tachyon terms will cancel out the Warner problem terms above, so long as the product of the tachyons is the superpotential W.

Previously in section 3.4 we briefly outlined how one would describe a tachyon in terms of part of a boundary action. Here, we are going to give a more complete answer and describe other terms in the boundary action. For simplicity, we shall assume that Chan-Paton bundles are trivial, with vanishing connections. The boundary action is then

$$\begin{aligned} -\frac{1}{4\alpha'}\int_{\partial\Sigma} dx \Big[& h_{\alpha\overline{\alpha}} h_{a\overline{a}} \overline{\eta}^{\overline{\alpha a}} d\eta^{\alpha a} \;+\; i\psi^i \left(\partial_i F_{\alpha a}\right) \eta^{\alpha a} \;+\; i\psi^{\overline{\imath}} \left(\partial_{\overline{\imath}} \overline{F}_{\overline{\alpha a}}\right) \overline{\eta}^{\overline{\alpha a}} \\ & + i\psi^i \left(\partial_i G^{\alpha a}\right) \overline{\eta}^{\overline{\alpha a}} h_{\alpha\overline{\alpha}} h_{a\overline{a}} \;+\; i\psi^{\overline{\imath}} \left(\partial_{\overline{\imath}} \overline{G}^{\overline{\alpha a}}\right) \eta^{\alpha a} h_{\alpha\overline{\alpha}} h_{a\overline{a}} \\ & - iF_{\alpha a} \overline{F}_{\overline{\alpha a}} h^{\alpha\overline{\alpha}} h^{a\overline{a}} \;-\; G^{\alpha a} \overline{G}^{\overline{\alpha a}} h_{\alpha\overline{\alpha}} h_{a\overline{a}} \Big] \end{aligned}$$

where $\psi^i = \psi^i_+ + \psi^i_-$, $\psi^{\overline{\imath}} = \psi^{\overline{\imath}}_+ + \psi^{\overline{\imath}}_-$, and η, $\overline{\eta}$ are fermions that only live along the boundary $\partial\Sigma$, known as *boundary fermions*. If we let $\mathcal{E}$, $\mathcal{F}$ denote the two (assumed trivial) holomorphic vector bundles appearing along the boundary, then $h_{\alpha\overline{\alpha}}$, $h_{a\overline{a}}$, respectively, are their hermitian fiber metrics (which we have assumed constant in stating that the connections vanish). The boundary fermions η, $\overline{\eta}$ couple to $\mathcal{E}^\vee \otimes \mathcal{F}$ and $\mathcal{E} \otimes \mathcal{F}^\vee$, respectively (which is slightly obscured by our notation). The fields $F_{\alpha a}$, $G^{\alpha a}$ are holomorphic sections of $\mathcal{E}^\vee \otimes \mathcal{F}$ and $\mathcal{E} \otimes \mathcal{F}^\vee$. F and G are the two tachyons mentioned earlier, connecting the brane to the anti-brane.

Take the supersymmetry variations of ϕ, ψ along the boundary to be the restriction to the boundary of the bulk supersymmetry transformations, and take the boundary fermions η, $\overline{\eta}$ to have supersymmetry variations

$$\begin{aligned} \delta\eta^{\alpha a} &= -ih^{\alpha\overline{\alpha}} h^{a\overline{a}} \overline{F}_{\overline{\alpha a}} \alpha \;-\; iG^{\alpha a} \tilde{\alpha} \\ \delta\overline{\eta}^{\overline{\alpha a}} &= -ih^{\alpha\overline{\alpha}} h^{a\overline{a}} F_{\alpha a} \tilde{\alpha} \;-\; i\overline{G}^{\overline{\alpha a}} \alpha \end{aligned}$$

then the supersymmetry variation of the boundary action above is given by

$$-\frac{1}{4\alpha'}\int_{\partial\Sigma} dx \left[-\alpha\psi^{\overline{\imath}} \partial_{\overline{\imath}} \left(\overline{F}_{\overline{\alpha a}} \overline{G}^{\overline{\alpha a}} \right) \;-\; \tilde{\alpha}\psi^i \partial_i \left(F_{\alpha a} G^{\alpha a} \right) \right] \tag{11}$$

Comparing to equation (10), it is easy to see that the Warner problem will be solved, the total boundary term in the supersymmetry variations will vanish, if we choose F, G such that

$$F_{\alpha a} G^{\alpha a} \; = \; W$$

This is the solution to the Warner problem: to introduce two bundles $\mathcal{E}$, $\mathcal{F}$, living on the submanifold defined by the Dirichlet boundary conditions, together with maps $F : \mathcal{E} \longrightarrow \mathcal{F}$ and $G : \mathcal{F} \longrightarrow \mathcal{E}$ such that $F \circ G = W\mathrm{Id}$. Since F and G are matrix-valued functions, this procedure is known as *matrix factorization*.

Each such matrix factorization defines a D-brane in the Landau-Ginzburg model.

We have only discussed the details for the special case that the bundles $\mathcal{E}$, $\mathcal{F}$ have trivial gauge field, but the results extend to more general cases. The only constraint that emerges more generally is that the ranks of $\mathcal{E}$ and $\mathcal{F}$ must match.

4.2.3 Orlov's theorem

To understand the motivation for some theorems of Orlov, let us first describe something we will examine more thoroughly later in section 5.3. Specifically, certain nonlinear sigma models and Landau-Ginzburg models are closely related. For example, a nonlinear sigma model on a degree d hypersurface $\{G = 0\}$ in $\mathbf{P}^n$ lies on the same Kähler moduli space as a $\mathbf{Z}_d$ orbifold of a Landau-Ginzburg model on $\mathbf{C}^{n+1}$ with superpotential $W = G$. Hopefully it should look reasonable to the reader that they are closely related (albeit not precisely the same); the precise reasons why they are on the same Kähler moduli space, will be described in section 5.3.

Now, let us consider the B twisted nonlinear sigma model, and assume in the example above that the degree $d = n + 1$, so that the hypersurface is Calabi-Yau (and hence the B model is well-defined). Since the B model is independent of Kähler moduli, one would expect that the open string sector of the nonlinear sigma model should be the same as that of the Landau-Ginzburg model. In other words, there should be an isomorphism between the derived category of the hypersurface, and some sort of $\mathbf{Z}_d$-equivariant derived category built from matrix factorizations in the Landau-Ginzburg model.

Such an equivalence does exist, and was proven by Orlov [50, 51] (and anticipated in [52]).

Without working through all of the ideas of those papers, let us try to briefly outline a dictionary in some simple cases. Start with a matrix factorization defined by locally-free sheaves $\mathcal{E}$, $\mathcal{F}$ over $\mathbf{C}^{n+1}$, with maps $F : \mathcal{E} \rightarrow \mathcal{F}$ and $G : \mathcal{F} \rightarrow \mathcal{E}$ such that $FG = GF = WI$ for I the identity. The combination $\mathcal{G} \equiv \mathcal{E} \oplus \mathcal{F}$ defines a single locally-free sheaf, with an automorphism Q defined by F and G. The automorphism Q is nearly nilpotent, modulo the action of W. We shall assume that the module defining $\mathcal{G}$ is integrally graded[25] Thanks to the grading, $\mathcal{G}$ defines a sheaf over a projective space, which we can restrict to the hypersurface $\{W = 0\}$. After restriction, the automorphism Q becomes nilpotent, and we can take its cohomology to get a sheaf on the hypersurface.

If $\mathcal{P}$ denotes a matrix factorization defined by (P_0, P_1, p_0, p_1) as above, where $p_0 p_1 = p_1 p_0 = WI$, and P_0, P_1 are locally-free sheaves of equal rank, and $\mathcal{Q}$ is another matrix factorization (Q_0, Q_1, q_0, q_1) over the same underlying space, then a morphism between the two matrix factorizations is a pair of maps $f_0 : P_0 \to Q_0$ and $f_1 : P_1 \to Q_1$ making all the obvious diagrams commute. Such a map is null-homotopic if there exist maps $s_0 : P_0 \to Q_1$, $s_1 : P_1 \to Q_0$ such that

$$f_0 \;=\; s_1 p_0 \;+\; q_1 s_0, \quad f_1 \;=\; q_0 s_1 \;+\; s_0 p_1$$

4.3 Exercises

1. Check that in the special case of a constant target-space metric, the supersymmetry transformations of the nonlinear sigma model with superpotential W close.
2. Check that for constant target-space metric $g_{i\bar{\jmath}}$, supersymmetry transformations of a Landau-Ginzburg model on a Riemann surface Σ with boundary give rise to the nonzero boundary terms in equation (10).

5 Gauged linear sigma models

Gauged linear sigma models are families of non-conformal quantum field theories. They define conformal field theories by virtue of renormalization group flow – given any[26] quantum field theory, we can apply the renormalization group to generate a succession of different quantum field theories whose endpoint is a conformally-invariant theory, a conformal field theory.

Gauged linear sigma models have been extraordinarily successful in the string theory community, in giving not only global information about moduli spaces of conformal field theories that would be difficult to extract otherwise, but also enabling numerous other computations.

In this section we shall briefly describe gauged linear sigma models and some of the insight they provide.

[25] Technically, for a closed-string Landau-Ginzburg model over a vector space to renormalization-group-flow to a nontrivial CFT, a necessary condition is that the superpotential be quasi-homogeneous. The grading assumption here is the extension of that quasi-homogeneity of the closed string theory to the boundary theory, combined with a specification of a finite group action.

[26] Of the form we consider in these notes, at least.

5.1 The $\mathbf{CP}^n$ model

The prototype for all gauged linear sigma models is the $\mathbf{CP}^n$ model, which gives an alternative description of the nonlinear sigma model on $\mathbf{P}^n$ (in the sense that both the nonlinear sigma model and the corresponding gauged linear sigma model have the same fixed point under renormalization group flow).

> The attentive reader will recall that since $\mathbf{P}^n$ is not Calabi-Yau, the nonlinear sigma model on it does not define a nontrivial conformal field theory – the beta function is nonzero, and under the renormalization group, the metric on $\mathbf{P}^n$ will shrink. The fixed point of renormalization group flow is believed to be a trivial conformal field theory, equivalent to a nonlinear sigma model on $\chi(\mathbf{P}^n) = n + 1$ points. Nonetheless, we can still write down a corresponding gauged linear sigma model, and it is the simplest gauged linear sigma model known.

The gauged linear sigma model for $\mathbf{P}^n$ revolves around the fact that $\mathbf{P}^n$ has three equivalent descriptions:

1. It can be described with homogeneous coordinates $[z_0, \cdots, z_n]$, not all simultaneously zero, obeying the equivalence relation
$$[z_0, \cdots, z_n] \sim [\lambda z_0, \cdots, \lambda z_n]$$
for $\lambda \in \mathbf{C}^\times$.
2. It can be described as the GIT quotient $\mathbf{C}^{n+1}//\mathbf{C}^\times$.
3. It can be described as the symplectic reduction $\mu^{-1}(r)/U(1) = S^{2n+1}/U(1)$, where $\mu : \mathbf{C}^{n+1} \to \mathbf{R}$ is a moment map.

The gauged linear sigma model for $\mathbf{P}^n$ contains $n + 1$ sets of bosonic fields and fermionic superpartners, one for each homogeneous coordinate on $\mathbf{P}^{n+1}$, together with a 'gauged' $U(1)$. A 'gauged' group is a group G together with an action on the fields of the quantum field theory, in which we re-interpret the fields as sections of appropriate bundles associated to principal G-bundles with connection. For example, to 'gauge' the $U(1)$ action on a complex boson ϕ given by
$$\phi \mapsto \exp(i\alpha)\phi$$
we replace the original kinetic term in the action
$$\partial_\mu \overline{\partial} \partial^\mu \phi$$

with a new kinetic term in which ordinary derivatives are replaced by 'covariant' derivatives:

$$D_\mu \overline{\phi} D^\mu \phi$$

where

$$D_\mu \phi \,=\, \left(\partial_\mu \,+\, i A_\mu \right) \phi$$

where A_μ is a connection on a principal G bundle. In the path integral for the theory, we now sum over equivalence classes of principal G bundles over spacetime, as well as perform a functional integral over connections A_μ for any fixed choice of principal G bundle.

