Lecture Notes in Physics 851

Marco Baumgartl
Ilka Brunner
Michael Haack *Editors*

Strings and Fundamental Physics

Springer

Lecture Notes in Physics

Volume 851

For further volumes:
http://www.springer.com/series/5304

The Lecture Notes in Physics

The series Lecture Notes in Physics (LNP), founded in 1969, reports new developments in physics research and teaching—quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge in an accessible way. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research and to serve three purposes:

- to be a compact and modern up-to-date source of reference on a well-defined topic
- to serve as an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas
- to be a source of advanced teaching material for specialized seminars, courses and schools

Both monographs and multi-author volumes will be considered for publication. Edited volumes should, however, consist of a very limited number of contributions only. Proceedings will not be considered for LNP.

Volumes published in LNP are disseminated both in print and in electronic formats, the electronic archive being available at springerlink.com. The series content is indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

Proposals should be sent to a member of the Editorial Board, or directly to the managing editor at Springer:

Christian Caron
Springer Heidelberg
Physics Editorial Department I
Tiergartenstrasse 17
69121 Heidelberg/Germany
christian.caron@springer.com

Marco Baumgartl · Ilka Brunner
Michael Haack
Editors

Strings and Fundamental Physics

Springer

Marco Baumgartl
Universität Hamburg
II. Institut für Theoretische Physik
Hamburg
Germany

Michael Haack
Department für Physik
Ludwig-Maximilians-Universität München
München
Germany

Ilka Brunner
Department für Physik
Ludwig-Maximilians-Universität München
München
Germany

ISSN 0075-8450
e-ISSN 1616-6361
ISBN 978-3-642-25946-3
e-ISBN 978-3-642-25947-0
DOI 10.1007/978-3-642-25947-0
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2011945027

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

One of the major open problems in theoretical physics is the lack of a unified description of the theory of general relativity and the theory of quantum fields. During the last decades string theory has provided a promising framework for the investigation of this issue. The basic idea is simple and revolutionary at the same time: by replacing the concept of a point particle with a one-dimensional string a whole new field of research has been opened up. The consequences of this are not yet fully conceivable today. Up to now string theory has offered a new way to view all particles as different excitations of the same fundamental object, it has celebrated success by discovering the graviton in its spectrum, it has forced us to consider space-times with more than four dimensions containing dynamical hypermanifolds, D-branes, in their vacuum structure and it has triggered numerous interesting developments in fields as different as condensed matter physics and pure mathematics. Still, as a physical theory string theory is not yet fully understood and remains a very active area of research.

In this book it is our aim to collect pedagogical lectures by leading experts in string theory introducing the reader to some of the newest developments in the field. In no way it is possible to give an overview over the whole research spectrum of string theory. Rather we have carefully selected topics which are at the cutting edge of research in string theory. This includes new developments in topics with long history like for example topological strings or AdS/CFT dualities, but also topics which appeared only recently like doubled field theory and holography in the hydrodynamical regime.

The contributions to this book are selected in a way so that it can be considered as a self-contained textbook. Readers with a basic familiarity with string theory will find it possible to use these lectures to catch up with some of the latest developments, enabling them to follow recent research articles on the subjects.

These lectures were given at the summer school “Strings and Fundamental Physics 2010” in the framework of the Excellence Cluster ‘Universe’, Munich, Germany, and was attended by numerous students and postdoctoral researchers. Videos of the lectures and additional material like exercises can be found on the webpage of the school: www.theorie.physik.uni-muenchen.de/activities/schools/archiv/sfp10

We want to thank all those who contributed to the success of this summer school and helped in one way or another to compose this book. Primarily we want to thank the contributors Ralph Blumenhagen, Atish Dabholkar, Johanna Erdmenger, Neil Lambert, Suresh Nampuri, Hirosi Ooguri, Dam Thanh Son and Barton Zwiebach.

In addition, we are most grateful to Rosa-Anna Friedl-Gründler for her invaluable support in all organizational matters before as well as during the school. Finally, we would like to thank Martin Ammon, Michael Kay, Nicolas Moeller and Daniel Plencner for their help in typing up some of the lectures and Stefan Theisen for giving us the opportunity to publish this set of tutorials in the Lecture Notes in Physics.

We acknowledge financial support by the Cluster of Excellence for Fundamental Physics "Origin and Structure of the Universe," the German Academic Exchange Service DAAD, the Elite Master Course Theoretical and Mathematical Physics TMP (at the Ludwig-Maximilians-University of Munich LMU) and the Arnold Sommerfeld Cernter for Theoretical Physics ASC (LMU). Finally we thank the Technical University of Munich TUM for providing the lecture hall and the TUM and LMU for providing the necessary infrastructure.

Hamburg, Munich, August 2011

Marco Baumgartl
Ilka Brunner
Michael Haack

Contents

Contributors

Ralph Blumenhagen Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Föhringer Ring 6, 80805 München, Germany, e-mail: blumenha@mpp.mpg.de

Atish Dabholkar Laboratoire de Physique Théorique et Hautes Energies (LPTHE), Université Pierre et Marie Curie-Paris 6, CNRS UMR 7589. Tour 13-14, 5ème étage, Boite 126, 4 Place Jussieu, 75252 Paris Cedex 05, France, e-mail: atish@lpthe.jussieu.fr

Johanna Erdmenger Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Föhringer Ring 6, 80805 München, Germany, e-mail: jke@mpp.mpg.de

Neil Lambert Theory Division, CERN, 1211 Geneva 23, Switzerland, e-mail: neil.lambert@cern.ch

Suresh Nampuri Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München, Theresienstrasse 37, 80333 Munich, Germany, e-mail: suresh.nampuri@physik.uni-muenchen.de

Hirosi Ooguri California Institute of Technology, 452-48, Pasadena, CA 91125, USA; University of Tokyo, Institute for the Physics and Mathematics of the Universe, Todai Institutes for Advanced Study, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan, e-mail: ooguri@theory.caltech.edu

Dam Thanh Son Institute for Nuclear Theory, University of Washington, Seattle, WA 98195-1550, USA, e-mail: dtson@u.washington.edu

Barton Zwiebach Massachusetts Institute of Technology, 77 Massachusetts Avenue, Bldg. 6-305, Cambridge, MA 02139, USA, e-mail: zwiebach@mit.edu

Chapter 1
String Theory 101

Neil Lambert

1.1 Introduction: Why String Theory?

The so-called Standard Model of Particle Physics is the most successful scientific theory of Nature in the sense that no other theory has such a high level of accuracy over such a complete range of physical phenomena using such a modest number of assumptions and parameters. It is unreasonably good and was never intended to be so successful. Since its formulation around 1970 there has not been a single experimental result that has produced even the slightest disagreement. Nothing, despite an enormous amount of effort. But there are skeletons in the closet. Let me mention just three.

The first is the following: Where does the Standard Model come from? For example it has quite a few parameters which are only fixed by experimental observation. What fixes these? It postulates a certain spectrum of fundamental particle states but why these? In particular these particle states form three families, each of which is a copy of the others, differing only in their masses. Furthermore only the lightest family seems to have much to do with life in the universe as we know it, so why the repetition? It is somewhat analogous to Mendelev's periodic table of the elements. There is clearly a discernible structure but this wasn't understood until the discovery of quantum mechanics. We are looking for the underlying principle that gives the somewhat bizarre and apparently ad hoc structure of the Standard Model. Moreover the Standard Model also doesn't contain Dark Matter that constitutes most of the 'stuff' in the observable universe.

The second problem is that, for all its strengths, the Standard Model does not include gravity. For that we must use General Relativity which is a classical theory and as such is incompatible with the rules of quantum mechanics. Observationally

N. Lambert (✉)
Theory Division, CERN,
1211 Geneva 23, Switzerland
e-mail: neil.lambert@cern.ch

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_1,

this is not a problem since the effect of gravity, at the energy scales which we probe, is smaller by a factor of 10^{-40} than the effects of the subnuclear forces which the Standard Model describes. You can experimentally test this assertion by lifting up a piece of paper with your little finger. You will see that the electromagnetic forces at work in your little finger can easily overcome the gravitational force of the entire earth which acts to pull the paper to the floor.

However this is clearly a problem theoretically. We can't claim to understand the universe physically until we can provide one theory which consistently describes gravity and the subnuclear forces. If we do try to include gravity into QFT then we encounter problems. A serious one is that the result is non-renormalizable, apparently producing an infinite series of divergences which must be subtracted by inventing an infinite series of new interactions, thereby removing any predictive power. Thus we cannot use the methods of QFT as a fundamental principle for gravity.

The third problem I want to mention is more technical. Quantum field theories generically only make mathematical sense if they are viewed as a low energy theory. Due to the effects of renormalization the Standard Model cannot be valid up to all energy scales, even if gravity was not a problem. Mathematically we know that there must be something else which will manifest itself at some higher energy scale. All we can say is that such new physics must arise before we reach the quantum gravity scale, which is some 10^{17} orders of magnitude above the energy scales that we have tested to date. To the physicists who developed the Standard Model the surprise is that we have not already seen such new physics many years ago. And we are all hoping to see it soon at the LHC.

With these comments in mind this course will introduce string theory, which, for good or bad, has become the dominant, and arguably only, framework for a complete theory of all known physical phenomena. As such it is in some sense a course to introduce the modern view of particle physics at its most fundamental level. Whether or not string theory is ultimately relevant to our physical universe is unknown, and indeed may never be known. However it has provided many deep and powerful ideas. Certainly it has had a profound effect upon pure mathematics. But an important feature of string theory is that it naturally includes gravitational and subnuclear-type forces consistently in a manner consistent with quantum mechanics and relativity (as far as anyone knows). Thus it seems fair to say that there is a mathematical framework which is capable of describing all of the physics that we know to be true. This is no small achievement.

However it is also fair to say that no one actually knows what string theory really is. In any event this course can only attempt to be a modest introduction that is aimed at students with no previous knowledge of string theory. There will be much that we will not have time to discuss: most notably the Veneziano amplitude, anomaly cancellation and compactification. The reader will undoubtedly benefit from the other courses in the School, in particular the notes of Ralph Blumenhagen on D-branes. Furthermore much more extensive and detailed discussions can be found in ref. [1–4]

We will first discuss the bosonic string in some detail. Although this theory is unphysical in several ways (it has a tachyon and no fermions) it is simpler to study than the superstring but has all the main ideas built-in. We then add worldsheet fermions and supersymmetry to obtain the superstring theories that are used in current research but our discussion will be relatively brief.