The effect of this procedure is that the functional integral over ϕ fields will end up integrating only over equivalence classes of ϕ fields, equivalence classes with respect to the $U(1)$ action, instead of all possible ϕ fields. (To completely explain and justify this is beyond the scope of these notes, but hopefully this description should suffice to enable the reader to understand the basic idea.)

Now, let us return to the $\mathbf{P}^n$ model. The gauged linear sigma model for this space is defined by $n+1$ sets of bosonic fields ϕ^i and superpartners $\psi^i_\pm$, together with a gauged[27] $U(1)$.

The complete action for this theory is rather complicated. For completeness, we give it here:

$$\begin{aligned}
\frac{1}{\alpha'} \int_\Sigma \Bigg[& \sum_i \left(-D_\mu \overline{\phi}^i D^\mu \phi^i \,+\, i \overline{\psi}^i_- D_z \psi^i_- \,+\, i \overline{\psi}^i_+ D_{\overline{z}} \psi^i_+ \right. \\
& - 2 |\sigma|^2 |\phi^i|^2 - \sqrt{2} \left(\overline{\sigma} \overline{\psi}^i_+ \psi^i_- \,+\, \sigma \overline{\psi}^i_- \psi^i_+ \right) \\
& \left. -\, i \sqrt{2} \overline{\phi}^i \left(\psi^i_- \lambda_+ \,-\, \psi^i_+ \lambda_- \right) \,-\, i \sqrt{2} \phi_i \left(\overline{\lambda}_- \overline{\psi}^i_+ \,-\, \overline{\lambda}_+ \overline{\psi}^i_- \right) \right) \\
& - \left(\sum_i |\phi^i|^2 \,-\, r \right)^2 \\
& + \frac{1}{2} F_{\mu\nu} F^{\mu\nu} \,+\, i \overline{\lambda}_+ \overline{\partial} \lambda_+ \,+\, i \overline{\lambda}_- \partial \lambda_- \,-\, \partial_\mu \overline{\sigma} \partial^\mu \sigma \Bigg]
\end{aligned}$$

(More general expressions can be found in [53].) This theory is called *gauged* because we have gauged a $U(1)$ group action. It is called a *linear* sigma model, rather than a nonlinear sigma model, because the metric in the kinetic terms is trivial, *i.e.* because the kinetic terms have the form

$$D_\mu \phi^{\overline{\imath}} D^\mu \phi^i$$

rather than

$$g_{i\overline{\jmath}}(\phi) D_\mu \phi^{\overline{\jmath}} D^\mu \phi^i$$

The bosonic fields σ and fermionic fields $\lambda_\pm$ were introduced to preserve supersymmetry in the presence of the gauge field A; for this reason, the $\lambda_\pm$ are sometimes known as *gauginos*.

[27] Why not gauge $\mathbf{C}^\times$ instead of $U(1)$? The answer is that we only understand how to gauge compact Lie groups, and in any event, we will see that physics realizes the symplectic reduction description of the projective space.

Given our discussion of quantum field theory at the beginning of week 1, this action may look formidable. In fact, with a bit of experience, it is easy to pick off some basic aspects. (Though, understanding all quantum effects in general gauged linear sigma models is an ongoing program to this day.) Because of the length of the action, however, we will usually avoid writing down the complete action in examples, and instead focus only on pertinent terms in the action.

One thing we can pick off immediately is the physical realization of the moment map in the symplectic reduction. The moment map $\mu : \mathbf{C}^{n+1} \to \mathbf{R}$ is given by

$$\mu(\phi^i) \;=\; \sum_i |\phi^i|^2$$

and so the points in $\mu^{-1}(r)$ are the solutions of

$$\sum_i |\phi^i|^2 \;=\; r$$

By comparison, notice that in the action above, we see a potential term

$$- \left(\sum_i |\phi^i|^2 \;-\; r \right)^2$$

Now, as stated earlier, we are interested in the low-energy behavior of the gauged linear sigma model: its renormalization-group-endpoints define the theories that we will be interested in. If we only wish to consider low-energy fluctuations, then we should treat potentials as if they were constraints, because we should stay close to the bottom of the potential. Thus, from the action above, we see that low-energy fluctuations obey the constraint

$$\sum_i |\phi^i|^2 \;=\; r$$

which defines an $S^{2n+1} \subset \mathbf{C}^{n+1}$, of radius $\sqrt{r}$. Furthermore, because we have gauged a $U(1)$ symmetry, we should only consider $U(1)$-equivalence classes of ϕ's, not ϕ's themselves, and so the low-energy degrees of freedom are really seeing

$$\left\{ \sum_i |\phi^i|^2 \;=\; r \right\} / U(1) \;=\; S^{2n+1}/U(1) \;=\; \mathbf{P}^n$$

This is how this particular quantum field theory realizes $\mathbf{P}^n$.

More generally, one can imagine a less symmetric $U(1)$ group action, in which the various ϕ fields have different weights. Physically, the weights of the $U(1)$ action are known as *charges* in the $U(1)$ gauge theory, and typically denoted Q_i. In general, one can write down a gauged linear sigma model with a variety of gauged $U(1)$'s each with a variety of different weights on the fields of the theory. For each gauged $U(1)$, there is a bosonic potential term of the form

$$\left(\sum_i Q_{i\alpha} |\phi^i|^2 \; - \; r_\alpha \right)^2$$

where $Q_{i\alpha}$ is the charge of the ith field with respect to the αth $U(1)$. Each parameter r_α is known as a *Fayet-Iliopoulos parameter*. They are real numbers that provides coordinates[28] on the Kähler moduli space. As before, the low-energy degrees of freedom are constrainted to lie at the bottom of each of these potentials; effectively, each such potential term defines a moment map along which we are performing a symplectic reduction. This potential, and the corresponding constraint on low-energy degrees of freedom, is known as a *D-term*, and there is one for each gauged $U(1)$.

For another easy example, suppose we wish to describe the weighted projective space $\mathbf{P}^n_{w_0,\cdots,w_n}$, defined by homogeneous coordinates $[z_0, \cdots, z_n]$ obeying the equivalence relation

$$[z_0, \cdots, z_n] \; \sim \; [\lambda^{w_0} z_0, \cdots, \lambda^{w_n} z_n]$$

We could do this with a gauged linear sigma model very similar to the one above for $\mathbf{P}^n$, except that the field ϕ^i would have $U(1)$ charge w_i rather than 1. The D-term constraint would then be

$$\sum_i w_i |\phi^i|^2 \; = \; r$$

In this fashion, it is possible to realize more general toric varieties, as we shall discuss later.

Advanced topic: The physical description of stacks. Suppose we start with the $\mathbf{P}^n$ model, but let the gauged $U(1)$ rotate each ϕ^i k times rather than once. This would correspond mathematically to something with homogeneous coordinates $[z_0, \cdots, z_n]$ obeying the equivalence relation

$$[z_0, \cdots, z_n] \; \sim \; [\lambda^k z_0, \cdots, \lambda^k z_n]$$

As an algebraic variety, this is the same as $\mathbf{P}^n$; however, as a stack, this is a $\mathbf{Z}_k$ gerbe over $\mathbf{P}^n$, and not $\mathbf{P}^n$ itself. One natural question to then ask is, which does the corresponding gauged linear sigma model (in which the ϕ^i fields all have charge k rather than charge 1) see, the variety or the stack? To further confuse matters, the conventional wisdom regarding $U(1)$ gauge theories with nonminimal charges was, for many years in many quarters, that they should be identical to $U(1)$ gauge theories with minimal charges – meaning here that the physical gauged linear sigma model would see the variety $\mathbf{P}^n$, and not the

[28] Strictly speaking, they only provide coordinates in the case that the GLSM is describing a Calabi-Yau. As we shall see later, for non-Calabi-Yau cases, such as $\mathbf{P}^n$, they vary as a function of cutoff Λ.

stack. However, in more recent years, a better understanding of such noneffective group actions has entered the physics literature, and nowadays we understand that physics actually sees the stack, not the variety, though the difference is a bit subtle. (Perturbatively, the two quantum field theories really are identical, the only difference lies in nonperturbative effects.) Put another way, if we start with the $\mathbf{P}^n$ model and give the fields all charge k rather than charge 1, the resulting gauged linear sigma model describes a $\mathbf{Z}_k$ gerbe on $\mathbf{P}^n$, instead of $\mathbf{P}^n$ itself. This is discussed in more detail in [14], for example. More generally, noneffective group actions are starting to also become important in other parts of physics – they play an important role in the recent physical understanding of the geometric Langlands program [54], for example.

We should also say something about how the gauged linear sigma model behaves under the renormalization group. In particular, the parameter r gets a quantum correction, proportional to the sum of the charges of the fields:

$$\delta r \propto \sum_i Q_i$$

(More generally, when there are multiple $U(1)$'s, the r associated to the αth $U(1)$ gets a correction proportional to the sum of the charges of the fields under the αth $U(1)$.) In particular, that means that for the $\mathbf{P}^n$ gauged linear sigma model, r is not a true parameter, in the sense that it is not renormalization-group invariant, but rather it flows. We should have expected this – we argued, after all, that nonlinear sigma models get quantum corrections that change the size of the metric, so that positively-curved spaces shrink. The computation that r is not a renormalization-group invariant is closely related to the corresponding fact about nonlinear sigma models.

One other quick note before proceeding: although we have described just the GLSM for $\mathbf{P}^n$, there also exist twisted versions, non-conformal topological field theories based on the GLSM. The twisting is done almost exactly as for the nonlinear sigma model, and as we will not use it, we will not mention it further.

5.2 Hypersurfaces in $\mathbf{CP}^n$

So far we have described GLSM's for $\mathbf{P}^n$ and various variants. Next, let us consider how to describe a hypersurface in $\mathbf{P}^n$.

The basic idea will be to add a potential function whose zero locus in $\mathbf{P}^n$ will be the desired hypersurface. To be specific, consider $\mathbf{P}^4$ with hypersurface $\{G(\phi) = 0\}$ where G is a degree 5 polynomial. Following our earlier discussion of Landau-Ginzburg models, the reader might guess that one should add a 'superpotential' $W = G$. That's a good idea, and it's close, but it is not quite correct:

- Referring back to the discussion of Landau-Ginzburg models, note that the potential terms (the F-terms) are of the form $\sum_i |\partial_i W|^2$ not $|W|^2$. Following the procedure that the quantum field theory we want is defined by the low-energy degrees of freedom of the GLSM, that would mean in the present case that the low-energy degrees of freedom must obey the constraint $\partial_i G = 0$ for all i, whereas instead we want them to obey $G = 0$.

- Technically, in a gauge theory, the superpotential must be gauge-invariant – but if we take $W = G$, then if all the ϕ's have charge 1, then W will have charge 5, which is certainly not gauge-invariant.

To fix these problems, we first add a new set of fields, consisting of a bosonic field p of charge -5 together with corresponding fermions, to the theory. Then, we take the superpotential to be $W = pG$. This fixes the two problems discussed above:

- The F-term constraints on low-energy degrees of freedom are now given by

$$\begin{aligned} G &= 0 \\ p\partial_i G &= 0 \end{aligned}$$

  and $G = 0$ is exactly the condition we were looking for. These two conditions arise from the two bosonic potential terms

$$|G|^2 + \sum_i |p\partial_i G|^2$$

  where the first term arises from $|\partial W/\partial p|^2$ and the other terms arise from $|\partial W/\partial \phi^i|^2$.

- The superpotential is now gauge-invariant: the product pG has total charge 0, since G has charge 5 and p has charge -5.

In the next section, we will perform a more detailed analysis to understand precisely why this choice reproduces the hypersurface $\{G = 0\}$ in $\mathbf{P}^4$.

Advanced topic: Following this procedure, one can build complete intersections of hypersurfaces in arbitrary toric varieties – for homework, you will work through a few simple examples. It is not too difficult to also describe, for example, partial flag manifolds, by replacing the $U(1)$ gauge group with some higher-rank nonabelian Lie group, as well as complete intersections of hypersurfaces therein, by adding a suitable superpotential. However, for a long time it was thought that only complete intersections of hypersurfaces in toric varieties or partial flag manifolds or small variants thereof were possible. Recently, that has changed. By taking advantage of some rather subtle physics, it has more recently been demonstrated that other constructions, distinct from complete intersections, are also possible. The first example appeared in [15][section 12.2] and was later described in more detail in [17, 18]. This example realized a branched double cover, via a $\mathbf{Z}_2$ gerbe as in [15]. The second example appeared in [19], and constructed a Pfaffian variety via strong-coupling nonabelian gauge physics. In both cases, not only were the geometric realizations unusual, but also the various geometries appearing as Kähler phases were not birational to one another, another surprise for people thinking about GLSM's. It has more recently been conjectured [17, 18] that in these circumstances 'birational' should be replaced by 'homologically projective dual' in the sense of Kuznetsov [20, 21, 22].

5.3 The relationship between nonlinear sigma models and Landau-Ginzburg models

Although many Landau-Ginzburg models appear to be closely related to certain nonlinear sigma models, they are *not* precisely the same.

Let us examine a GLSM for a quintic hypersurface $\{G(\phi) = 0\}$ in $\mathbf{P}^4$. In addition to a nonlinear sigma model on the quintic, there is also a Landau-Ginzburg model on $\mathbf{C}^5$ with potential function $G(\phi)$, that looks like it ought to be closely related, as well as orbifoldds thereof. Examining the GLSM will illuminate the precise relationship.

The GLSM contains 5 fields ϕ_i each of charge 1 under a gauged $U(1)$, plus another field p of charge -5. As a result, it has a D-term condition given by

$$\sum_i |\phi_i|^2 \; - \; 5|p|^2 \; = \; r \tag{12}$$

for a real number r (the Fayet-Iliopoulos parameter).

The GLSM has a superpotential $W = pG(\phi)$, from which we get the F-term conditions

$$\begin{aligned} G &= 0 \qquad (13) \\ p\frac{\partial G}{\partial \phi_i} &= 0 \text{ for all } i \qquad (14) \end{aligned}$$

In the limit $r \gg 0$, from the D-term condition (12) we see that the ϕ_i cannot all vanish. If the hypersurface is smooth, then the only way the F-term conditions (13,14) can both be satisfied is if $G = 0$ and $p = 0$. (After all, the condition for a singularity in the hypersurface $\{G = 0\}$ is to have a simultaneous solution of $G = 0$ and $dG = 0$.) So, when $r \gg 0$, one of the fields p has effectively disappeared, and we are left at low energies with the fields of the $\mathbf{CP}^4$ model forced to lie along the hypersurface $\{G = 0\}$.

We conclude that the theory in the limit $r \gg 0$ is in the same universality class of renormalization group flow as a nonlinear sigma model on the quintic.

Next, let us examine the opposite limit, $r \ll 0$. The D-term condition (12) now implies that p cannot vanish, though all the fields ϕ_i are allowed to vanish. Since p cannot vanish, and the hypersurface $\{G = 0\} \subset \mathbf{P}^4$ is assumed to be smooth, the only way for the F-term constraints (13,14) to be satisfied is for the fields ϕ_i to all have vanishing value.

This theory appears to be very different from the nonlinear sigma model in the $r \gg 0$ limit. We are certainly not describing a nonlinear sigma model; rather, what we are describing is more nearly a Landau-Ginzburg model over the space $\mathbf{C}^5$ (corresponding to the ϕ_i) with superpotential $W = G(\phi_i)$.

In fact, there is one slight complication still. The $U(1)$ gauge symmetry is broken by the fact that p is nonzero, but it is not completely broken. Since p has charge 5, the $U(1)$ gauge symmetry is broken to a $\mathbf{Z}_5$ subgroup (under which p is neutral, since it has charge 5). A gauge theory with a discrete gauge group is an orbifold, so we see that the theory at the $r \ll 0$ limit is in the same universality class as a $\mathbf{Z}_5$ orbifold of a Landau-Ginzburg model on $\mathbf{C}^5$.

To review: the GLSM for the quintic describes a family of CFT's, obtained by renormalization group flow for various values of the Fayet-Iliopoulos parameter. When $r \gg 0$, those theories are nonlinear sigma models on the quintic. When $r \ll 0$, one obtains an orbifold of a Landau-Ginzburg model associated to the quintic.

In particular, a nonlinear sigma model on the quintic is *not* the same as an orbifold of a Landau-Ginzburg model associated to the quintic – they lie on different points on a moduli space of CFT's. There are other ways to see that they are different theories, but hopefully this example should help to illustrate their precise connection. (In fact, historically it was Witten's paper introducing gauged linear sigma models that first made the precise relationship between nonlinear sigma models and Landau-Ginzburg models clear.)

5.4 A flop

Let us work through a simple example of a flop, realized with a gauged linear sigma model. (This example first appeared in [53].) We are going to write down a GLSM that describes the two small resolutions of the singularity

$$\mathbf{C}[a, b, c, d]/(ab - cd)$$

Our GLSM will have 4 chiral superfields, call them x, y, u, v, and a single gauged $U(1)$, with charges

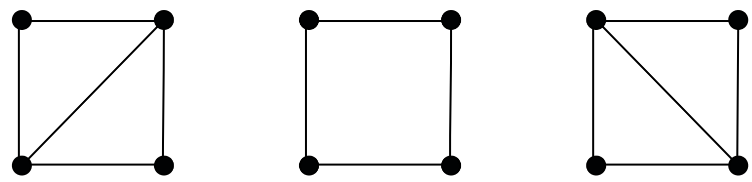


Figure 6: Cross section of fans for the two sides of the flop.

x	y	u	v
1	1	-1	-1

We will not add a superpotential. The D-term constraint is

$$|x|^2 \,+\, |y|^2 \,-\, |u|^2 \,-\, |v|^2 \,=\, r$$

When $r \gg 0$, x and y cannot both vanish. They form homogeneous coordinates on a $\mathbf{P}^1$, and the geometry is the total space of the bundle $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \rightarrow \mathbf{P}^1$.

When $r \ll 0$, u and v cannot both vanish. They form homogeneous coordinates on a $\mathbf{P}^1$, and the geometry is the total space of the bundle $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \rightarrow \mathbf{P}^1$.

For both $r \gg 0$ and $r \ll 0$, we have the same geometry on either side of the transition. Note that the transition is nontrivial – the excluded set changes. Moving in either direction, the excluded set of one phase is the zero section of the vector bundle in the other phase.

For completeness, we have illustrated cross-sections of the fans of the corresponding toric varieties in figure 6. Each of the two sides of the flop ($r \ll 0$, $r \gg 0$) have a fan, as does the singular point at $r = 0$. Each fan is a three-dimensional cone, and we have illustrated cross-sections through the plane $x = 1$. The corner points in each case, generating the edges of the fan, are at $(1, \pm 1, \pm 1)$. For the two fans in which $|r| \gg 0$, there are two separate coordinate charts, and so the fan has two maximal-dimension cones. The singular case $r = 0$ is described by a single coordinate chart.

5.5 Realization of toric varieties

Any toric variety can be described with a gauged linear sigma model, with gauge group $U(1)^n$ for some n. We can do this explicitly as follows. First, to each edge of the toric fan, we associate a chiral superfield. Next, we need to determine how many $U(1)$'s there are, and the charges of each chiral superfield. To do this, construct a matrix A whose rows are the coordinates of generators of each edge of the toric fan. Then, the number of $U(1)$'s in the corresponding GLSM is the kernel of A (with respect to left-multiplication), and the charges of the chiral superfields are determined by elements of that kernel.

Let us consider a simple example. The projective space $\mathbf{P}^N$ can be described by a toric variety with $N+1$ edges in N-dimensional space. N of those edges are the edges of the first octant in N-dimensional space, and the other points opposite the first octant. Thus, the matrix A is of the form

$$A = \left[\begin{array}{ccccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ -1 & -1 & -1 & \cdots & -1 \end{array} \right]$$

Now, construct a vector

$$v = [a_1, a_2, \cdots, a_N]$$

and solve for the a_i such that $vA = 0$. We get a sequence of relations:

$$\begin{array}{rcl} a_1 - a_N & = & 0 \\ a_2 - a_N & = & 0 \\ a_3 - a_N & = & 0 \\ \vdots & = & 0 \\ a_{N-1} - a_N & = & 0 \end{array}$$

This set of equations has a one-dimensional solution space, given by $a_1 = a_2 = a_3 = \cdots = a_N$, so the corresponding GLSM has a single gauged $U(1)$, and if we wish to describe the variety $\mathbf{P}^N$, then we make minimal integer choices of the a_i, in other words, we take $a_i = 1$ for all i.

As an aside, we can also describe toric stacks similarly [11, 14]. The primary distinction is that charges need not be minimal. Physically, a GLSM always sees the corresponding toric stack, so for example specifying a single $U(1)$ and $N+1$ chiral superfields each with charge k is distinct from a single $U(1)$ and $N+1$ chiral superfields each with charge 1 – the latter describes the variety $\mathbf{P}^N$, whereas the former describes a $\mathbf{Z}_k$ gerbe over $\mathbf{P}^N$. (The physics reason for the distinction is subtle, and wholly nonperturbative in nature. See for example [14] for details.)

5.6 Exercises

1. Describe GLSM's for the toric varieties $\mathbf{F}_n$, the Hirzebruch surfaces. Also explain (both in toric language and GLSM language) why Hirzebruch surfaces are $\mathbf{P}^1$ bundles over $\mathbf{P}^1$. A fan for $\mathbf{F}_n$ is shown in figure 7.

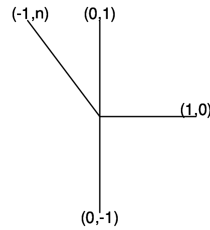


Figure 7: Fan for $\mathbf{F}_n$.

2. Check that the fans shown in figure 6 are described by the GLSM indicated in section 5.4.

3. Show that the one-loop renormalization of the 'parameter' r in the gauged linear sigma model for $\mathbf{P}^n$ is nonzero.

4. Describe the GLSM for a complete intersection of hypersurfaces in $\mathbf{P}^n$. How does the Calabi-Yau condition compare to the condition for the Fayet-Iliopoulos parameter r to not be renormalized?

5. Describe the GLSM for a hypersurface in a product of projective spaces. How does the Calabi-Yau condition compare to the condition for the Fayet-Iliopoulos parameter r to not be renormalized?

6 Mirror symmetry

6.1 Generalities

Mirror symmetry is a duality that exchanges pairs[29] of (possibly distinct) Calabi-Yau's manifolds of the same dimension. It can sometimes also be extended to more general spaces, exchanging projective spaces with certain Landau-Ginzburg models for example. In Calabi-Yau cases, two spaces X, Y are said to be mirror to one another if they are described by the same two-dimensional conformal field theory. This implies, for reasons we shall not attempt to explain here, that Dolbeault cohomology groups of the pair X, Y will be exchanged:

$$H^{p,q}(X) \;=\; H^{q,p}(Y)$$

In particular, Kähler moduli ($H^{1,1}$) are exchanged with complex moduli ($H^{2,1}$). Mirror symmetry also has the property of exchanging classical and quantum effects: sums over world-

[29] In low-dimensional cases, namely elliptic curves and K3's, there is a richer structure, but we shall not attempt to describe that here.

sheet instantons become classical computations in the dual theory. In terms of topological field theories, the mirror to the A model on X is the B model on Y, and the mirror to the A model on Y is the B model on X. (That last statement is true even for non-Calabi-Yau's.)