1.2 Classical and Quantum Dynamics of Point Particles

1.2.1 Classical Action

We want to describe a single particle moving in spacetime. For now we simply consider flat D-dimensional Minkowski space

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + \ldots + (dx^{D-1})^2 \tag{1.2.1}$$

A particle has no spatial extent but it does trace out a curve—its worldline—in spacetime. Furthermore in the absence of external forces this will be a straight line (geodesic if you know GR). In other words the equation of motion should be that the length of the worldline is extremized. Thus we take

$$\begin{aligned} S_{pp} &= -m \int ds \\ &= -m \int \sqrt{-\eta_{\mu\nu}\dot{X}^\mu \dot{X}^\nu} d\tau \end{aligned} \tag{1.2.2}$$

where τ parameterizes the points along the worldline and $X^\mu(\tau)$ gives the location of the particle in spacetime, i.e. the embedding coordinates of the worldline into spacetime.

Let us note some features of this action. Firstly it is manifestly invariant under spacetime Lorentz transformations $X^\mu \to \Lambda^\mu_\nu X^\nu$ where $\Lambda^T \eta \Lambda = \eta$. Secondly it is reparameterization invariant under $\tau \to \tau'(\tau)$ for any invertible change of worldline coordinate

$$d\tau = \frac{d\tau}{d\tau'} d\tau', \qquad \dot{X}^\mu = \frac{dX^\mu}{d\tau} = \frac{d\tau'}{d\tau}\frac{dX^\mu}{d\tau'} \tag{1.2.3}$$

thus

$$\begin{aligned} S_{pp} &= -m \int \sqrt{-\eta_{\mu\nu}\frac{dX^\mu}{d\tau}\frac{dX^\nu}{d\tau}} d\tau \\ &= -m \int \sqrt{-\eta_{\mu\nu}\left(\frac{d\tau'}{d\tau}\right)^2 \frac{dX^\mu}{d\tau'}\frac{dX^\nu}{d\tau'}\frac{d\tau}{d\tau'}} d\tau' \\ &= -m \int \sqrt{-\eta_{\mu\nu}\frac{dX^\mu}{d\tau'}\frac{dX^\nu}{d\tau'}} d\tau' \end{aligned} \tag{1.2.4}$$

Thirdly we can see why the m appears in front and with a minus sign by looking at the non-relativistic limit. In this case we choose a gauge for the worldline reparameterization invariance such that $\tau = X^0$ i.e. worldline 'time' is the same as spacetime 'time'. This is known as static gauge. It is a gauge choice since, as we have seen, we are free to take any parameterization we like. The nonrelativistic limit corresponds to assuming that $\dot{X}^i \ll 1$. In this case we can expand

$$S_{pp} = -m \int \sqrt{1 - \delta_{ij}\dot{X}^i\dot{X}^j} d\tau = \int -m + \frac{1}{2}m\delta_{ij}\dot{X}^i\dot{X}^j d\tau + \ldots \qquad (1.2.5)$$

where the ellipses denotes terms with higher powers of the velocities $\dot{X}^i$. The second term is just the familiar kinetic energy $\frac{1}{2}mv^2$. The first term is simply a constant and doesn't affect the equations of motion. However it can be interpreted as a constant potential energy equal to the rest mass of the particle. Thus we see that the m and minus signs are correct.

Moving on let us consider the equations of motion and conservation laws that follow from this action. The equations of motion follow from the usual Euler-Lagrange equations applied to S_{pp}:

$$\frac{d}{d\tau}\left(\frac{\dot{X}^\nu}{\sqrt{-\eta_{\lambda\rho}\dot{X}^\lambda\dot{X}^\rho}}\right) = 0 \qquad (1.2.6)$$

These equations can be understood as conservation laws since the Lagrangian is invariant under constant shifts $X^\mu \to X^\mu + b^\mu$. The associated charge is

$$p^\mu = \frac{m\dot{X}^\mu}{\sqrt{-\eta_{\lambda\rho}\dot{X}^\lambda\dot{X}^\rho}} \qquad (1.2.7)$$

so that indeed the equation of motion is just $\dot{p}^\mu = 0$. Note that I have called this a charge and not a current. In this case it doesn't matter because the Lagrangian theory we are talking about, the worldline theory of the point particle, is in zero spatial dimensions. So I could just as well called p^μ a conserved current with the conserved charge being obtained by integrating the temporal component of p^μ over space. Here there is no space p^μ only has temporal components.

Warning: We are thinking in terms of the worldline theory where the index μ labels the different scalar fields X^μ, it does not label the coordinates of the worldline. In particular p^0 is not the temporal component of p^μ from the worldline point of view. This confusion between worldvolume coordinates and spacetime coordinates arises throughout string theory.

If we go to static gauge again, where $\tau = X^0$ and write $v^i = \dot{X}^i$ then we have the equations of motion

$$\frac{d}{d\tau}\frac{v^i}{\sqrt{1 - v^2}} = 0 \qquad (1.2.8)$$

and conserved charge

$$p^i = m\frac{v^i}{\sqrt{1-v^2}} \tag{1.2.9}$$

which is simply the spatial momentum. These are the standard relativistic expressions.

We can solve the equation of motion in terms of the constant of motion p^i by writing

$$\frac{v^i}{\sqrt{1-v^2}} = p^i/m \iff p^2/m^2 = \frac{v^2}{1-v^2} \iff v^2 = \frac{p^2}{p^2+m^2} \tag{1.2.10}$$

hence

$$X^i(\tau) = X^i(0) + \frac{p^i\tau}{\sqrt{p^2+m^2}} \tag{1.2.11}$$

and we see that v^i is constant with $v^2<1$.

1.2.2 Electromagnetic Field

Next we can consider a particle interacting with an external electromagnetic field. An electromagnetic field is described by a vector potential A_μ and its field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The natural action of a point particle of mass m and charge q in the presence of such an electromagnetic field is

$$S_{pp} = -m\int\sqrt{-\eta_{\mu\nu}\dot{X}^\mu\dot{X}^\nu}d\tau + q\int A_\mu(X)\dot{X}^\mu d\tau \tag{1.2.12}$$

For those who know differential geometry the vector potential is a connection one-form on spacetime and $A_\mu\dot{X}^\mu d\tau$ is simply the pull-back of A_μ to the worldline of the particle.

The equation of motion is now

$$-m\frac{d}{d\tau}\left(\frac{\eta_{\mu\nu}\dot{X}^\nu}{\sqrt{-\eta_{\lambda\rho}\dot{X}^\lambda\dot{X}^\rho}}\right) - q\frac{d}{d\tau}A_\mu + q\partial_\mu A_\nu\dot{X}^\nu = 0 \tag{1.2.13}$$

which we rewrite as

$$m\frac{d}{d\tau}\left(\frac{\eta_{\mu\nu}\dot{X}^\nu}{\sqrt{-\eta_{\lambda\rho}\dot{X}^\lambda\dot{X}^\rho}}\right) = qF_{\mu\nu}\dot{X}^\nu \tag{1.2.14}$$

To be more concrete we could choose static gauge again and we find

$$m\frac{d}{d\tau}\left(\frac{v^i}{\sqrt{1-v^2}}\right) = qF_{i0} + qF_{ij}v^j \tag{1.2.15}$$

Here we see the Lorentz force magnetic law arising as it should from the second term on the right hand side. The first term on the right hand side shows that an electric field provides a driving force.

At this point we should pause to mention a subtlety. In addition to (1.2.15) there is also the equation of motion for $X^0 = \tau$. However the reparameterization gauge symmetry implies that this equation is automatically satisfied. In particular the X^0 equation of motion is

$$-m\frac{d}{d\tau}\left(\frac{1}{\sqrt{1-v^2}}\right) = qF_{0i}v^i \tag{1.2.16}$$

Exercise 1 *Show that if* (1.2.15) *is satisfied then so is* (1.2.16)

Exercise 2 *Show that, in static gauge* $X^0 = \tau$, *the Hamiltonian for a charged particle is*

$$H = \sqrt{m^2 + (p^i - qA^i)(p^i - qA^i)} - qA_0. \tag{1.2.17}$$

1.2.3 Quantization

Next we'd like to quantize the point particle. This is made difficult by the highly non-linear form of the action. To this end we will consider a new action which is classically equivalent to the old one. In particular consider

$$S_{HT} = -\frac{1}{2}\int d\tau e\left(-\frac{1}{e^2}\dot{X}^\mu\dot{X}^\nu\eta_{\mu\nu} + m^2\right) \tag{1.2.18}$$

Here we have introduced a new field $e(\tau)$ which is non-dynamical, i.e. has no kinetic term. This action is now just quadratic in the fields X^μ. The point of it is that it reproduces the same equations of motion as before. To see this consider the e equation of motion:

$$\frac{1}{e^2}\dot{X}^\mu\dot{X}^\nu\eta_{\mu\nu} + m^2 = 0 \tag{1.2.19}$$

we can solve this to find $e = m^{-1}\sqrt{-\dot{X}^\mu\dot{X}^\nu\eta_{\mu\nu}}$. We now compute the X^μ equation of motion

$$
\begin{aligned}
0 &= \frac{d}{d\tau}\left(\frac{1}{e}\dot{X}^{\mu}\right) \\
&= m\frac{d}{d\tau}\left(\frac{\dot{X}^{\mu}}{\sqrt{-\dot{X}^{\lambda}\dot{X}^{\rho}\eta_{\lambda\rho}}}\right)
\end{aligned}
\tag{1.2.20}
$$

This is exactly what we want. Thus we conclude that S_{HT} is classically equivalent to S_{pp}.

One way to think about this action is that we have introduced a worldline metric $\gamma_{\tau\tau} = -e^2$ and its inverse $\gamma^{\tau\tau} = -1/e^2$ so that infinitesimal distances along the worldline have length

$$
ds^2 = \gamma_{\tau\tau} d\tau^2 \tag{1.2.21}
$$

Note that previously we never said that $d\tau$ represented the physical length of a piece of worldline, just that τ labeled points along the worldline.

There is another advantage to this form of the action; we can smoothly set $m^2 = 0$ and describe massless particles, which was impossible with the original form of the action.