We shall not attempt to give a thorough history of mirror symmetry here, as the subject has a long history. We should note that one of the most useful constructions of mirrors involved exchanging polytopes derived from toric varieties, and is due to Batyrev [55]. However, computationally working out how to exchange polytopes can often be very cumbersome. The work we shall describe next has the virtue of being much easier to perform.

6.2 Hori-Vafa procedure

The papers[30] [8, 9, 10] describe an ansatz for constructing a Landau-Ginzburg model from any given gauged linear sigma model. As such, because it maps an entire gauged linear sigma model to a Landau-Ginzburg model, it should be noted at the beginning that they are not precisely constructing mirrors. For example, mirrors of typical nonlinear sigma models are other nonlinear sigma models, and not Landau-Ginzburg models, which are distinct theories. However, given a Landau-Ginzburg model, one can often construct a gauged linear sigma model containing that Landau-Ginzburg model at some special point in the moduli space, and so their construction serves as an existence argument for mirror symmetry. Furthermore, if instead of computing mirrors of full physical theories, if one is only interested in computing the mirror of the topological subsector of an A model topological field theory, then their construction will often determine the corresponding topological subsector of the B model of the mirror, basically because the B model is independent of Kähler moduli and so does not see the degrees of freedom that their construction discards.

The ansatz in question is relatively simple to describe. Suppose we have a gauged linear sigma model with some chiral superfields ϕ_i, obeying a collection of D-term relations of the form

$$\sum_i q_{\alpha i} |\phi_i|^2 \;=\; r_\alpha$$

for some integers $q_{\alpha i}$ and a real number r_α. For each field ϕ_i, the dual theory will contain a coordinate Y_i. If there are N chiral superfields ϕ_i, then the Landau-Ginzburg model in question is over a base X which is a compactification of $(\mathbf{C}^\times)^N$, with superpotential

$$W \;=\; \exp(-Y_1) \;+\; \cdots \;+\; \exp(-Y_N)$$

[30] Strictly speaking, the reference [8] attempted to construct dual gauged linear sigma models, rather than dualize a GLSM into a Landau-Ginzburg model. Their construction was not a complete success, however; the later work [9, 10] used the same technology to achieve very slightly different ends, namely a dualization of GLSM's into a Landau-Ginzburg model.

and a collection of relations (one for each D-term relation) each of the form

$$\sum_i q_{\alpha i} Y_i \;=\; r_\alpha$$

Determining which compactification X to take of $(\mathbf{C}^\times)^N$ is important to get the physics right, but is much more subtle. In these lectures, we shall simply state the answer in various examples, we shall not try to give a systematic description of the compactification procedure.

Let us first work through the example of $\mathbf{P}^n$. This is described as a GLSM by $n+1$ chiral superfields ϕ_i, each of charge 1, so that the D-term constraint is

$$\sum_i |\phi_i|^2 \;=\; r$$

Following the ansatz described above, the dual Landau-Ginzburg model will be defined over some $(n+1)$-dimensional space X, with superpotential

$$W \;=\; \exp(-Y_1) \;+\; \cdots \;+\; \exp(-Y_{n+1})$$

and relations

$$\sum_i Y_i \;=\; q$$

Let us use the relation to solve for Y_{n+1}:

$$Y_{n+1} \;=\; r \;-\; Y_1 \;-\; \cdots \;-\; Y_n$$

so we see the theory is effectively described by n variables $Y_1, \cdots, Y_n$, and a superpotential

$$W \;=\; \exp(-Y_1) \;+\; \cdots \;+\; \exp(-Y_n) \;+\; \exp(-r)\exp(Y_1 \;+\; \cdots \;+\; Y_n)$$

The attentive reader will note that this is the same Landau-Ginzburg model as that we discussed in section 4.1 defined with $y_i = -Y_i$ and $q = \exp(-r)$. In that section, we computed the (B model) correlation functions in this Landau-Ginzburg model, and discovered that they matched those of the A model on $\mathbf{P}^n$. Now we see why – the A model on $\mathbf{P}^n$ is mirror to the Landau-Ginzburg model above, and so their correlation functions must match.

Let us work through another example, namely the quintic hypersurface in $\mathbf{P}^4$. Here, the gauged linear sigma model has five fields ϕ_i of charge 1 and one field p of charge -5. The dual Landau-Ginzburg model has fields Y_i (corresponding to ϕ_i and Y_p (corresponding to p), with superpotential

$$W \;=\; \exp(-Y_1) \;+\; \cdots \;+\; \exp(= Y_5) \;+\; \exp(-Y_p)$$

and constraint

$$\sum_i Y_i \;-\; 5Y_p \;=\; r$$

Let us solve for Y_p and remove it from the superpotential:

$$Y_p \;=\; \frac{1}{5}\sum_i Y_i \;-\; r$$

so the superpotential becomes

$$W \;=\; \sum_i \exp(-Y_i) \;+\; \exp(r)\exp\left(-(1/5)\sum_i Y_i\right)$$

Now, for reasons we shall not try to explain, the correct thing to do at this point is to pick a particular compactification of the base. Define $x_i = \exp(-Y_i/5)$, then the superpotential becomes

$$W \;=\; x_1^5 \;+\; \cdots \;+\; x_5^5 \;+\; \psi x_1 x_2 x_3 x_4 x_5$$

where $\psi = \exp(r)$. Furthermore, we also need to perform a set of $\mathbf{Z}_5$ orbifolds (essentially because we defined x_i to be a fifth root of $\exp(-Y_i)$ – the ambiguity in fifth roots of unity is the origin of the $\mathbf{Z}_5$ orbifolds). The final result is that the 'mirror' to the quintic is a $\mathbf{Z}_5^{\oplus 4}$ orbifold of a Landau-Ginzburg model on $\mathbf{C}^5$ with superpotential given by a quintic polynomial. It turns out that this is precisely the Landau-Ginzburg point of the GLSM that is mirror to that of the original quintic, so the result is of the desired form.

6.3 Generalizations

One generalization of mirror symmetry that has been attempted is known as "(0,2) mirror symmetry." The "(0,2)" refers to the fact that it involves nonlinear sigma models that are generalizations of those discussed in these notes, with (0,2) supersymmetry instead of (2,2) supersymmetry. The reduced supersymmetry implies that the CFT is specified by a space X together with some holomorphic vector bundle $\mathcal{E}$ over X, obeying the constraint $\mathrm{ch}_2(\mathcal{E}) = \mathrm{ch}_2(TX)$. Here, (0,2) mirror symmetry is believed to exchange pairs $(X, \mathcal{E})$, $(Y, \mathcal{F})$, such that each pair defines the same conformal field theory. Rather than exchanging Dolbeault cohomology, instead sheaf cohomology is exchanged. Very little is known about (0,2) mirror symmetry at this point in time.

Another generalization is known as "homological mirror symmetry" [28]. Homological mirror symmetry was originally proposed by Kontsevich as a duality between derived categories of coherent sheaves on X and Fukaya categories on the (ordinary) mirror Y. The modern understanding of homological mirror symmetry is as the duality between D-branes implied by ordinary mirror symmetry, where derived categories arise from the open string B model and Fukaya categories from the open string A model. It is a tribute to Kontsevich's insight that he made this proposal *before* D-branes became widely known or accepted among physicists, and indeed, understanding the physical meaning of Kontsevich's proposal was one of the motivations behind the original papers on D-branes and derived categories [34].

6.4 Exercises

1. (Research problem) To what extent is the Hori-Vafa mirror ansatz independent of presentation? For example, $\mathbf{P}^1 = \mathbf{P}^2[2]$ and $\mathbf{P}^3[2] = \mathbf{P}^1 \times \mathbf{P}^1$; in those cases, how are the dual Landau-Ginzburg models one obtains by Hori-Vafa related?

A Asymptotic series

Asymptotic series play a crucial role in understanding quantum field theory, as Feynman diagram expansions are typically asymptotic series expansions. As I will occasionally refer to asymptotic series, I have included in this appendix some basic information on the subject.

See [57] sections 5.9, 5.10, 7.3 (7.4 in 5th edition), 8.3 (10.3 in 5th edition) for some of the material below.

A.1 Definition

By now as graduate students you have seen infinite series appear many times. However, in most of those appearances, you have probably made the assumption that the series converged, or that the series is only useful when convergent.

Asymptotic series are non-convergent series, that nevertheless can be made useful, and play an important role in physics. The infinite series one gets in quantum field theory by summing Feynman diagrams, for example, are asymptotic series.

To be precise, consider a function $f(z)$ with an expansion as

$$f(z) \;=\; A_0 \,+\, \frac{A_1}{z} \,+\, \frac{A_2}{z^2} \,+\, \cdots$$

where the A_i are numbers. We can think of the series $\sum A_i/z^i$ as approximating $f(z)/\varphi(z)$ for large values of z.

We say that the series $\sum A_i/z^i$ represents $f(z)$ *asymptotically* if, for a given n, the first n terms of the series may be made as close as desired to $f(z)$ by making z large enough, *i.e.*

$$\lim_{z\rightarrow\infty} z^n \left[f(z) \,-\, \sum_{p=0}^{n} \frac{A_p}{z^p} \right] \;=\; 0$$

Such series need not converge for large z; in fact, in typical cases of interest, an asymptotic series will never converge.

It is important to note that asymptotic series are distinct from convergent series: a convergent series need not be asymptotic. For example, consider the Taylor series for $\exp(z)$. This is a convergent power series, but the same power series does *not* define an asymptotic series for $\exp(z)$. After all,

$$\lim_{z\to\infty} z^n \left[\exp(z) \; - \; \sum_{p=0}^{n} \frac{z^p}{n!} \right] \;\to\; \infty$$

and so the series is not asymptotic to $\exp(z)$, though it does converge to $\exp(z)$.

Not all functions have an asymptotic expansion; $\exp(z)$ is one such function. If a function does have an asymptotic expansion, then that asymptotic expansion is unique. However, several different functions can have the same asymptotic expansion; the map from functions to asymptotic expansions is many-to-one, when it is well-defined.

Example: Consider the function

$$-\mathrm{Ei}(-x) \;=\; E_1(x) \;=\; \int_x^\infty \frac{\exp(-t)}{t} dt$$

We can generate a series approximating this function by a series of integrations by parts:

$$\begin{aligned} E_1(x) \;&=\; \frac{\exp(-x)}{x} \;-\; \int_x^\infty \frac{\exp(-t)}{t^2} dt \\ &=\; \frac{\exp(-x)}{x} \left[1 \;-\; \frac{1}{x} \;+\; \frac{2!}{x^2} \;-\; \frac{3!}{x^3} \;+\; \cdots \; \frac{(-)^n n!}{x^n} \right] \;+\; (-)^{n+1}(n+1)! \int_x^\infty \frac{\exp(-t)}{t^{n+2}} dt \end{aligned}$$

The series

$$\sum_{n=0}^{\infty} (-)^n \frac{n!}{x^n}$$

is not convergent, in any standard sense. For example, when $x = 1$, this is the series

$$1 \;-\; 2! \;+\; 3! \;-\; 4! \;+\; \cdots$$

More generally, for any fixed x the magnitude of the terms *increases* as n grows, so this alternating series necessarily diverges.

However, although the series diverges, it is asymptotic to $E_1(x) x \exp(x)$. To show this, we must prove that for fixed n,

$$\lim_{x\to\infty} x^n \left[E_1(x) \;-\; \frac{\exp(-x)}{x} \sum_{p=0}^{n} \frac{(-)^n n!}{x^n} \right] \;=\; 0$$

Using our expansion from successive integrations by parts, we have that the limit is given by

$$\lim_{x\to\infty} x^n \left[(-)^{n+1}(n+1)! \int_x^\infty \frac{\exp(-t)}{t^{n+2}} dt \right]$$

We can evaluate this limit using the fact that

$$\int_x^\infty \frac{\exp(-t)}{t^{n+2}} dt \; < \; \frac{1}{x^{n+2}} \int_x^\infty \exp(-t) dt \; = \; \frac{\exp(-x)}{x^{n+2}}$$

Since

$$\lim_{x \to \infty} \frac{x^n (n+1)! \exp(-x)}{x^{n+2}} \; = \; 0$$

we see that the series is asymptotic.