Now the action is quadratic in the fields X^{μ} we calculate the Hamiltonian and quantize more easily. The first step here is to obtain the momentum conjugate to each of the X^{μ}

$$
\begin{aligned}
p_{\mu} &= \frac{\partial L}{\partial \dot{X}^{\mu}} \\
&= \frac{1}{e}\eta_{\mu\nu}\dot{X}^{\nu}
\end{aligned}
\tag{1.2.22}
$$

There is no conjugate momentum to e, it acts as a constraint and we will deal with it later. The Hamiltonian is

$$
\begin{aligned}
H &= p_{\mu}\dot{X}^{\mu} - L \\
&= \frac{e}{2}\left(\eta_{\mu\nu}p^{\mu}p^{\nu} + m^2\right)
\end{aligned}
\tag{1.2.23}
$$

To quantize this system we consider wavefunctions $\Psi(X, \tau)$ and promote X^{μ} and p_{μ} to the operators

$$
\hat{X}^{\mu}\Psi = X^{\mu}\Psi \qquad \hat{p}_{\mu}\Psi = -i\frac{\partial \Psi}{\partial X^{\mu}} \tag{1.2.24}
$$

We then arrive at the Schrödinger equation

$$
i\frac{\partial \Psi}{\partial \tau} = \frac{e}{2}\left(-\eta^{\mu\nu}\frac{\partial^2 \Psi}{\partial X^{\mu}\partial X^{\nu}} - m^2\Psi\right) \tag{1.2.25}
$$

Lastly we must deal with e which we saw has no conjugate momentum. Classically its equation of motion imposes the constraint

$$p^{\mu}p_{\mu} + m^2 = 0 \tag{1.2.26}$$

which is the mass-shell condition for the particle. Quantum mechanically we realize this by restricting our physical wavefunctions to those that satisfy the corresponding constraint

$$-\eta^{\mu\nu}\frac{\partial^2 \Psi}{\partial X^{\mu}\partial X^{\nu}} + m^2\Psi = 0 \tag{1.2.27}$$

However this is just the condition that $\hat{H}\Psi = 0$ so that the wavefunctions are τ independent. If you trace back the origin of this time-independence it arises as a consequence of the reparameterization invariance of the worldline theory. It simply states that wavefunctions must also be reparameterization invariant, i.e. they can't depend on τ. This is deep issue in quantum gravity. In effect it says that there is no such thing as time in the quantum theory.

This equation should be familiar if you have learnt quantum field theory. In particular if we consider a free scalar field Ψ in D-dimensional spacetime the action is

$$S = -\int d^D x \left(\frac{1}{2}\partial_{\mu}\Psi^{*}\partial^{\mu}\Psi + \frac{1}{2}m^2\Psi^{*}\Psi\right) \tag{1.2.28}$$

and the corresponding equation of motion is

$$\partial^2\Psi - m^2\Psi = 0 \tag{1.2.29}$$

which is the same as our Schrodinger equation (when restricted to the physical Hilbert space).

Thus we see that there is a natural identification of a free scalar field with a quantum point particle. In particular the quantum states of the point particle are in a one-to-one correspondence with the classical solutions of the free spacetime effective action. However one important difference should be stressed. The quantum point particle gave a Schrodinger equation which could be identified with the classical equation of motion for the scalar field. In quantum field theory one performs a second quantization whereby particles are allowed to be created and destroyed. This is beyond the quantization of the point particle that we considered since by default we studied the effective action on the worldline of a single particle: it would have made no sense to create or destroy particles. Thus the second quantized spacetime action provides a more complete physical theory.

Here we also can see that the quantum description of a point particle in one dimension leads to a classical spacetime effective action in D-dimensions. This is an important concept in string theory where the quantum dynamics of the two-dimensional worldvolume theory, with interactions included, leads to interesting and non-trivial spacetime effective actions.

Exercise 3 *Find the Schödinger equation, constraint and effective action for a quantized particle in the background of a classical electromagnetic field using the action*

$$S_{pp} = -\int \frac{1}{2}e\left(-\frac{1}{e^2}\dot{X}^{\mu}\dot{X}^{\nu}\eta_{\mu\nu} + m^2\right) - A_{\mu}\dot{X}^{\mu}. \tag{1.2.30}$$

1.3 Classical and Quantum Dynamics of Strings

1.3.1 Classical Action

Having studied point particles from their worldline perspective we now turn to our main subject: strings. Our starting point will be the action the worldvolume of a string, which is two-dimensional. The natural starting point is to consider the action

$$S_{string} = \frac{1}{2\pi\alpha'}\int d^2\sigma\sqrt{-\det(\partial_{\alpha}X^{\mu}\partial_{\beta}X^{\nu}\eta_{\mu\nu})} \tag{1.3.1}$$

which is simply the area of the two-dimensional worldvolume that the string sweeps out. Here σ^{α}, $\alpha = 0, 1$ labels the spatial and temporal coordinates of the string: τ, σ. Here $\sqrt{\alpha'}$ is a length scale that determines the size of the string.

Again we don't want to work directly with such a highly non-linear action. We saw above that we could change this by coupling to an auxiliary worldvolume metric $\gamma_{\alpha\beta}$.

Exercise 4 *Show that by solving the equation of motion for the metric $\gamma_{\alpha\beta}$ on a d-dimensional worldsheet the action*

$$S_{HT} = -\frac{1}{2}\int d^d\sigma\sqrt{-\gamma}\left(\gamma^{\alpha\beta}\partial_{\alpha}X^{\mu}\partial_{\beta}X^{\nu}\eta_{\mu\nu} + m^2(d-2)\right) \tag{1.3.2}$$

one finds the action

$$S_{NG} = m^{2-d}\int d^d\sigma\sqrt{-\det\left(\partial_{\alpha}X^{\mu}\partial_{\beta}X^{\nu}\eta_{\mu\nu}\right)} \tag{1.3.3}$$

for the remaining fields X^{μ}, i.e. calculate and solve the $\gamma_{\alpha\beta}$ equation of motion and then substitute the solution back into S_{HT} to obtain S_{NG}. Note that the action S_{HT} is often referred to as the Howe-Tucker form for the action whereas S_{NG} is the Nambu-Goto form. (Hint: You will need to use the fact that $\delta\sqrt{-\gamma}/\delta\gamma^{\alpha\beta} = -\frac{1}{2}\gamma_{\alpha\beta}\sqrt{-\gamma}$). If you have not yet learnt much about metrics just consider the case of $d=1$ where $\gamma_{\alpha\beta}$ just has a single component $\gamma_{\tau\tau}$.

So we might instead start with

$$S_{string} = -\frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{-\gamma}\gamma^{\alpha\beta}\partial_\alpha X^\mu\partial_\beta X^\nu\eta_{\mu\nu} \tag{1.3.4}$$

where we have taken $d=2$ in (1.3.2).

Exercise 5 *What transformation law must $\gamma_{\alpha\beta}$ have to ensure that S_{string} is reparameterization invariant? Hint: Use the fact that*

$$\frac{\partial\sigma'^\gamma}{\partial\sigma^\alpha}\frac{\partial\sigma^\beta}{\partial\sigma'^\gamma} = \delta^\beta_\alpha. \tag{1.3.5}$$

However this case is very special. If we evaluate the $\gamma_{\alpha\beta}$ equation of motion we find

$$T_{\alpha\beta} = \partial_\alpha X^\mu\partial_\beta X^\nu\eta_{\mu\nu} - \frac{1}{2}\gamma_{\alpha\beta}\gamma^{\gamma\delta}\partial_\gamma X^\mu\partial_\delta X^\nu\eta_{\mu\nu} = 0 \tag{1.3.6}$$

Once again we see that $\gamma_{\alpha\beta} = b\partial_\alpha X^\mu\partial_\beta X^\nu\eta_{\mu\nu}$ for some b. However in this case nothing fixes b, it is arbitrary. This occurs because there is an addition symmetry of the action that is unique to two-dimensions: it is conformally invariant. That means that under a worldvolume conformal transformation

$$\gamma_{\alpha\beta} \to e^{2\varphi}\gamma_{\alpha\beta} \tag{1.3.7}$$

(here φ is any function of the worldvolume coordinates) the action is invariant.

There are other features that are unique to two-dimensions. The first is that, up to a reparameterization, we can always choose the metric $\gamma_{\alpha\beta} = e^{2\rho}\eta_{\alpha\beta}$. To see this we note that under a reparameterization we have

$$\gamma'_{\alpha\beta} = \frac{\partial\sigma^\gamma}{\partial\sigma'^\alpha}\frac{\partial\sigma^\delta}{\partial\sigma'^\beta}\gamma_{\gamma\delta} \tag{1.3.8}$$

Thus we simply choose our new coordinates to fix $\gamma'_{01} = 0$ and $\gamma'_{00} = -\gamma'_{11}$. This requires that

$$\frac{\partial\sigma^\gamma}{\partial\sigma'^0}\frac{\partial\sigma^\delta}{\partial\sigma'^1}\gamma_{\gamma\delta} = 0 \tag{1.3.9}$$

and

$$\frac{\partial\sigma^\gamma}{\partial\sigma'^1}\frac{\partial\sigma^\delta}{\partial\sigma'^1}\gamma_{\gamma\delta} + \frac{\partial\sigma^\gamma}{\partial\sigma'^0}\frac{\partial\sigma^\delta}{\partial\sigma'^0}\gamma_{\gamma\delta} = 0 \tag{1.3.10}$$

These are two (complicated) differential equation for two functions $\sigma^0(\sigma'^0, \sigma'^1)$ and $\sigma^1(\sigma'^0, \sigma'^1)$. Hence there will be a solution (at least locally).

The second feature is that in two dimensions the Einstein equation

$$R_{\alpha\beta} - \frac{1}{2}\gamma_{\alpha\beta}R = 0 \tag{1.3.11}$$

vanishes identically. The reason for this is that in two dimensions there is only one independent component for the Riemann tensor: $R_{0101} = -R_{0110} = -R_{1001} = R_{1010}$. Therefore $R_{00} = R_{0101}\gamma^{11}$, $R_{11} = R_{0101}\gamma^{00}$ and $R_{01} = -R_{0101}\gamma^{01}$. Thus we see that

$$\begin{aligned} R &= 2R_{0101}(\gamma^{00}\gamma^{11} - \gamma^{01}\gamma^{01}) \\ &= 2R_{0101}\det(\gamma^{-1}) \\ &= \frac{2}{\det(\gamma)}R_{0101} \end{aligned} \tag{1.3.12}$$

Now we note that

$$\begin{pmatrix} \gamma^{00} & \gamma^{01} \\ \gamma^{01} & \gamma^{11} \end{pmatrix} = \frac{1}{\det(\gamma)}\begin{pmatrix} \gamma_{11} & -\gamma_{01} \\ -\gamma_{01} & \gamma_{00} \end{pmatrix} \tag{1.3.13}$$

and the result follows.

Thus Einstein's equation

$$R_{\alpha\beta} - \frac{1}{2}\gamma_{\alpha\beta}R = T_{\alpha\beta} \tag{1.3.14}$$

will imply that $T_{\alpha\beta} = 0$. Hence even if we include two-dimensional gravity the $\gamma_{\alpha\beta}$ equation of motion imposes the constraint that the energy-momentum tensor vanishes

$$T_{\alpha\beta} = \frac{\partial\mathscr{L}}{\partial\gamma^{\alpha\beta}} = 0 \tag{1.3.15}$$

Thus we can use worldsheet diffeomorphisms to set $\gamma_{\alpha\beta} = e^{2\rho}\eta_{\alpha\beta}$ and then use worldsheet conformal invariance to set $\rho = 0$, i.e. $\gamma_{\alpha\beta} = \eta_{\alpha\beta}$. This means that the worldvolume metric $\gamma_{\alpha\beta}$ actually decouples from the fields X^μ. This conformal invariance of two-dimensional gravity coupled to the embedding coordinates (viewed as scalar fields) will be our fundamental principle. It allows us to simply set $\gamma_{\alpha\beta} = \eta_{\alpha\beta}$. Thus to consider strings propagating in flat spacetime we use the action (known as the Polyakov action)

$$S_{string} = -\frac{1}{4\pi\alpha'}\int d^2\sigma\,\eta^{\alpha\beta}\partial_\alpha X^\mu\partial_\beta X^\nu\eta_{\mu\nu} \tag{1.3.16}$$

subject to the constraint (1.3.15) which becomes

$$\partial_\alpha X^\mu\partial_\beta X^\nu\eta_{\mu\nu} - \frac{1}{2}\eta_{\alpha\beta}\eta^{\gamma\delta}\partial_\gamma X^\mu\partial_\delta X^\nu\eta_{\mu\nu} = 0 \tag{1.3.17}$$

1.3.2 Spacetime Symmetries and Conserved Charges

We should also pause to outline how the spacetime symmetries lead to conserved currents and hence conserved charges in the worldsheet theory.

First we summarize Noether's theorem. Suppose that a Lagrangian $\mathscr{L}(\Phi_A, \partial_\alpha \Phi_A)$, where we denoted the fields by Φ_A, has a symmetry: $\mathscr{L}(\Phi_A) = \mathscr{L}(\Phi_A + \delta\Phi_A)$. This implies that

$$\frac{\partial \mathscr{L}}{\partial \Phi_A}\delta\Phi_A + \frac{\partial \mathscr{L}}{\partial(\partial_\alpha \Phi_A)}\delta\partial_\alpha\Phi_A = 0 \tag{1.3.18}$$

This allows us to construct a current:

$$J^\alpha = \frac{\partial \mathscr{L}}{\partial(\partial_\alpha \Phi_A)}\delta\Phi_A \tag{1.3.19}$$

which is conserved

$$\begin{aligned} \partial_\alpha J^\alpha &= \partial_\alpha\left(\frac{\partial \mathscr{L}}{\partial(\partial_\alpha \Phi_A)}\right)\delta\Phi_A + \frac{\partial \mathscr{L}}{\partial(\partial_\alpha \Phi_A)}\partial_\alpha\delta\Phi_A \\ &= \partial_\alpha\left(\frac{\partial \mathscr{L}}{\partial(\partial_\alpha \Phi_A)}\right)\delta\Phi_A - \frac{\partial \mathscr{L}}{\partial \Phi_A}\delta\Phi_A \\ &= 0 \end{aligned} \tag{1.3.20}$$

by the equation of motion. This means that the integral over space of J^0 is a constant defines a charge

$$Q = \int_{space} \sigma J^0 \tag{1.3.21}$$

which is conserved

$$\begin{aligned} \frac{dQ}{dt} &= \int_{space} \partial_0 J^0 \\ &= -\int_{space} \partial_i J^i \\ &= 0 \end{aligned}$$

Let us now consider the action

$$S_{string} = -\frac{1}{4\pi\alpha'}\int d^2\sigma\, \eta^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu \eta_{\mu\nu} \tag{1.3.22}$$

This has the spacetime Poincare symmetries: translations $\delta X^\mu = a^\mu$ and Lorentz transformations $\delta X^\mu = \Lambda^\mu_\nu X^\nu$. In the first case the conserved current is

$$P^\alpha_{a^\mu} = -\frac{1}{2\pi\alpha'}\partial^\alpha X_\mu a^\mu \tag{1.3.23}$$

The associated conserved charge is just the total momentum along the direction a^μ and in particular there are D independent choices

$$p_\mu = \frac{1}{2\pi\alpha'}\int d\sigma \dot{X}_\mu \tag{1.3.24}$$

We can also consider the spacetime Lorentz transformations which lead to the conserved currents

$$J^\alpha_\Lambda = -\frac{1}{2\pi\alpha'}\partial^\alpha X_\mu \Lambda^\mu_\nu X^\nu \tag{1.3.25}$$

The independent conserved charges are therefore given by (here $Q_\Lambda = \int d\sigma J^0_\Lambda = M^\mu{}_\nu \Lambda^\nu{}_\mu$)

$$M^\mu_\nu = \frac{1}{4\pi\alpha'}\int d\sigma \dot{X}^\mu X_\nu - X^\mu \dot{X}_\nu \tag{1.3.26}$$

The Poisson brackets of these worldsheet charges will, at least at the classical level, satisfy the algebra Poincare algebra. In the quantum theory they are lifted to operators that commute with the Hamiltonian.

1.3.3 *Quantization*

Next we wish to quantize this action. Unlike the point particle this action is a field theory in $(1+1)$-dimensions. As such we must use the quantization techniques of quantum field theory rather than simply constructing a Schrodinger equation. There are several ways to do this. The most modern way is the path integral formulation and Fadeev-Popov ghosts. However this requires some techniques that are possibly unfamiliar. So here we will use the method of canonical quantization.

Canonical quantization is essentially the Heisenberg picture of quantum mechanics where the fields X^μ and their conjugate momenta P_μ are promoted to operators which satisfy the equal time commutation relations

$$\begin{aligned}
[\hat{X}^\mu(\tau,\sigma), \hat{P}_\nu(\tau,\sigma')] &= i\delta(\sigma-\sigma')\delta^\mu_\nu \\
[\hat{X}^\mu(\tau,\sigma), \hat{X}^\nu(\tau,\sigma')] &= 0 \\
[\hat{P}_\mu(\tau,\sigma), \hat{P}_\nu(\tau,\sigma')] &= 0
\end{aligned} \tag{1.3.27}$$

as well as the Heisenberg equation

$$\frac{d\hat{X}^\mu}{d\tau} = i[\hat{H}, \hat{X}^\mu] \qquad \frac{d\hat{P}_\mu}{d\tau} = i[\hat{H}, \hat{P}_\mu] \tag{1.3.28}$$

In the case at hand we have

$$\hat{L} = \frac{1}{4\pi\alpha'}\int d\sigma\, \eta_{\mu\nu}\dot{\hat{X}}^{\mu}\dot{\hat{X}}^{\nu} - \eta_{\mu\nu}\hat{X}'^{\mu}\hat{X}'^{\nu} \tag{1.3.29}$$

hence

$$\hat{P}_{\mu} = \frac{1}{2\pi\alpha'}\eta_{\mu\nu}\dot{\hat{X}}^{\nu} \tag{1.3.30}$$

and

$$\begin{aligned}\hat{H} &= \int d\sigma\, \hat{P}_{\mu}\dot{\hat{X}}^{\mu} - \hat{L}\\ &= \int d\sigma 2\pi\alpha'\eta^{\mu\nu}\hat{P}_{\mu}\hat{P}_{\nu} - \int d\sigma\frac{1}{4\pi\alpha'}(2\pi\alpha')^2\eta^{\mu\nu}\hat{P}_{\mu}\hat{P}_{\nu} + \frac{1}{4\pi\alpha'}\eta_{\mu\nu}\hat{X}'^{\mu}\hat{X}'^{\nu}\\ &= \int d\sigma\pi\alpha'\eta^{\mu\nu}\hat{P}_{\mu}\hat{P}_{\nu} + \frac{1}{4\pi\alpha'}\eta_{\mu\nu}\hat{X}'^{\mu}\hat{X}'^{\nu}\end{aligned} \tag{1.3.31}$$

We can now calculate

$$\begin{aligned}\dot{\hat{X}}^{\mu}(\sigma) &= i[\hat{H}, \hat{X}^{\mu}(\sigma)]\\ &= \pi\alpha' i\int d\sigma'\eta^{\lambda\nu}[\hat{P}_{\lambda}(\sigma')\hat{P}_{\nu}(\sigma'), \hat{X}^{\mu}(\sigma)]\\ &= 2\pi\alpha' i\int d\sigma'\eta^{\lambda\nu}\hat{P}_{\lambda}(\sigma')[\hat{P}_{\nu}(\sigma'), \hat{X}^{\mu}(\sigma)]\\ &= 2\pi\alpha'\int d\sigma'\eta^{\lambda\nu}\hat{P}_{\lambda}(\sigma')\delta^{\mu}_{\nu}\delta(\sigma-\sigma')\\ &= 2\pi\alpha'\eta^{\mu\nu}\hat{P}_{\nu}(\sigma)\end{aligned} \tag{1.3.32}$$

which we already knew. But also we can now calculate

$$\begin{aligned}\dot{\hat{P}}_{\mu}(\sigma) &= i[\hat{H}, \hat{X}_{\mu}(\sigma)]\\ &= \frac{i}{4\pi\alpha'}\int d\sigma'\eta_{\lambda\nu}[\hat{X}'^{\lambda}(\sigma')\hat{X}'^{\nu}(\sigma'), \hat{P}_{\mu}(\sigma)]\\ &= \frac{i}{2\pi\alpha'}\int d\sigma'\eta_{\lambda\nu}\hat{X}'^{\lambda}(\sigma')[\hat{X}'^{\nu}(\sigma'), \hat{P}_{\mu}(\sigma)]\\ &= \frac{i}{2\pi\alpha'}\int d\sigma'\eta_{\lambda\nu}\hat{X}'^{\lambda}(\sigma')\frac{\partial}{\partial\sigma'}[\hat{X}^{\nu}(\sigma'), \hat{P}_{\mu}(\sigma)]\\ &= -\frac{i}{2\pi\alpha'}\int d\sigma'\eta_{\lambda\nu}\hat{X}''^{\lambda}(\sigma')[\hat{X}^{\nu}(\sigma'), \hat{P}_{\mu}(\sigma)]\\ &= \frac{1}{2\pi\alpha'}\int d\sigma'\eta_{\lambda\nu}\hat{X}''^{\lambda}(\sigma')\delta^{\nu}_{\mu}\delta(\sigma-\sigma')\\ &= \frac{1}{2\pi\alpha'}\eta_{\mu\nu}\hat{X}''^{\nu}(\sigma)\end{aligned} \tag{1.3.33}$$

or equivalently

$$-\ddot{\hat{X}}^{\mu} + \hat{X}''^{\mu} = 0 \tag{1.3.34}$$

Of course this is just the classical equation of motion reinterpreted in the quantum theory as an operator equation. In two-dimensions the solution to this is simply that

$$\hat{X}^{\mu} = \hat{X}^{\mu}_{L}(\tau + \sigma) + \hat{X}^{\mu}_{R}(\tau - \sigma) \tag{1.3.35}$$

i.e. we can split $\hat{X}^{\mu}$ into a left and right moving part.