Example: Consider the ordinary differential equation

$$y' + y = \frac{1}{x}$$

The solutions of this ODE have an asymptotic expansion, as we shall now verify.

To begin, assume that the solutions have a power series expansion of the form

$$y(x) \; = \; \sum_{n=0}^\infty \frac{a_n}{x^n}$$

for some constants a_n. Plugging this ansatz into the differential equation above and solving for the coefficients, we find

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 1 \\ a_2 &= a_1 \; = \; 1 \\ a_3 &= 2a_2 \; = \; 2 \\ a_4 &= 3a_3 \; = \; 3! \\ a_5 &= 4a_4 \; = \; 4! \end{aligned}$$

and so forth, leading to the expression

$$y(x) \; = \; \sum_{n=1}^\infty \frac{(n-1)!}{x^n}$$

First, let us check convergence of this series. Apply the ratio test to find

$$\lim_{n \to \infty} \frac{n!/x^{n+1}}{(n-1)!/x^n} \; = \; \lim_{n \to \infty} \frac{n}{x} \; \longrightarrow \; \infty$$

In particular, by the ratio test, this series diverges for all x (strictly speaking, all x for which it is well-defined, *i.e.* all $x \neq 0$).

We can derive this asymptotic series in an alternate fashion, which will explain its close resemblance to the previous example. Recall the method of variation of parameters for solving inhomogeneous equations: first find the solutions of the associated homogeneous equations, then make an ansatz that the solution to the inhomogeneous equation is given by multiplying the solutions to the homogeneous solutions by functions of x. In the present case, the associated homogeneous equation is given by

$$y' + y = 0$$

which has solution $y(x) \propto \exp(-x)$. Following the method of variation of parameters, we make the ansatz

$$y(x) = A(x) \exp(-x)$$

for some function $A(x)$, and plug back into the (inhomogeneous) differential equation to solve for $A(x)$. In the present case, that yields

$$A' \exp(-x) = \frac{1}{x}$$

which we can solve as

$$A(x) = \int_{-\infty}^{x} \frac{\exp(t)}{t} dt$$

(Note that I am implicitly setting a value of the integration constant by setting a lower limit of integration. Also note that the integral above is ill-defined if x is positive, a matter I will gloss over for the purposes of this discussion.) Thus, the solution to the inhomogeneous equation is given by

$$y(x) = \exp(-x) \int_{-\infty}^{x} \frac{\exp(t)}{t} dt$$

whose resemblance to the previous example should now be obvious.

A.2 The gamma function and the Stirling series

An important example of an asymptotic series is the asymptotic series for the gamma function, known as the Stirling series. The gamma function is a meromorphic function on the complex plane that generalizes the factorial function. Denoted $\Gamma(z)$, it has the properties

$$\begin{aligned} \Gamma(z+1) &= z\Gamma(z) \\ \Gamma(1/2) &= \sqrt{\pi} \\ \Gamma(1) &= 1 \\ \Gamma(n+1) &= n! \text{ for } n \text{ a positive integer} \end{aligned}$$

Because of that last property, because the gamma function generalizes the factorial function, people sometimes define $z! \equiv \Gamma(z+1)$ for any complex number z. The gamma function has simple poles at $z = 0, -1, -2, -3, \cdots$. It also obeys numerous curious identities, including

$$\Gamma(z)\Gamma(1-z) \;=\; \frac{\pi}{\sin \pi z}$$

and "Legendre's duplication formula"

$$\Gamma(1+z)\Gamma(z+1/2) \;=\; 2^{-2z}\sqrt{\pi}\,\Gamma(2z+1)$$

The gamma function has several equivalent definitions. It can be expressed as an integral, using an expression due to Euler:

$$\Gamma(z) \;=\; \int_0^\infty \exp(-t) t^{z-1} dt, \quad \Re z > 0$$

It can also be expressed as a limit, using another expression also due to Euler:

$$\Gamma(z) \;=\; \lim_{n\rightarrow\infty} \frac{1\cdot 2\cdot 3 \cdots n}{z(z+1)(z+2)\cdots(z+n)} n^z, \quad n \neq 0, -1, -2, -3, \cdots$$

It can also be expressed as an infinite product, using an expression due to Weierstrass:

$$\frac{1}{\Gamma(z)} \;=\; z\exp(\gamma z) \prod_{n=1}^\infty \left(1 \,+\, \frac{z}{n} \right) \exp(-z/n)$$

where γ is the Euler-Mascheroni constant

$$\begin{aligned} \gamma &\equiv \lim_{n\rightarrow\infty} \left(\sum_{m=1}^n \frac{1}{m} \,-\, \log n \right) \\ &\approx 0.577215661901\cdots \end{aligned}$$

where we use log to denote the natural logarithm. (The regions of validity of each definition are slightly different; analytic continuation defines the function globally.)

We can also define the digamma and polygamma functions, which are various derivatives of the gamma function. The digamma function, denoted either ψ or F, is defined by

$$\psi(z+1) \;\equiv\; F(z) \;\equiv\; \frac{d}{dz} \log \Gamma(z+1)$$

It can be shown that

$$\psi(z+1) \;=\; -\gamma \,+\, \sum_{n=1}^\infty \frac{z}{n(n+z)}$$

The polygamma function $\psi^{(n)}$ is defined to be a higher-order derivative:

$$\psi^{(n)}(z+1) \;\equiv\; F^{(n)}(z) \;\equiv\; \frac{d^{n+1}}{dz^{n+1}} \log \Gamma(z+1)$$

The Stirling series for the gamma function is derived using the Euler-Maclaurin integration formula, which we shall digress briefly to explain. (See section 5.9 of [57] for this background, and section 8.3 on the resulting asymptotic series, known as the Stirling series.)

First, recall the Bernoulli numbers B_n are defined by the generating function

$$\frac{x}{\exp(x) - 1} \;=\; \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}$$

It is straightforward to compute that $B_0 = 1$, $B_1 = -1/2$, and $B_2 = 1/6$.

Next, the Bernoulli functions $B_n(s)$ are defined by the generating function

$$\frac{x\exp(xs)}{\exp(x) - 1} \;=\; \sum_{n=0}^{\infty} B_n(s) \frac{x^n}{n!}$$

so that, for example, $B_0(s) = 1$, $B_1(s) = s - 1/2$, and $B_2(s) = x^2 - x + 1/6$. It is trivial to see that the Bernoulli numbers and functions are related by $B_n(0) = B_n$. It is also easy to check that

$$\frac{d}{ds} B_n(s) \;=\; n B_{n-1}(s)$$

and that

$$B_n(1) \;=\; (-)^n B_n(0)$$

Now, we can derive the Euler-Maclaurin integration formula, which we will use to derive an asymptotic series for the gamma function. Consider the integral

$$\int_0^1 f(x) dx \;=\; \int_0^1 f(x) B_0(x) dx$$

The idea is to use the relation $B_n'(x) = n B_{n-1}(x)$ and successive integrations by parts to generate an infinite series. At the first step, use $B_1'(x) = B_0(x)$ and integrate by parts to get

$$\begin{aligned}
\int_0^1 f(x) dx &= \int_0^1 f(x) B_1'(x) dx \\
&= f(x) B_1(x)]_0^1 \;-\; \int_0^1 f'(x) B_1(x) dx \\
&= \frac{1}{2}\left(f(1) + f(0)\right) \;-\; \frac{1}{2}\int_0^1 f'(x) B_2'(x) dx \\
&= \frac{1}{2}\left(f(1) + f(0)\right) \;-\; \frac{1}{2}\left[f'(x) B_2(x)\right]_0^1 \;+\; \frac{1}{2}\int_0^1 f''(x) B_2(x) dx
\end{aligned}$$

and so forth. Continuing this process, we find

$$\begin{aligned}
\int_0^1 f(x) dx \;=\; & \frac{1}{2}\left[f(1) + f(0)\right] \;-\; \sum_{p=1}^{q} \frac{B_{2p}}{(2p)!}\left(f^{(2p-1)}(1) - f^{(2p-1)}(0) \right) \\
& + \frac{1}{(2q)!} \int_0^1 f^{(2q)}(x) B_{2q}(x) dx
\end{aligned}$$

where the B_{2p} in the sum above are the Bernoulli numbers, not the functions, and where we have used the relations

$$\begin{aligned} B_{2n}(1) &= B_{2n}(0) = B_{2n} \\ B_{2n+1}(1) &= B_{2n+1}(0) = 0 \end{aligned}$$

(Compare equation (5.168a) in [57].)

By replacing $f(x)$ with $f(x+1)$ we can shift the integration region from $[0,1]$ to $[1,2]$, and by adding up results for different regions, we finally get the Euler-Maclaurin integration formula:

$$\begin{aligned} \int_0^n f(x)dx &= \frac{1}{2}f(0) + f(1) + f(2) + \cdots + f(n-1) + \frac{1}{2}f(n) \\ &\quad - \sum_{p=1}^{q} \frac{B_{2p}}{(2p)!}\left(f^{(2p-1)}(n) - f^{(2p-1)}(0)\right) \\ &\quad + \frac{1}{(2q)!}\int_0^n B_{2q}(x)\sum_{s=0}^{n-1} f^{(2q)}(x+s)dx \end{aligned}$$

(Compare equation (5.168b) in [57].)

Apply the Euler-Maclaurin integration formula above to the right-hand side of the equation

$$\frac{1}{z} = \int_0^\infty \frac{1}{(z+x)^2}dx$$

to get the series

$$\begin{aligned} \frac{1}{z} &= \frac{1}{2}f(0) + \sum_{n=1}^{\infty} f(n) \\ &\quad - \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)!}\left(\lim_{x\to\infty} f^{(2n-1)}(x) - f^{(2n-1)}(0)\right) \end{aligned}$$

for $f(x) = (z+x)^{-2}$. Thus,

$$\frac{1}{z} = \frac{1}{2z^2} + \sum_{n=1}^{\infty} \frac{1}{(z+n)^2} - \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)!}\left(\frac{(2n)!}{z^{2n+1}}\right)$$

From equation (10.41) in [57],

$$\sum_{n=1}^{\infty} \frac{1}{(z+n)^2} = F^{(1)}(z)$$

so we can write

$$F^{(1)}(z) = \frac{d}{dz}F(z) = \frac{1}{z} - \frac{1}{2z^2} + \sum_{n=1}^{\infty} \frac{B_{2n}}{z^{2n+1}}$$

(Compare equation (10.50) in [57].) Integrating once we get

$$F(z) \;=\; C_1 \;+\; \log z \;+\; \frac{1}{2z} \;-\; \sum_{n=1}^{\infty} \frac{B_{2n}}{2n z^{2n}}$$

It can be shown (see [57]) that $C_1 = 0$. Since

$$F(z) \;=\; \frac{d}{dz} \log \Gamma(z+1)$$

we can integrate again to get

$$\log \Gamma(z+1) \;=\; C_2 \;+\; \left(z + \frac{1}{2}\right) \log z \;-\; z \;+\; \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)(2n-1) z^{2n-1}}$$

for some constant C_2, where we have used the fact that

$$\frac{d}{dz} z \,(\log z \;-\; 1) \;=\; \log z.$$

We can solve for C_2 by substituting the expression above into the Legendre duplication formula

$$\Gamma(z+1)\Gamma(z + \frac{1}{2}) \;=\; 2^{-2z} \pi^{1/2} \Gamma(2z+1)$$

(this corrects a minor typo in equation (10.53)) from which one can derive that $C_2 = (1/2)\log(2\pi)$. Thus,

$$\log \Gamma(z+1) \;=\; \frac{\log 2\pi}{2} \;+\; \left(z \;+\; \frac{1}{2}\right) \log z \;-\; z \;+\; \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)(2n-1) z^{2n-1}}$$

which is Stirling's series, an asymptotic series for the natural logarithm of the gamma function.