To proceed we expand the string in a Fourier series

$$\hat{X}^{\mu} = x^{\mu} + \alpha' p^{\mu}\tau + \sqrt{\frac{\alpha'}{2}}i\sum_{n\neq 0}\left(\frac{a^{\mu}_{n}}{n}e^{-in(\tau+\sigma)} + \frac{\tilde{a}^{\mu}_{n}}{n}e^{-in(\tau-\sigma)}\right) \tag{1.3.36}$$

The various factors of n and α' will turn out to be useful later on. We have also included linear terms since $\hat{X}^{\mu}$ need not be periodic (more on this later). Or if you prefer

$$\begin{aligned}\hat{X}^{\mu}_{L} &= x^{\mu}_{L} + \frac{1}{2}\alpha' p^{\mu}(\tau + \sigma) + \sqrt{\frac{\alpha'}{2}}i\sum_{n\neq 0}\frac{a^{\mu}_{n}}{n}e^{-in(\tau+\sigma)}\\ \hat{X}^{\mu}_{R} &= x^{\mu}_{R} + \frac{1}{2}\alpha' p^{\mu}(\tau - \sigma) + \sqrt{\frac{\alpha'}{2}}i\sum_{n\neq 0}\frac{\tilde{a}^{\mu}_{n}}{n}e^{-in(\tau-\sigma)}\end{aligned} \tag{1.3.37}$$

Note that we have dropped the hat on the operators a^{μ} and $\tilde{a}^{\mu}$ since they will appear frequently. But don't forget that they are operators! Note also that we haven't yet said what n is, *e.g.* whether or not it is an integer, we will be more specific later. The a^{μ}_{n} and $\tilde{a}^{\mu}_{n}$ have the interpretation as left and right moving oscillators. Just as in quantum mechanics and quantum field theory these will be related to particle creation and annihilation operators.

Since X^{μ} is an observable we require that it is Hermitian in the quantum theory. This in turn implies that

$$(a^{\mu}_{n})^{\dagger} = a^{\mu}_{-n}, \qquad (\tilde{a}^{\mu}_{n})^{\dagger} = \tilde{a}^{\mu}_{-n} \tag{1.3.38}$$

and $(x^{\mu})^{\dagger} = x^{\mu}$, $(p^{\mu})^{\dagger} = p^{\mu}$. In this basis

$$\begin{aligned}\hat{P}^{\mu} &= \frac{1}{2\pi\alpha'}\dot{\hat{X}}^{\mu}\\ &= \frac{1}{2\pi\alpha'}\left(\alpha' p^{\mu} + \sqrt{\frac{\alpha'}{2}}\sum_{-n\neq 0}a^{\mu}_{n}e^{-in(\tau+\sigma)} + \sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\tilde{a}^{\mu}_{n}e^{-in(\tau-\sigma)}\right)\end{aligned} \tag{1.3.39}$$

We can work out the commutator. First we take $x^\mu = p^\mu = 0$

$$
\begin{aligned}
[\hat{X}^\mu(\tau,\sigma), \hat{P}_\nu(\tau,\sigma')] &= \frac{i}{4\pi}\sum_n\sum_m \frac{1}{n} e^{-i(n+m)\tau} e^{-i(n\sigma+m\sigma')}[a_n^\mu, a_m^\nu] \\
&+ \frac{i}{4\pi}\sum_n\sum_m \frac{1}{n} e^{-i(n+m)\tau} e^{i(n\sigma+m\sigma')}[\tilde{a}_n^\mu, \tilde{a}_m^\nu] \\
&+ \frac{i}{4\pi}\sum_n\sum_m \frac{1}{n} e^{-i(n+m)\tau} e^{i(n\sigma-m\sigma')}[\tilde{a}_n^\mu, a_m^\nu] \\
&+ \frac{i}{4\pi}\sum_n\sum_m \frac{1}{n} e^{-i(n+m)\tau} e^{-i(n\sigma-m\sigma')}[a_n^\mu, \tilde{a}_m^\nu] \quad (1.3.40)
\end{aligned}
$$

In order for the τ-dependent terms to cancel we see that we need the commutators to vanish if $n \neq -m$. The sum now reduces to

$$
\begin{aligned}
[\hat{X}^\mu(\tau,\sigma), \hat{P}^\nu(\tau,\sigma')] &= \frac{i}{4\pi}\sum_n \frac{1}{n} e^{-in(\sigma-\sigma')}[a_n^\mu, a_{-n}^\nu] \\
&+ \frac{i}{4\pi}\sum_n \frac{1}{n} e^{in(\sigma-\sigma')}[\tilde{a}_n^\mu, \tilde{a}_{-n}^\nu] \\
&+ \frac{i}{4\pi}\sum_n \frac{1}{n} e^{in(\sigma+\sigma')}[\tilde{a}_n^\mu, a_{-n}^\nu] \\
&+ \frac{i}{4\pi}\sum_n \frac{1}{n} e^{-in(\sigma+\sigma')}[a_n^\mu, \tilde{a}_{-n}^\nu] \quad (1.3.41)
\end{aligned}
$$

Next translational invariance implies that the $\sigma + \sigma'$ terms vanish and hence

$$[a_n^\mu, \tilde{a}_m^\nu] = 0 \quad (1.3.42)$$

A slight rearrangement of indices shows that we are left with

$$[\hat{X}^\mu(\tau,\sigma), \hat{P}^\nu(\tau,\sigma')] = \frac{i}{4\pi}\sum_n \frac{1}{n} e^{-in(\sigma-\sigma')}([a_n^\mu, a_{-n}^\nu] + [\tilde{a}_n^\mu, \tilde{a}_{-n}^\nu]) \quad (1.3.43)$$

In a Fourier basis

$$\delta(\sigma - \sigma') = \frac{1}{2\pi}\sum_n e^{-in(\sigma-\sigma')} \quad (1.3.44)$$

Note that there is a contribution from $n = 0$ here that doesn't come from the oscillators, we'll deal with it in a moment. Therefore we see that we must take

$$[a_n^\mu, a_m^\nu] = n\eta^{\mu\nu}\delta_{n,-m}, \qquad [\tilde{a}_n^\mu, \tilde{a}_m^\nu] = n\eta^{\mu\nu}\delta_{n,-m} \quad (1.3.45)$$

Next it remains to consider the zero-modes (including the $n = 0$ contribution in (1.3.44)).

Exercise 6 *Show that if* $x^\mu, p^\mu \neq 0$ *then we also have*

$$[x^\mu, p^\nu] = i\eta^{\mu\nu} \tag{1.3.46}$$

with the other commutators vanishing.

We also have to consider the constraint $\hat{T}_{\alpha\beta} = 0$. Its components are

$$\begin{aligned}
\hat{T}_{00} &= \frac{1}{2}\dot{\hat{X}}^\mu \dot{\hat{X}}^\nu \eta_{\mu\nu} + \frac{1}{2}\hat{X}'^\mu \hat{X}'^\nu \eta_{\mu\nu} \\
\hat{T}_{11} &= \frac{1}{2}\hat{X}'^\mu P \hat{X}'^\nu \eta_{\mu\nu} + \frac{1}{2}\dot{\hat{X}}^\mu \dot{\hat{X}}^\nu \eta_{\mu\nu} \\
\hat{T}_{01} &= \dot{\hat{X}}^\mu \hat{X}'^\nu \eta_{\mu\nu}
\end{aligned} \tag{1.3.47}$$

It is helpful to change coordinates to

$$\begin{pmatrix} \sigma^+ = \tau + \sigma \\ \sigma^- = \tau - \sigma \end{pmatrix} \iff \begin{pmatrix} \tau = \frac{\sigma^+ + \sigma^-}{2} \\ \sigma = \frac{\sigma^+ - \sigma^-}{2} \end{pmatrix} \tag{1.3.48}$$

Exercise 7 *Show that in these coordinates*

$$\begin{aligned}
\hat{T}_{++} &= \partial_+ \hat{X}^\mu \partial_+ \hat{X}^\nu \eta_{\mu\nu} \\
\hat{T}_{--} &= \partial_- \hat{X}^\mu \partial_- \hat{X}^\nu \eta_{\mu\nu} \\
\hat{T}_{+-} &= T_{-+} = 0
\end{aligned} \tag{1.3.49}$$

Let us now calculate T_{++} in terms of oscillators. We have

$$\partial_+ \hat{X}^\mu = \sqrt{\frac{\alpha'}{2}} \sum_{n=-\infty}^{\infty} a_n^\mu e^{-in(\tau+\sigma)} \tag{1.3.50}$$

where we have introduced

$$a_0^\mu = \sqrt{\frac{\alpha'}{2}} p^\mu + \sqrt{\frac{1}{2\alpha'}} w^\mu \tag{1.3.51}$$

thus

$$\begin{aligned}
\hat{T}_{++} &= \frac{\alpha'}{2} \sum_{nm} a_n^\mu a_m^\nu e^{-i(n+m)(\tau+\sigma)} \eta_{\mu\nu} \\
&= \alpha' \sum_n L_n e^{-in(\tau+\sigma)}
\end{aligned} \tag{1.3.52}$$

with

$$L_n = \frac{1}{2} \sum_m a_{n-m}^\mu a_m^\nu \eta_{\mu\nu} \tag{1.3.53}$$

where again we've dropped a hat on L_n, even though it is an operator. Similarly we find

$$T_{--} = \alpha' \sum_n \tilde{L}_n e^{-in(\tau-\sigma)} \tag{1.3.54}$$

with

$$\tilde{L}_n = \frac{1}{2}\sum_m \tilde{a}^\mu_{n-m}\tilde{a}^\nu_m \eta_{\mu\nu} \tag{1.3.55}$$

and

$$\tilde{a}^\mu_0 = \sqrt{\frac{\alpha'}{2}}p^\mu - \sqrt{\frac{2}{\alpha'}}w^\mu \tag{1.3.56}$$

We can rewrite the commutators (1.3.45) using (1.3.38) as

$$[a^\mu_n, a^{\nu\dagger}_n] = n\eta^{\mu\nu} \qquad [\tilde{a}^\mu_n, \tilde{a}^{\nu\dagger}_n] = n\eta^{\mu\nu} \tag{1.3.57}$$

with $n > 0$. Thus we can think of a^μ_n and $\tilde{a}^\mu_n$ annihilation operators and $a^{\mu\dagger}_n$ and $\tilde{a}^{\mu\dagger}_n$ as creation operators. Following the standard practice of QFT we consider the ground state $|0\rangle$ to be annihilated by a_n and $\tilde{a}_n$:

$$a_n|0\rangle = 0, \qquad \tilde{a}_n|0\rangle = 0, \qquad n > 0 \tag{1.3.58}$$

The zero modes also act on the ground state. Since x^μ and p^μ don't commute we can only chose $|0\rangle$ to be an eigenstate of one, we take

$$\hat{p}^\mu|0\rangle = p^\mu|0\rangle \tag{1.3.59}$$

when we want to be precise we label the ground state $|0; p\rangle$. You will have to excuse the clumsy notion where I have reintroduce a hat on an operator to distinguish it from its eigenvalue acting on a state. We can now construct a Fock space of multi-particle states by acting on the ground state with the creation operators a^μ_{-n} and $\tilde{a}^\mu_{-n}$. For example

$$a^\mu_{-1}\tilde{a}^\nu_{-1}|0\rangle, \qquad a^\mu_{-2}\tilde{a}^\lambda_{-1}\tilde{a}^\rho_{-1}|0\rangle, \qquad etc. \tag{1.3.60}$$

These elements should be familiar from the study of the harmonic oscillator. In a string theory each classical vibrational mode is mapped in the quantum theory to an individual harmonic oscillator with the same frequency.

Note that we really should considering normal ordered operators, where the annihilation operators always appear to the right of the creation operators. For L_n and $\tilde{L}_n$ with $n \neq 0$ there is no ambiguity as a^μ_m and a^ν_{n-m} will commute. However for L_0 and $\tilde{L}_0$ one finds

$$L_0 = \frac{1}{2}a_0^\mu a_0^\nu \eta_{\mu\nu} + \sum_{m>0} a_{-m}^\mu a_m^\nu \eta_{\mu\nu} - \frac{1}{2}\sum_{m>0}[a_{-m}^\mu, a_m^\nu]\eta_{\mu\nu} \tag{1.3.61}$$

The last term is an infinite divergent sum

$$\frac{D}{2}\sum_{m>0} m \tag{1.3.62}$$

This can be thought of as sum over the zero-point energies of the infinite number of harmonic oscillators. We must renormalize. Clearly $\tilde{L}_0$ has the same problem and this introduces the same sum. Since this is just a number the end result is that we define the normal ordered L_0 and $\tilde{L}_0$ to be

$$\begin{aligned} :L_0: &= \frac{1}{2}a_0^\mu a_0^\nu \eta_{\mu\nu} + \alpha' \sum_{m>0} a_{-m}^\mu a_m^\nu \eta_{\mu\nu} \\ :\tilde{L}_0: &= \frac{1}{2}\tilde{a}_0^\mu \tilde{a}_0^\nu \eta_{\mu\nu} + \alpha' \sum_{m>0} \tilde{a}_{-m}^\mu \tilde{a}_m^\nu \eta_{\mu\nu} \end{aligned} \tag{1.3.63}$$

In string theory $:L_n:$ and $:\tilde{L}_n:$ play a central role.

How do we deal with constraints in the quantum theory? We should proceed by reducing to the so-called physical Hilbert space of states which are those states that are animated by $:\hat{T}_{\alpha\beta}:$. However this turns out to be too strong a condition and would remove all states. Instead we impose that the positive frequency components of $:\hat{T}_{\alpha\beta}:$ annihilates any physical state

$$:L_n:|phys\rangle =:\tilde{L}_n:|phys\rangle = 0, n>0 \qquad (:L_0:-a)|phys\rangle = (:\tilde{L}_0:-a)|phys\rangle = 0 \tag{1.3.64}$$

Here we have introduced a parameter a since $:L_0:$ differs from L_0 by an infinite constant that we must regularize to the finite value a. For historical reasons the parameter a is called the intercept (and α' the slope). However it is not a parameter but rather is fixed by consistency conditions. Indeed it can be calculated by a variety of methods (such as ζ-function regularization or by using the modern BRST approach to quantization). We will see that the correct value is $a = 1$.

This is then sufficient to show that the expectation value of $:\hat{T}_{\alpha\beta}:$ vanishes

$$\langle phys| :L_n: |phys\rangle = \langle phys| :\tilde{L}_n: |phys\rangle = 0 \qquad \forall n \neq 0 \tag{1.3.65}$$

since the state on the right is annihilated by the postiche frequency parts where as by taking the Hermitian conjugates one sees that the state on the left is annihilated by the negative frequency part.

It is helpful to calculate the commutator $[:L_m:, :L_n:]$. There will be a similar expression for $[:\tilde{L}_m:, :\tilde{L}_m:]$ and clearly one has $[:L_m:, :\tilde{L}_n:] = 0$. To do this we first consider the case without worrying about normal orderings

$$
\begin{aligned}
[L_m,\ L_n] &= \frac{1}{4}\sum_{pq}[a^{\mu}_{m-p}a^{\nu}_{p},\, a^{\lambda}_{n-q}a^{\rho}_{q}]\eta_{\mu\nu}\eta_{\lambda\rho} \\
&= \frac{1}{4}\sum_{pq}\eta_{\mu\nu}\eta_{\lambda\rho}\Big([a^{\mu}_{m-p}a^{\nu}_{p},\, a^{\lambda}_{n-q}]a^{\rho}_{q} + a^{\lambda}_{n-q}[a^{\mu}_{m-p}a^{\nu}_{p},\, a^{\rho}_{q}]\Big) \\
&= \frac{1}{4}\sum_{pq}\eta_{\mu\nu}\eta_{\lambda\rho}\Big(a^{\mu}_{m-p}[a^{\nu}_{p},\, a^{\lambda}_{n-q}]a^{\rho}_{q} + [a^{\mu}_{m-p},\, a^{\lambda}_{n-q}]a^{\nu}_{p}a^{\rho}_{q} \\
&\quad + a^{\lambda}_{n-q}a^{\mu}_{m-p}[a^{\nu}_{p},\, a^{\rho}_{q}] + a^{\lambda}_{n-q}[a^{\mu}_{m-p},\, a^{\rho}_{q}]a^{\nu}_{p}\Big) \\
&= \frac{1}{4}\sum_{p}\eta_{\mu\rho}\Big(pa^{\mu}_{m-p}a^{\rho}_{n+p} + (m-p)a^{\mu}_{p}a^{\rho}_{n+m-p} \\
&\quad + pa^{\rho}_{n+p}a^{\mu}_{m-p} + (m-p)a^{\rho}_{n+m-p}a^{\mu}_{p}\Big) \\
&= \frac{1}{2}\sum_{p}\eta_{\mu\rho}\Big((p-n)a^{\mu}_{m+n-p}a^{\rho}_{p} + (m-p)a^{\mu}_{p}a^{\rho}_{n+m-p}\eta_{\mu\rho}\Big)
\end{aligned}
\tag{1.3.66}
$$

Here we have used the identities

$$
[A, BC] = [A, B]C + B[A, C], \qquad [AB, C] = A[B, C] + [A, C]B \tag{1.3.67}
$$

and shifted the $p =$ variable in the sum. Thus we find

$$
[L_m, L_n] = (m-n)L_{m+n} \tag{1.3.68}
$$

This is called the classical Virasoro algebra and is of crucial importance in string theory and conformal field theory in general. Recall that it is the algebra of constraints that arose from the condition $\hat{T}_{\alpha\beta} = 0$ which is the statement of conformal invariance.

In the quantum theory we must consider the issues associated with normal ordering. We saw that this only affected $: L_0 :$. It follows that the only effect this can have on the Virasoro algebra is in terms with an $: L_0 :$. Since the effect on $: L_0 :$ is a shift by an infinite constant it won't appear in the commutator on the left hand side. Thus any new terms can only appear with $: L_0 :$ on the right hand side. Thus the general form is

$$
[: L_m :, : L_n :] = (m-n) : L_{m+n} : + C(n)\delta_{m-n} \tag{1.3.69}
$$

The easiest way to determine the $C(n)$ is to note the following (one can also perform a direct calculation but it is notoriously complicated and messy). First one imposes the Jacobi identity

$$
[: L_k :, [: L_m :, : L_n :]] + [: L_m :, [: L_n :, : L_k :]] + [: L_m :, [: L_n :, : L_k :]] = 0 \tag{1.3.70}
$$

If we impose that $k + m + n = 0$ with $k, m, n \neq 0$ (so that no pair of them adds up to zero) then this reduces to

$$(m-n)C(k)+(n-k)C(m)+(k-m)C(n)=0 \tag{1.3.71}$$

If we pick $k=1$ and $m=-n-1$ one finds

$$-(2n+1)C(1)+(n-1)C(-n-1)+(n+2)C(n)=0 \tag{1.3.72}$$

Now we note that $C(-n)=-C(n)$ by definition. Hence we learn that $C(0)=0$ and

$$C(n+1)=\frac{(n+2)C(n)-(2n+1)C(1)}{n-1} \tag{1.3.73}$$

This is just a difference equation and given $C(2)$ it will determine $C(n)$ for $n>1$ (note that it can't determine $C(2)$ given $C(1)$). We can look for a solution to this by considering polynomials. Since it must be odd in n the simplest guess is

$$c(n)=c_1n^3+c_2n \tag{1.3.74}$$

In this case the right hand side becomes

$$\begin{aligned}&\frac{(n+1)(c_1n^3+c_2n)-(2n+1)(c_1+c_2)}{n-1}\\ &=\frac{c_1n^4+2c_1n^3+c_2n^2-2c_1n-(c_1+c_2)}{n-1}\\ &=\frac{(n-1)(c_1n^3+3c_1n^2+(3c_1+c_2)n+c_1+c_2)}{n-1}\end{aligned} \tag{1.3.75}$$

Expanding out the left hand side gives

$$c_1(n+1)^3+c_2(n+1)=c_1n^3+3c_1n^2+(3c_1+c_2)n+c_1+c_2 \tag{1.3.76}$$

and hence they agree.