We can also derive a more commonly used expression for Stirling's series by exponentiating the series above. We get

$$\Gamma(z+1) \;=\; \sqrt{2\pi} z^{z+1/2} \exp(-z) \exp\left(\sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)(2n-1) z^{2n-1}} \right)$$

We can simplify the last factor as follows. Recall the Taylor expansion

$$\log\,(1 \;+\; x) \;=\; x \;-\; \frac{x^2}{2} \;+\; \frac{x^3}{3} \;-\; \frac{x^4}{4} \;+\; \cdots$$

If we find an x such that

$$\sum_{n=1}^{\infty} (-)^{n+1} \frac{x^n}{n} \;=\; \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)(2n-1) z^{2n-1}}$$

then we can write

$$\exp\left(\sum_{n=1}^{\infty}\frac{B_{2n}}{(2n)(2n-1)z^{2n-1}}\right) \;=\; 1+x$$

Although finding a closed-form expression is impossible, we can find a series in z for x. From the first terms, clearly

$$x \;=\; \frac{B_2}{2z} \;+\; \mathcal{O}(z^{-2})$$

and if we work out the expansion more systematically, we discover

$$\begin{aligned} x \;&=\; \frac{B_2}{2z} \;+\; \frac{B_2^2}{8z^2} \;+\; \mathcal{O}(z^{-3}) \\ &=\; \frac{1}{12z} \;+\; \frac{1}{288z^2} \;+\; \mathcal{O}(z^{-3}) \end{aligned}$$

Thus, we have that

$$\Gamma(z+1) \;=\; \sqrt{2\pi}z^{z+1/2}\exp(-z)\left(1 \;+\; \frac{1}{12z} \;+\; \frac{1}{288z^2} \;+\; \mathcal{O}(z^{-3})\right)$$

another form of Stirling's asymptotic series.

A.3 Method of steepest descent

Consider a contour integral of the form

$$G(z) \;=\; \int_C \exp(zf(t))dt$$

The *method of steepest descent* is a systematic procedure for generating an asymptotic series that approximates integrals of this form.

One application will be to the gamma function. Recall the definite integral description of the gamma function:

$$\Gamma(z+1) \;=\; \int_0^\infty \exp(-t)t^z dt$$

Change integration variables to $t = \tau z$:

$$\Gamma(z+1) \;=\; z^{z+1}\int_0^\infty \exp(-z\tau z)\tau^z d\tau$$

and write

$$\exp(-\tau/z)\tau^z \;=\; \exp(-z\tau + z\log\tau)$$

giving an integral of the form

$$\int_C \exp(zf(t))dt$$

with $f(t) = -t + \log t$.

Another application of the method of steepest descent is to the Feynman path integral description of quantum mechanics and quantum field theory, where it is used to recover the classical limit.

Now that we have described the setup, just what exactly is the method of steepest descent? The general idea is that in an integral over a complex exponential of the form $\exp(zf(t))$, for large z, the part of the integration contour that mostly just changes the phase will not significantly contribute to the integral, but rather will tend to cancel out. A little more systematically, if the integral involves integrating over all phases of the complex exponential, then the different contributions should sum to zero, on the grounds that

$$\int_0^{2\pi} \cos\theta d\theta \;=\; \int_0^{2\pi} \sin\theta d\theta \;=\; 0$$

At this same, level of approximation, the leading contribution to the integral should come from parts of the contour where the phase does not change significantly.

A little more mechanically, if the contour contains or passes near a spot where $\partial f(t)\partial t = 0$ (known as a 'saddle point'), and that is at $t = t_0$, then for large z the leading contribution to the contour integral

$$\int_C \exp(zf(t))dt$$

from that section of the contour should be proportional to $\exp(zf(t_0))$

The description we have given is very far from being a thorough argument, but it turns out to be essentially correct.

Let us describe how to derive this leading contribution in the case of the gamma function, then we shall describe how to more systematically use these general ideas to create an asymptotic expansion for such contour integrals.

In the case of the gamma function, recall

$$\Gamma(z+1) \;=\; z^{z+1}\int_0^\infty \exp(zf(t))dt$$

for $f(t) = \log t - t$. Solving $\partial f/\partial t = 0$, we find that the only possible saddle point is at

$$\frac{1}{t} - 1 \;=\; 0$$

i.e. $t_0 = 1$. Let us expand $f(t)$ about this saddle point. Write $t = 1 + x$, then

$$\begin{aligned} f(t) &= \log(1+x) \;-\; (1+x) \\ &= \left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \right) - (1+x) \\ &= -1 - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \end{aligned}$$

from which we see the leading order approximation to the integral should be proportional to $\exp(-z)$. We can now approximate the gamma function by

$$\begin{aligned}
\Gamma(z+1) &= z^{z+1}\exp(zf(1))\int_{-1}^{\infty}\exp(zf''(1)x^2/2!)dx \\
&= z^{z+1}\exp(-z)\int_{-1}^{\infty}\exp(-zx^2/2)dx \\
&\cong z^{z+1}\exp(-z)\int_{-\infty}^{\infty}\exp(-zx^2/2)dx \\
&= z^{z+1}\exp(-z)\sqrt{\frac{2\pi}{z}} \\
&= \sqrt{2\pi z}z^z\exp(-z)
\end{aligned}$$

which is the leading term in Stirling's expansion of the factorial function.

Reference [57], in section 7.4, discusses how to treat a contour integral in which the contour does not lie along the real line. An important part of the treatment, which I will omit here but which is discussed in [57], involves replacing the contour integral with an integral along an infinite line that is tangent to a particular direction at the saddle point.

Let us instead outline how to more systematically derive an asymptotic series, not just the leading term, using these methods. In terms of our original function $G(z)$, if our contour crosses a single saddle point t_0, and the contour line is along a "path of steepest descent" along which the imaginary part of $zf(t)$ is constant. Write

$$f(t) = f(t_0) - w^2$$

for some variable w, which is real Im $f(t) = \text{Im } f(t_0)$ everywhere along the contour. Then,

$$G(z) = \exp(zf(t_0))\int_C \exp(-zw^2)\left(\frac{dt}{dw}\right)dw$$

Assume that the contour C is such that the w integral can be taken to run over the real numbers from $-\infty$ to ∞. Next, we need to write dt/dw as a function of w, rather than t. In general, we can accomplish such an inversion at the power-series level, so write

$$\frac{dt}{dw} = \sum_{n=0}^{\infty} a_n w^n$$

for some constants a_n. (Note that from the definition of w above, one only expects even powers of w to appear in this power series.) Substituting in one has

$$\begin{aligned}
G(z) &= \exp(zf(t_0))\int_{-\infty}^{\infty}\exp(-zw^2)\sum_{n=0}^{\infty} a_n w^n dw \\
&= \exp(zf(t_0))\sum_{n=0}^{\infty} a_n z^{-(n+1)/2}\Gamma\left(\frac{n+1}{2}\right) \\
&= \frac{\exp(zf(t_0))}{\sqrt{z}}\sum_{m=0}^{\infty} a_{2m}\Gamma\left(m+\frac{1}{2}\right)\left(\frac{1}{z}\right)^m
\end{aligned}$$

where in the last line we have used the fact that only even powers of w appear in dt/dw.

Let us outline how to derive the Stirling series using this method. Recall there that

$$\Gamma(z+1) \;=\; z^{z+1}\int_0^\infty \exp(zf(t))dt$$

for $f(t) = \log t - t$, with only saddle point at $t_0 = 1$. Expand in a Taylor series about $f(t_0) = -1$ to get

$$f(t) \;=\; -1 \;-\; \frac{(t-1)^2}{2!} \;+\; 2\frac{(t-1)^3}{3!} \;-\; (3!)\frac{(t-1)^4}{4!} \;+\; \cdots$$

so in the notation above,

$$w^2 \;=\; \frac{(t-1)^2}{2} \;-\; \frac{(t-1)^3}{3} \;+\; \cdots$$

It can be shown (see [58][section 4.6] for details, but note their a_0 at the bottom of p. 442 is $\sqrt{2}$ not $1/\sqrt{2}$) that

$$\frac{dt}{dw} \;=\; \sqrt{2}\left(1 \;+\; \frac{w^2}{6} \;+\; \frac{w^4}{216} \;+\; \cdots\right)$$

and plugging these values into the general expression for the asymptotic series we find

$$\begin{aligned}\Gamma(z+1) \;&=\; z^{z+1}\frac{\exp(-z)}{\sqrt{z}}\sqrt{2}\left[\Gamma(1/2) \;+\; \frac{1}{6}\Gamma(3/2)\frac{1}{z} \;+\; \frac{1}{216}\Gamma(5/2)\frac{1}{z^2} \;+\; \cdots\right] \\ &=\; \sqrt{2\pi}z^{z+1/2}\exp(-z)\left[1 \;+\; \frac{1}{12z} \;+\; \frac{1}{288z^2} \;+\; \cdots\right]\end{aligned}$$

and so we recover Stirling's series.

For more information, see [57], section 7.4, or [58][section 4.6].

A.4 Uniqueness (or lack thereof)

One very important property of asymptotic series is that they do *not* uniquely determine a function. It is easy to check, for example, that if $\mathrm{Re}(z) > 0$, then the same series can be simultaneously asymptotic to both $f(z)/\varphi(z)$ and $f(z)/\varphi(z) + \exp(-z)$.

This fact is very important in quantum field theory, and is a reflection of nonperturbative effects in the theory. Summing over Feynman diagrams yields a series in which the coupling constant of the theory plays the part of $1/z$. Now, a typical quantum field theory has 'nonperturbative effects,' which cannot be seen in a (perturbative) Feynman diagram expansion. Nonperturbative effects, which are not uniquely determined by the perturbative theory, are exponentially small in the coupling constant, *i.e.* multiplied by factors of

$\exp(-1/g) = \exp(-z)$. Since the Feynman diagram expansion is only an asymptotic series, and the nonperturbative effects are exponentially small, adding nonperturbative effects does not change the asymptotic expansion, *i.e.* does not change the Feynman diagram expansion.

Properties of asymptotic series:

Asymptotic series can be added, multiplied, and integrated term-by-term. However, asymptotic series can only be differentiated term-by-term to obtain an asymptotic expansion for the derivative only if it is known that the derivative possesses an asymptotic expansion.

A.5 Summation of asymptotic series

How can we sum, in any sense, a divergent series?

One approach is as follows. Given a divergent series

$$F(z) \; = \; \sum_{n=0}^{\infty} A_n z^n$$

for some constants A_n, consider the related series

$$B(z) \; = \; \sum_{n=0}^{\infty} A_n \frac{z^n}{n!}$$

Depending upon how badly divergent the original series $F(z)$ was, one might hope that the new series $B(z)$ might actually converge in some region. Assuming that $B(z)$ converges and can be resummed, how might one recover $F(z)$? Well, use the formula

$$\int_0^\infty \exp(-t/z) t^n dt \; = \; z^{n+1} n!$$

to show that, *formally*,

$$z F(z) \; = \; \int_0^\infty \exp(-t/z) B(t) dt$$

To calculate $F(z)$ using the formal trick above, we need $B(t)$ for real positive values of t less than or of order z. So long as any singularities in $B(t)$ on the complex t plane are at distances greater than $|z|$ from the origin, this should be OK. (In quantum field theory, singularities in $B(t)$ are typically associated with nonperturbative effects – instantons – so again we see that nonperturbative effects limit the usefulness of resummation methods for the (asymptotic) Feynman series. See [59] for more information.)

This particular resummation technique is more or less known as Borel summation. There are other techniques one can apply to try to resum asymptotic series (modulo the fundamental nonuniqueness issue), see [61] for more information.