Note that if we shift L_0 by a constant l then $C(n)$ is shifted by $2nl$ (note that in so doing we'd have to shift a as well). This means that we can change the value of c_2. Therefore we will fix it to be $c_1=-c_2$. Finally we must calculate c_1. To do this we consider the ground state with no momentum $|0;0,0\rangle$ This state is annihilated by $:L_n:$ for all $n\geq 0$. Thus we have

$$\begin{aligned}\langle 0,0;0|:L_2::L_{-2}:|0;0,0\rangle &= \langle 0,0;0|[:L_2:,:L_{-2}:]|0;0,0\rangle\\ &=4\langle 0,0;0|:L_0:|0;0,0\rangle+6c_1\langle 0,0;0|0;0,0\rangle\\ &=6c_1\end{aligned} \tag{1.3.77}$$

where we assume that the ground state has unit norm.

Exercise 8 *Show that*

$$\langle 0,0;0|:L_2::L_{-2}:|0;0,0\rangle=\frac{D}{2} \tag{1.3.78}$$

So we deduce that

$$[: L_m :, : L_n :] = (m - n) : L_{m+n} : + \frac{D}{12}(m^3 - m)\delta_{m-n} \tag{1.3.79}$$

Of course there is a similar expression for $[: \tilde{L}_m :, : \tilde{L}_m :]$. This is called the central extension of the Virasoro algebra and D is the central charge which has arisen as a quantum effect. From now on we will always take operators to be normal ordered and we will drop the :: symbol, unless otherwise stated.

Let us return to our Fock space of states. It is built up out of the ground state which we take to have unit norm $\langle 0|0\rangle = 1$. One sees that the one-particle state $a^\mu_{-1}|0\rangle$ has norm

$$\langle 0|a^\mu_1 a^\mu_{-1}|0\rangle = \langle 0|[a^\mu_1, a^\mu_{-1}]|0\rangle = \eta^{\mu\mu} \tag{1.3.80}$$

where we do not sum over μ. Thus the state $a^0_{-1}|0\rangle$ has negative norm!

Problem Show that the state $(a^0_{-1} + a^1_{-1})|0\rangle$ has zero norm.

Thus the natural inner product on the Fock space is not positive definite because the time-like oscillators come with the wrong sign. This also occurs in other quantum theories such as QED and doesn't necessarily represent any kind of sickness.

There are stranger states still. A physical state $|\chi\rangle$ that satisfies $\langle\chi|phys\rangle = 0$ for all physical states is called null (or spurious if it only satisfies the $n = 0$ physical state condition). It then follows that a null state has zero norm (as it must be orthogonal to itself).

There are many such states. To construct an example just consider

$$|\chi\rangle = L_{-1}|0; p\rangle \text{ with } p^2 = 0 \tag{1.3.81}$$

Note that the zero-momentum ground state satisfies $L_n|0; 0\rangle = 0$ and for all $n \geq 0$ and this remains true if for $|0; p\rangle$ if $p^2 = 0$. First we verify that $|\chi\rangle$ is physical. We have for $m \geq 0$

$$\begin{aligned} L_m|\chi\rangle &= L_m L_{-1}|0; p\rangle \\ &= [L_m, L_{-1}]|0; p\rangle \\ &= (m+1)L_{m-1}|0; p\rangle + \frac{D}{12}(m^3 - m)\delta_{m1}|0; p\rangle \end{aligned} \tag{1.3.82}$$

The last term will vanish automatically whereas the first term can only be non-zero for $m = 0$ (since $L_n|0; p\rangle = 0$ for all $n \geq 0$). Here we find $L_0|\chi\rangle = |\chi\rangle$ which is the physical state condition for $a = 1$ which will turn out to be the case. Next we see that $\langle\chi|phys\rangle = \langle 0|L_1|phys\rangle = 0$. Note that we could have used any state instead of $|0; p\rangle$ that was annihilated by L_n for all $n \geq 0$ to construct a null state.

Thus if we calculate some amplitude between two physical states $\langle phys'|phys\rangle$ we can shift $|phys\rangle \to |phys\rangle + |\chi\rangle$ where $|\chi\rangle$ is a null state. The new state $|phys\rangle + |\chi\rangle$ is still physical but the amplitude will remain the same—for any other choice of physical state $|phys'\rangle$. Thus we have a stringy gauge symmetry whereby two physical states are equivalent if their difference is a null state. This will turn out to be the origin of Yang-Mills and other gauge symmetries within string theory. And furthermore one can prove a no-ghost theorem which asserts that there are no physical states with negative norm (at least for a=1 and D=26).

1.3.4 Open Strings

Strings come in two varieties: open and closed. To date we have tried to develop as many formulae and results as possible which apply to both. However now we must make a decision and proceed along slightly different but analogous roots. Open strings have two end points which traditionally arise at $\sigma = 0$ and $\sigma = \pi$. We must be careful to ensure that the correct boundary conditions are imposed. In particular we must choose boundary conditions so that the boundary value problem is well defined. This requires that

$$\eta_{\mu\nu}\delta X^{\mu}\partial_{\sigma}X^{\nu} = 0 \tag{1.3.83}$$

at $\sigma = 0, \pi$.

Problem Show this!

There are essentially two boundary conditions that one can impose. The first is Dirichlet: we hold X^{μ} fixed at the end points so that $\delta X^{\mu} = 0$. The second is Neumann: we set $\partial_{\sigma}X^{\mu} = 0$ at the end points. The first condition implies that somehow the end points of the string are fixed in spacetime, like a flag to a flag pole. At first glance this seems unphysical and we will ignore it for now, although such boundary conditions turn out to be profoundly important. So we will start by considering second boundary condition, which states that no momentum leaks off the ends of the string.

The condition that $\partial_{\sigma}\hat{X}^{\mu}(\tau, 0) = 0$ implies that

$$a_n^{\mu} = \tilde{a}_n^{\mu} \tag{1.3.84}$$

i.e. the left and right oscillators are not independent. If we look at the boundary condition at $\sigma = \pi$ then we determine that

$$\sum_{n\neq 0} a_n^{\mu} e^{in\tau}\sin(n\pi) = 0 \tag{1.3.85}$$

Thus n is indeed an integer. The mode expansion is therefore

$$X^{\mu} = x^{\mu} + 2\alpha' p^{\mu}\tau + \sqrt{2\alpha'}i\sum_{n\neq 0}\frac{a_n^{\mu}}{n}e^{in\tau}\cos(n\sigma) \tag{1.3.86}$$

(Note the slightly redefined value of p^{μ} as compared to before.)

For the open string the physical states are constrained to satisfy

$$L_n|phys\rangle = 0, n > 0 \quad \text{and} \quad (L_0 - 1)|phys\rangle = 0 \tag{1.3.87}$$

in particular there is only one copy of the constraints required since the $\tilde{L}_n$ constraints will automatically be satisfied. The second condition is the most illuminating as it gives the spacetime mass shell condition. To see this we note that translational invariance $X^\mu \to X^\mu + x^\mu$ gives rise to the conserved current $\hat{P}^\mu = \frac{1}{2\pi\alpha'}\dot{X}^\mu$. This is a worldsheet current and hence the conserved charge (from the worldsheet point of view) is

$$\begin{aligned} p^\mu &= \frac{1}{2\pi\alpha'}\int_0^\pi d\sigma \dot{X}^\mu \\ &= \frac{1}{2\pi\alpha'}\int_0^\pi d\sigma 2p^\mu + \sqrt{2\alpha'}\sum_{n\neq 0} a_n^\mu e^{in\tau}\cos(n\sigma) \\ &= p^\mu \end{aligned} \tag{1.3.88}$$

where again we have abused notation and confused the operator $\hat{p}^\mu$ that appears in the mode expansion of X^μ with its eigenvalue p^μ which we have now identified with the conserved charge. In any case we do this because we have shown that p^μ is indeed the spacetime momentum of the string. Note that this also explains why we put in the extra factor of 2 in front of $p^\mu\tau$ in the mode expansion.

Next we let

$$N = \sum_{n>0}\eta_{\mu\nu}a_{-n}^\mu a_n^\nu \tag{1.3.89}$$

Which is the analogue of the number operator that appears in the Harmonic oscillator. Again this is an operator even though we are being lazy and dropping the hat. It is easy to see that for $m > 0$

$$\begin{aligned} [N, a_{-m}^\lambda] &= \sum_{n>0}\eta_{\mu\nu}a_{-n}^\mu[a_n^\nu, a_{-m}^\lambda] \\ &= m a_{-m}^\lambda \end{aligned} \tag{1.3.90}$$

Thus if $|n\rangle$ is a state with $N|n\rangle = n|n\rangle$ then

$$\begin{aligned} N a_{-m}^\lambda|n\rangle &= ([N, a_{-m}^\lambda] + a_{-m}^\lambda N)|n\rangle \\ &= (m a_{-m}^\lambda + a_{-m}^\lambda n)|n\rangle \\ &= (m+n)a_{-m}^\lambda|n\rangle \end{aligned} \tag{1.3.91}$$

Therefore $a_{-m}^\lambda|n\rangle$ is a state with N-eigenvalue $n+m$. You can think of N as counting the number of oscillator modes in a given state.

With this definition we can write the physical state condition $(L_0 - 1)|phys\rangle = 0$ as

$$(p_\mu p^\mu + \frac{1}{\alpha'}(N-1))|phys\rangle = 0 \qquad (1.3.92)$$

Thus we can identify the spacetime mass-squared of a physical state to be the eigenvalue of

$$M^2 = \frac{1}{\alpha'}(N-1) \qquad (1.3.93)$$

We call the eigenvalue of N the level of the state. In other words the higher oscillator modes give more and more massive states in spacetime. In practice one takes $a^{-1/2}$ to be a very high mass scale so that only the massless modes are physically relevant. Note that the number of states at level n grows exponentially in n as the number of possible oscillations will be of order of the number of partitions of n into smaller integers. This exponentially growing tower of massive modes a unique feature of strings as opposed to point particles.

Of course we must not forget the other physical state condition $L_n|phys\rangle = 0$ for $n > 0$. This constraint will take the form of a gauge fixing condition . Let us consider the lowest lying states.