A.6 Stokes' phenomenon

Stokes' phenomenon is the observation that the operations of analytic continuation and asymptotic series expansion do not commute with one another.

Let us work through a simple example, following [58][section 5.3]. Consider the confluent hypergeometric function defined by

$$M(a,c;z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n} \frac{z^n}{n!}$$

where $(a)_n$ is the Pochhammer symbol:

$$\begin{aligned} (a)_n &= a(a+1)(a+2)\cdots(a+n-1) \\ &= \frac{(a+n-1)!}{(a-1)!} = \frac{\Gamma(a+n)}{\Gamma(a)} \\ (a)_0 &= 1 \end{aligned}$$

First, let us derive the leading term in an asymptotic expansion as $z \rightarrow +\infty$ along the real axis. The n term of the series above is given by

$$\frac{(a)_n}{(c)_n} \frac{z^n}{n!} = \frac{(a+n-1)!}{(a-1)!} \frac{(c-1)!}{(c+n-1)!} \frac{z^n}{n!}$$

It can be shown [57][problem 8.3.8] that

$$\lim_{x \rightarrow \infty} x^{b-a} \frac{(x+a)!}{(x+b)!} = 1$$

so we see that for large n, the n term becomes

$$\begin{aligned} &\rightarrow \frac{(c-1)!}{(a-1)!} \frac{n^{(a-1)-(c-1)}}{n!} z^n \\ &\rightarrow \frac{(c-1)!}{(a-1)!} \frac{z^n}{(n-a+c)!} \end{aligned}$$

At large positive z, the terms with large n dominate, so that

$$M(a,c;z) \approx \sum_{n=1}^{\infty} \frac{(c-1)!}{(a-1)!} \frac{z^n}{(n-a+c)!}$$

Assume $a - c \in \mathbf{Z}$, and ignoring the first few terms (small compared to the rest), we have that

$$\begin{aligned} M(a,c;z) &\approx \frac{(c-1)!}{(a-1)!} \sum_{m=0}^{\infty} \frac{z^m}{m!} z^{a-c} \quad (m = n-a+c) \\ &= \frac{\Gamma(c)}{\Gamma(a)} z^{a-c} \exp(z) \end{aligned}$$

We have been sloppy, but a more careful analysis reveals this result is correct; the leading term in an asymptotic series expansion of $M(a, c; z)$ for $z \to +\infty$ along the real axis is given by the expression above.

Next, let us find the leading term in an asymptotic expansion of $M(a, c; z)$ for $z \to -\infty$ along the real axis. The fast way to do this is to use the identity

$$M(a, c; z) \; = \; \exp(z) M(c - a, c; -z)$$

Using this and our previous result, we see that as $z \to -\infty$ along the real axis,

$$M(a, c; z) \; \to \; \frac{\Gamma(c)}{\Gamma(c - a)} (-z)^{-a}$$

This is the leading term in an asymptotic series expansion of $M(a, c; z)$ for $z \to -\infty$.

But if we compare our results for the two different limits, we see that this is not what we would have gotten by analytically continuing either separately.

For example, if we started with the $z \to +\infty$ limit, and analytically continued, we would have found

$$\frac{\Gamma(c)}{\Gamma(a)} (-z)^{a-c} \exp(-z) \; \neq \; \frac{\Gamma(c)}{\Gamma(c - a)} (-z)^{-a}$$

Thus, analytic continuation does *not* commute with asymptotic series expansions. This is known as Stokes' phenomenon. This is very unlike convergent Taylor series, for example, where analytic continuation does commute with series expansion.

Let us also understand this more systematically by taking asymptotic series expansions at all $\arg(z)$, not just along the real axis. To do this, we shall use the integral representation [57][problem 13.5.10a]

$$M(a, c; z) \; = \; \frac{\Gamma(c)}{\Gamma(a)\Gamma(c - a)} \int_0^1 \exp(zt) t^{a-1} (1 - t)^{c-a-1} dt$$

Deform the integration path to go from first, $t = 0$ to $t = -\infty \exp(-i\phi)$, then, from $t = -\infty \exp(-i\phi)$ to $t = 1$, where $z = |z| \exp(i\phi)$. Then we find that

$$\begin{aligned} M(a, c; z) \; = \; \frac{\Gamma(c)}{\Gamma(a)\Gamma(c - a)} \Big\{ & \int_0^{-\infty \exp(-i\phi)} \exp(zt) t^{a-1} (1 - t)^{c-a-1} dt \\ & + \int_{-\infty \exp(-i\phi)}^1 \exp(zt) t^{a-1} (1 - t)^{c-a-1} dt \Big\} \end{aligned}$$

When $0 < \phi < \pi$, make the following changes of variables. In the first integral, define w by

$$t \; = \; -\frac{w \exp(-i\phi)}{|z|} \; = \; -\frac{w}{z}$$

and in the second integral, define

$$t \;=\; 1 \;-\; \frac{u\exp(-i\phi)}{|z|} \;=\; 1 - \frac{u}{z}$$

Plugging these in, we find

$$\begin{aligned}
M(a,c;z) \;&=\; \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\left\{\int_0^\infty \exp(-w)\left(\frac{-w}{z}\right)^{a-1}\left(1+\frac{w}{z}\right)^{c-a-1}\left(-\frac{dw}{z}\right)\right. \\
&\qquad \left. +\;\int_\infty^0 \exp(z(1-u/z))\left(1-\frac{u}{z}\right)^{a-1}\left(\frac{u}{z}\right)^{c-a-1}\left(-\frac{du}{z}\right)\right\} \\
&=\; \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\left\{\frac{\exp(ia\pi)}{z^a}\int_0^\infty \exp(-w)w^{a-1}\left(1+\frac{w}{z}\right)^{c-a-1}dw\right. \\
&\qquad \left. +\;\frac{\exp(z)}{z^{c-a}}\int_0^\infty \exp(-u)\left(1-\frac{u}{z}\right)^{a-1}u^{c-a-1}du\right\} \\
&\to\; \frac{\Gamma(c)}{\Gamma(c-a)}\frac{\exp(ia\pi)}{z^a} \;+\; \frac{\Gamma(c)}{\Gamma(a)}\exp(z)z^{a-c}
\end{aligned}$$

The first term dominates when $\phi=\pi$, at which the second term is negligible. When $\phi=0$, the opposite is true: the second term dominates, the other term is negligible. For $\phi=\pi/2$, the two terms are comparable.

The analysis can be repeated for $-\pi<\phi<0$; but as it is very similar, for brevity we shall not repeat it here.

So, what we have found in general is that for general ϕ, the leading term is a combination of the two terms, but in the two limits, one dominates and the other is much less than the corrections, so that the leading term in the asymptotic series expansion is defined by only term, not both. Which term dominates, varies as ϕ changes. Thus, analytic continuation does not commute with asymptotic series expansion.

B A_∞ algebras

A strict, or strong A_∞ algebra is defined by a **Z**-graded vector space A together with a collection of multiplications $m_n : A^{\otimes n} \to A$, for $n \geq 1$, which are homogeneous of degree $2-n$ and are subject to the constraint

$$\sum_{k+l=n+1 j=0\cdots k-1} (-)^s\, m_k\left(u_1,\cdots,u_j, m_l(u_{j+1},\cdots,u_{j+l}), u_{j+l+1},\cdots,u_n\right) \;=\; 0$$

for $n \geq 1$, where

$$s \;=\; l\left(|u_1|+\cdots+|u_j|\right) \;+\; j(l-1) \;+\; (k-1)l$$

and $|\cdot|$ denotes the degree of a homogeneous element in A.

The first six constraints can be written as follows. For $n=1$,

$$m_1 \circ m_1 \;=\; 0$$

If we think of m_1 as a differential d, then this is just the condition $d^2=0$. For $n=2$,

$$m_1\left(m_2(u_1,u_2)\right) \;-\; m_2\left(m_1(u_1),u_2\right) \;+\; (-)^{|u_1|+1} m_2\left(u_1,m_1(u_2)\right) \;=\; 0$$

Note that this amounts to the product rule on the product defined by m_2, with differential m_1. For $n=3$,

$$\begin{aligned}
&m_1\left(m_3(u_1,u_2,u_3)\right) \;+\; m_2\left(m_2(u_1,u_2),u_3\right) \;-\; m_2\left(u_1,m_2(u_2,u_3)\right) \\
&\quad + m_3\left(m_1(u_1),u_2,u_3\right) \;+\; (-)^{|u_1|} m_3\left(u_1,m_1(u_2),u_3\right) \;+\; (-)^{|u_1|+|u_2|} m_3\left(u_1,u_2,m_1(u_3)\right) \;=\; 0
\end{aligned}$$

Let us take a moment to interpret this. This condition is just the statement that the product defined by m_2 is associative *up to* the multiplication m_3. In particular, an A_∞ algebra is not associative in general. For $n=4$,

$$\begin{aligned}
&m_1\left(m_4(u_1,u_2,u_3,u_4)\right) \;-\; m_2\left(m_3(u_1,u_2,u_3),u_4\right) \;+\; (-)^{|u_1|+1} m_2\left(u_1,m_3(u_1,u_2,u_3)\right) \\
&\quad + m_3\left(m_2(u_1,u_2),u_3,u_4\right) \;-\; m_3\left(u_1,m_2(u_2,u_3),u_4\right) \;+\; m_3\left(u_1,u_2,m_2(u_3,u_4)\right) \\
&\quad - m_4\left(m_1(u_1),u_2,u_3,u_4\right) \;+\; (-)^{|u_1|+1} m_4\left(u_1,m_1(u_2),u_3,u_4\right) \\
&\quad + (-)^{|u_1|+|u_2|+1} m_4\left(u_1,u_2,m_1(u_3),u_4\right) \;+\; (-)^{|u_1|+|u_2|+|u_3|+1} m_4\left(u_1,u_2,u_3,m_1(u_4)\right) \;=\; 0
\end{aligned}$$

For $n=5$,

$$\begin{aligned}
&m_1\left(m_5(u_1,u_2,u_3,u_4,u_5)\right) \;+\; m_2\left(m_4(u_1,u_2,u_3,u_4),u_5\right) \;-\; m_2\left(u_1,m_4(u_2,u_3,u_4,u_5)\right) \\
&\quad + m_3\left(m_3(u_1,u_2,u_3),u_4,u_5\right) \;+\; (-)^{|u_1|} m_3\left(u_1,m_3(u_2,u_3,u_4),u_5\right) \\
&\quad + (-)^{|u_1|+|u_2|} m_3\left(u_1,u_2,m_3(u_3,u_4,u_5)\right) \;+\; m_4\left(m_2(u_1,u_2),u_3,u_4,u_5\right) \;-\; m_4\left(u_1,m_2(u_2,u_3),u_4,u_5\right) \\
&\quad + m_4\left(u_1,u_2,m_2(u_3,u_4),u_5\right) \;-\; m_4\left(u_1,u_2,u_3,m_2(u_4,u_5)\right) \;+\; m_5\left(m_1(u_1),u_2,u_3,u_4,u_5\right) \\
&\quad + (-)^{|u_1)} m_5\left(u_1,m_1(u_2),u_3,u_4,u_5\right) \;+\; (-)^{|u_1|+|u_2|} m_5\left(u_1,u_2,m_1(u_3),u_4,u_5\right) \\
&\quad + (-)^{|u_1|+|u_2|+|u_3|} m_5\left(u_1,u_2,u_3,m_1(u_4),u_5\right) \\
&\quad + (-)^{|u_1|+|u_2|+|u_3|+|u_4|} m_5\left(u_1,u_2,u_3,u_4,m_1(u_5)\right) \;=\; 0
\end{aligned}$$

For $n=6$,

$$\begin{aligned}
&m_1\left(m_6(u_1,u_2,u_3,u_4,u_5,u_6)\right) \;-\; m_2\left(m_5(u_1,u_2,u_3,u_4,u_5),u_6\right) \;+\; (-)^{|u_1|+1} m_2\left(u_1,m_5(u_2,u_3,u_4,u_5,u_6)\right) \\
&\quad + m_3\left(m_4(u_1,u_2,u_3,u_4),u_5,u_6\right) \;-\; m_3\left(u_1,m_4(u_2,u_3,u_4,u_5),u_6\right) \;+\; m_3\left(u_1,u_2,m_4(u_3,u_4,u_5,u_6)\right) \\
&\quad - m_4\left(m_3(u_1,u_2,u_3),u_4,u_5,u_6\right) \;+\; (-)^{|u_1|+1} m_4\left(u_1,m_3(u_2,u_3,u_4),u_5,u_6\right) \\
&\quad + (-)^{|u_1|+|u_2|+1} m_4\left(u_1,u_2,m_3(u_3,u_4,u_5),u_6\right) \;+\; :(-)^{|u_1|+|u_2|+|u_3|+1} m_4\left(u_1,u_2,u_3,m_3(u_4,u_5,u_6)\right) \\
&\quad + m_5\left(m_2(u_1,u_2),u_3,u_4,u_5,y_6\right) \;-\; m_5\left(u_1,m_2(u_2,u_3),u_4,u_5,u_6\right)
\end{aligned}$$

$$
\begin{aligned}
&+ m_5\,(u_1, u_2, m_2(u_3, u_4), u_5, u_6) \;-\; m_5\,(u_1, u_2, u_3, m_2(u_4, u_5), u_6) \\
&+ m_5\,(u_1, u_2, u_3, u_4, m_2(u_5, u_6)) \;-\; m_6\,(m_1(u_1), u_2, u_3, u_4, u_5, u_6) \\
&+ (-)^{|u_1|+1} m_6\,(u_1, m_1(u_2), u_3, u_4, u_5, u_6) \;+\; (-)^{|u_1|+|u_2|+1} m_6\,(u_1, u_2, m_1(u_3), u_4, u_5, u_6) \\
&+ (-)^{|u_1|+|u_2|+|u_3|+1} m_6\,(u_1, u_2, u_3, m_1(u_4), u_5, u_6) \\
&+ (-)^{|u_1|+|u_2|+|u_3|+|u_4|+1} m_6\,(u_1, u_2, u_3, u_4, m_1(u_5), u_6) \\
&+ (-)^{|u_1|+|u_2|+|u_3|+|u_4|+|u_5|+1} m_6\,(u_1, u_2, u_3, u_4, u_5, m_1(u_6)) \;=\; 0
\end{aligned}
$$

In the special case that $m_1 = 0$, we say the A_∞ algebra is a *minimal* A_∞ algebra or *minimal model*. Another interesting special case is when $m_k = 0$ for $k \geq 3$. In this special case, the A_∞ algebra is known as a *differential graded algebra*, or dga for short.

An A_∞ morphism is a set of maps $f_k : A^{\otimes k} \to B$ such that (up to signs)

$$
\sum_{r+s+t=n} f_u(1^{\otimes r} \otimes m_s \otimes 1^{\otimes t}) \;=\; \sum_{1 \leq r \leq n, i_1 + \cdots + i_r = n} m_r(f_{i_1} \otimes f_{i_2} \otimes \cdots \otimes f_{i_r})
$$

for any $n > 0$ and $u = n + 1 - s$.

Given a dga, one may construct a minimal model $H^*(A)$, by taking the cohomology with respect to m_1. It is possible to define [56] an A_∞ structure on $H^*(A)$ such that there is an A_∞ morphism $f : H^*(A) \to A$ with f_1 equal to an embedding $i : H^*(A) \hookrightarrow A$. The A_∞ structure on $H^*(A)$ is unique up to A_∞ isomorphisms.

The notion of an A_∞ algebra can be extended to an A_∞ category. The difference between an ordinary category and an A_∞ category is that the morphisms in the latter have an A_∞ structure. In particular, an A_∞ category is not a category in the usual sense, as composition of morphisms need not be associative.

References

[1] B. Schutz, *A first course in general relativity*, Cambridge University Press, 1990.

[2] D. Gross, E. Witten, "Superstring modifications of Einstein's equations," Nucl. Phys. **B277** (1986) 1-10.

[3] D. Nemeschansky, A. Sen, "Conformal invariance of supersymmetric sigma models on Calabi-Yau manifolds," Phys. Lett. **B178** (1985) 365-369.

[4] M. Peskin, D. Schroeder, *An introduction to quantum field theory*, Addison-Wesley, 1995.

[5] P. Aspinwall, D. Morrison, "Topological field theory and rational curves," Comm. Math. Phys. **151** (1993) 245-262, `hep-th/9110048`.

[6] E. Witten, "Mirror manifolds and topological field theory," hep-th/9112056.

[7] E. Sharpe, "Notes on certain other (0,2) correlation functions," hep-th/0605005.

[8] D. Morrison, R. Plesser, "Towards mirror symmetry as duality for two-dimensional abelian gauge theories," Nucl. Phys. Proc. Suppl. **46** (1996) 177-186, hep-th/9508107.

[9] K. Hori, C. Vafa, "Mirror symmetry," hep-th/0002222.

[10] M. Aganagic, C. Vafa, "Mirror symmetry and supermanifolds," hep-th/0403192.

[11] L. Borisov, L. Chen, G. Smith, "The orbifold Chow ring of toric Deligne-Mumford stacks," math.AG/0309229.

[12] T. Pantev, E. Sharpe, "Notes on gauging noneffective group actions," hep-th/0502027.

[13] T. Pantev, E. Sharpe, "String compactifications on Calabi-Yau stacks," hep-th/0502044.

[14] T. Pantev, E. Sharpe, "GLSM's for gerbes (and other toric stacks)," Adv. Theor. Math. Phys. **10** (2006) 77-121, hep-th/0502053.

[15] S. Hellerman, A. Henriques, T. Pantev, E. Sharpe, M. Ando, "Cluster decomposition, T-duality, and gerby CFT's," hep-th/0606034.

[16] E. Sharpe, "Derived categories and stacks in physics," hep-th/0608056.

[17] R. Donagi, E. Sharpe, "GLSM's for partial flag manifolds," arXiv: 0704.1761.

[18] A. Caldararu, J. Distler, S. Hellerman, T. Pantev, E. Sharpe, "Non-birational twisted derived equivalences in abelian GLSM's," to appear.

[19] K. Hori, D. Tong, "Aspects of non-abelian gauge dynamics in two-dimensional $\mathcal{N} = (2,2)$ theories," hep-th/0609032.

[20] A. Kuznetsov, "Homological projective duality," math.AG/0507292.

[21] A. Kuznetsov, "Derived categories of quadric fibrations and intersections of quadrics," math.AG/0510670.

[22] A. Kuznetsov, "Homological projective duality for Grassmannians of lines," math.AG/0610957.

[23] M. Atiyah, *The geometry and physics of knots*, Cambridge University Press, 1990.

[24] K. Costello, "Topological conformal field theories and Calabi-Yau categories," math.QA/0412149.

[25] K. Costello, “Topological conformal field theories and gauge theories,” math.QA/0605647.

[26] A. Abouelsaood, C. Callan, C. Nappi, and S. Yost, “Open strings in background gauge fields,” Nucl. Phys. **B280** (1987) 599-624.

[27] J. Harvey, G. Moore, “On the algebras of BPS states,” Comm. Math. Phys. **197** (1998) 489-519, hep-th/9609017.

[28] M. Kontsevich, “Homological algebra of mirror symmetry,” in *Proceedings of the International Congress of Mathematicians*, pp. 120-139, Birkhäuser, 1995, alg-geom/9411018.

[29] E. Witten, “Chern-Simons gauge theory as a string theory,” Progr. Math. **133** (1995) 637-678, hep-th/9207094.

[30] D. Freed, E. Witten, “Anomalies in string theory with D-branes,” hep-th/9907189.

[31] S. Katz, E. Sharpe, “D-branes, open string vertex operators, and Ext groups,” Adv. Theor. Math. Phys. **6** (2003) 979-1030, hep-th/0208104.

[32] R. Donagi, S. Katz, and E. Sharpe, “Spectra of D-branes with Higgs vevs,” Adv. Theor. Math. Phys. **8** (2005) 813-859, hep-th/0309270.

[33] Y. Manin, *Gauge field theory and complex geometry*, Springer-Verlag, 1988.

[34] E. Sharpe, “D-branes, derived categories, and Grothendieck groups,” Nucl. Phys. **B561** (1999) 433-450, hep-th/9902116.

[35] M. Douglas, “D-branes, categories, and $\mathcal{N} = 1$ supersymmetry,” J. Math. Phys. **42** (2001) 2818-2843, hep-th/0011017.

[36] P. Aspinwall, A. Lawrence, “Derived categories and zero-brane stability,” JHEP 0108 (2001) 004, hep-th/0104147.

[37] E. Sharpe, “Lectures on D-branes and sheaves,” hep-th/0307245.

[38] A. Kapustin, D. Orlov, “Remarks on A-branes, mirror symmetry, and the Fukaya category,” J. Geom. Phys. **48** (2003) 84-99, hep-th/0109098.

[39] K. Hori, “Linear models of supersymmetric D-branes,” hep-th/0012179.

[40] P. Kraus, F. Larsen, “Boundary string field theory of the $D\overline{D}$ system,” hep-th/0012198.

[41] T. Takayanagi, S. Terashima, T. Uesugi, “Brane-antibrane action from boundary string field theory,” hep-th/0012210.

[42] K. Hori, M. Herbst, D. Page, to appear.

[43] A. Bondal, M. Kapranov, "Enhanced triangulated categories," Math. USSR Sbornik **70** (1991) 93-107.

[44] C. Lazaroiu, "Generalized complexes and string field theory," `hep-th/0102122`.

[45] C. Lazaroiu, "Graded lagrangians, exotic topological D-branes and enhanced triangulated categories," `hep-th/0105063`.

[46] E. Diaconescu, "Enhanced D-brane categories from string field theory," JHEP 0106 (2001) 016, `hep-th/0104200`.

[47] P. Aspinwall, S. Katz, "Computation of superpotentials for D-branes," `hep-th/0412209`.

[48] C. Vafa, "Topological Landau-Ginzburg models," Mod. Phys. Lett. **A6** (1991) 337-346.

[49] N. P. Warner, "Supersymmetry in boundary integrable models," Nucl. Phys. **B450** (1995) 663-694, `hep-th/9506064`.

[50] D. Orlov, "Derived categories of coherent sheaves and triangulated categories of singularities," `math.AG/0503632`.

[51] D. Orlov, "Triangulated categories of singularities and D-branes in Landau-Ginzburg orbifolds," Proc. Steklov Inst. Math. **246** (2004) 227-248, `math.AG/0302304`.

[52] J. Walcher, "Stability of Landau-Ginzburg branes," `hep-th/0412274`.

[53] E. Witten, "Phases of $\mathcal{N} = 2$ theories in two dimensions," Nucl. Phys. **B403** (1993) 159-222, `hep-th/9301042`.

[54] A. Kapustin, E. Witten, "Electric-magnetic duality and the geometric Langlands program," `hep-th/0604151`.

[55] V. Batyrev, "Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties," `alg-geom/9310003`.

[56] T. V. Kadeishvili, "The algebraic structure in the homology of an A_∞ algebra," Soobshch. Akad. Nauk. Gruzin. SSR **108** (1982) 249-252.

[57] G. Arfken, H. Weber, *Mathematical methods for physicists*, Academic Press, 2001.

[58] P. Morse and H. Feshbach, *Methods of Theoretical Physics* volume I, McGraw-Hill, New York, 1953.

[59] S. Weinberg, *The Quantum Theory of Fields*, volume II, Cambridge University Press, 1996, section 20.7.

[60] B. Shivamoggi, *Perturbation Methods for Differential Equations*, Birkhäuser, Boston, 2003.

[61] G. H. Hardy, *Divergent Series*, Oxford University Press, 1949.

Index