At level zero we have the vacuum $|0; p\rangle$. We see that the mass-shell condition is

$$p^2 - \alpha'^{-1} = 0 \qquad (1.3.94)$$

The other constraint, $L_n|0; p\rangle = 0$ with $n > 0$, is automatically satisfied. This has a negative mass-squared! Such a mode is called a tachyon. Tachyons arise in field theory if rather than expanding a scalar field about a minimum of the potential one expands about a maximum. Thus they are interpreted as instabilities. The problem is that no one knows in general whether or not the instability associated to this open string tachyon is ever stabilized. We will simply ignore the tachyon. Our reason for doing this is that it naturally disappears once one includes worldsheet fermions and considers the superstring theories. However the rest of the physics of bosonic strings remains useful in the superstring. Hence we continue to study it.

Next consider level 1. Here we have

$$|A_\mu\rangle = A_\mu(p)a^\mu_{-1}|0; p\rangle \qquad (1.3.95)$$

Since these modes have $N = 1$ it follows from the mass shell condition that they are massless (for $a = 1$!), i.e. the L_0 constraint implies that $p^2 A_\mu = 0$. Note that this depends crucially on the fact that $a = 1$. If $a > 1$ then $|A_\mu\rangle$ would be tachyonic whereas if $a\langle 1\ |A_\mu\rangle$ would be massive. In either case there is no known constituent theory of a massive (or tachyonic) vector field.

But we must also check that $L_n|A\rangle = 0$ for $n > 0$. Thus

$$\begin{aligned} L_n|A_\mu\rangle &= \frac{1}{2}A_\mu \sum_m \eta_{\nu\lambda} a^\nu_{n-m} a^\lambda_m a^\mu_{-1}|0; p\rangle \\ &= \frac{1}{2}A_\mu \eta_{\nu\lambda} \sum_{m\leq 1} a^\nu_{n-m} a^\lambda_m a^\mu_{-1}|0; p\rangle \\ &= \frac{1}{2}A_\mu \eta_{\nu\lambda} \sum_{n-1\leq m\leq 1} a^\nu_{n-m} a^\lambda_m a^\mu_{-1}|0; p\rangle \end{aligned} \tag{1.3.96}$$

In the second line we've noted that if $m > 1$ we can safely commute a^λ_m past a^μ_{-1} where it annihilates the vacuum. In the third line we've observed that if $n - m > 1$ then we can safely commute a^ν_{n-m} through the other two oscillators to annihilate the vacuum (recall that for $n > 0$ a^ν_{n-m} always commutes through a^λ_m). Thus for $n > 1$ we automatically have $L_n|A_\mu\rangle = 0$. For $n = 1$ we find just two terms

$$\begin{aligned} L_1|A\rangle &= \frac{1}{2}A_\mu \eta_{\nu\lambda}(a^\nu_1 a^\lambda_0 a^\mu_{-1} + a^\nu_0 a^\lambda_1 a^\mu_{-1})|0; p\rangle \\ &= A_\mu a^\mu_0|0; p\rangle \\ &= \sqrt{2\alpha'}p^\mu A_\mu|0; p\rangle \end{aligned} \tag{1.3.97}$$

Thus we see that $|A_\mu\rangle$ is represent a massless vector mode with $p^\mu A_\mu = 0$. In position space this is just $\partial^\mu A_\mu = 0$ and this looks like the Lorentz gauge condition for an electromagnetic potential.

Indeed recall that before we found the null state, with $p^2 = 0$,

$$\begin{aligned} |\Lambda\rangle &= i\Lambda(p)L_{-1}|0; p\rangle \\ &= i\eta_{\mu\nu}\Lambda a^\mu_0 a^\nu_{-1}|0; p\rangle \\ &= i\sqrt{2\alpha'}p_\mu \Lambda a^\mu_{-1}|0; p\rangle \end{aligned} \tag{1.3.98}$$

provided that $p^2 = 0$. Thus we must identify $A_\mu \equiv A_\mu + i\sqrt{2\alpha'}p_\mu\Lambda$ which in position space is the electromagnetic gauge symmetry $A_\mu \equiv A_\mu + \sqrt{2\alpha'}\partial_\mu\Lambda$. Again this occurs precisely when $a = 1$, otherwise $L_{-1}|0; p\rangle$ is not a null state and their would not be a gauge symmetry.

There is one more thing that can be done. Since and open string has two preferred points, its end points, we can attach discrete labels to the end points so that the ground state, of the open string carries two indices

$$|0; p, ab\rangle \tag{1.3.99}$$

where $a = 1, .., N$ refers the $\sigma = 0$ end and $b = 1, \ldots, N$ refers to the $\sigma = \pi$ end. It then follows that all the Fock space elements built out of $|0; p, ab\rangle$ will carry these indices. These are called Chan–Paton indices. The level one states now have the form

$$|A_\mu^{ab}\rangle = A_\mu^{ab} a_{-1}^\mu |0; p, ab\rangle \tag{1.3.100}$$

The null states take the form

$$|\Lambda^{ab}\rangle = i\Lambda^{ab} L_{-1}|0; p, ab\rangle \tag{1.3.101}$$

and the gauge symmetry is

$$A_\mu^{ab} \equiv A_\mu^{ab} + \sqrt{2\alpha'}\partial_\mu \Lambda^{ab} \tag{1.3.102}$$

These are the gauge symmetries of a non-Abelian Yang–Mills field with gauge group $U(N)$ (at lowest order in the fields). Thus we see that we can obtain non-Abelian gauge field dynamics from open strings.

1.3.5 Closed Strings

Let us now consider a closed string, so that $\sigma \sim \sigma + 2\pi$. The resulting "boundary condition" is more simple: we simply demand that $\hat{X}^\mu(\tau, \sigma + 2\pi) = \hat{X}^\mu(\tau, \sigma)$. This is achieved by again taking n to be an integer. However we now have two independent sets of left and right moving oscillators. Thus the mode expansion is given by

$$X^\mu = x^\mu + \alpha' p^\mu \tau + \sqrt{\frac{\alpha'}{2}} i \sum_{n\neq 0} \left(\frac{a_n^\mu}{n} e^{-in(\tau+\sigma)} + \frac{\tilde{a}_n^\mu}{n} e^{-in(\tau-\sigma)} \right) \tag{1.3.103}$$

note the absence of the factor of 2 in front of $p^\mu \tau$. The total momentum of such a string is calculated as before to give

$$\begin{aligned} p^\mu &= \frac{1}{2\pi\alpha'} \int_0^{2\pi} d\sigma \dot{X}^\mu \\ &= \frac{1}{2\pi\alpha'} \int_0^{2\pi} d\sigma p^\mu + \sqrt{\frac{\alpha'}{2}} \sum_{n\neq 0} a_n^\mu e^{-in(\tau+\sigma)} + \tilde{a}_n^\mu e^{-in(\tau-\sigma)} \\ &= p^\mu \end{aligned} \tag{1.3.104}$$

so again p^μ is the spacetime momentum of the string.

We now have double the constraints:

$$\begin{aligned} (L_0 - 1)|phys\rangle &= (\tilde{L}_0 - 1)|phys\rangle = 0 \\ L_n|phys &= \tilde{L}_n|phys\rangle = 0 \end{aligned} \tag{1.3.105}$$

with $n > 0$. If we introduce the right-moving number operator $\tilde{N}$

$$\tilde{N} = \sum_{n>0} \eta_{\mu\nu} \tilde{a}_{-n}^\mu \tilde{a}_n^\nu \tag{1.3.106}$$

then the first conditions can be rewritten as

$$(p_\mu p^\mu + \frac{4}{\alpha'}(N-1))|phys\rangle = 0 \qquad (N-\tilde{N})|phys\rangle = 0 \tag{1.3.107}$$

where we have recalled that, if $w^\mu = 0$, $a_0^\mu = \tilde{a}_0^\mu = \sqrt{\frac{\alpha'}{2}}p^\mu$ and $L_0 = \frac{1}{2}\eta_{\mu\nu}a_0^\mu a_0^\nu + N$, $L_0 = \frac{1}{2}\eta_{\mu\nu}\tilde{a}_0^\mu \tilde{a}_0^\nu + \tilde{N}$. The second condition is called level matching. It simply says that any physical state must be made up out of an equal number of left and right moving oscillators. Again the remaining constraints will give gauge fixing conditions.

Let us consider the lowest modes of the closed string. At level 0 (which means level 0 on both the left and right moving sectors by level matching) we simply have the ground state $|0; p\rangle$. This is automatically annihilated by both L_n and $\tilde{L}_n$ with $n > 0$. For $n = 0$ we find

$$p^2 - \frac{4}{\alpha} = 0 \tag{1.3.108}$$

Thus we again find a tachyonic ground state. No one knows what to do with this instability. It turns out to be much more serious than the open string tachyon that we saw, which can sometimes be dealt with. Most people today would say that the bosonic string is inconsistent although this hasn't been demonstrated. However for us the cure is the same as for the open string: in the superstring this mode is projected out. So we continuing by simply ignoring it, as our discussion of the other modes still holds in the superstring.

ext we have level 1. Here the states are of the form

$$|G_{\mu\nu}\rangle = G_{\mu\nu}a_{-1}^\mu \tilde{a}_{-1}^\nu|0; p\rangle \tag{1.3.109}$$

Just as for the open string these will be massless, i.e. $p^2 = 0$ (again only if $a = 1$). Next we consider the constraints $L_m|G_{\mu\nu}\rangle = \tilde{L}_m|G_{\mu\nu}\rangle = 0$ with $m > 0$.

Exercise 9 *Show that these constraints imply that* $p^\mu G_{\mu\nu} = p^\nu G_{\mu\nu} = 0$

The matrix $G_{\mu\nu}$ is a spacetime tensor. Under the Lorentz group $SO(1, D-1)$ it will decompose into a symmetric traceless, anti-symmetric and trace part. What this means is that under spacetime Lorentz transformations the tensors $g_{\mu\nu}$, $b_{\mu\nu}$ and ϕ will transform into themselves. Here

$$\begin{aligned} g_{\mu\nu} &= G_{(\mu\nu)} - \frac{1}{D}\eta^{\lambda\rho}G_{\lambda\rho}\eta_{\mu\nu} \\ b_{\mu\nu} &= G_{[\mu\nu]} \\ \phi &= \eta^{\lambda\rho}G_{\lambda\rho} \end{aligned} \tag{1.3.110}$$

i.e. $G_{\mu\nu} = g_{\mu\nu} + b_{\mu\nu} + \frac{1}{D}\eta_{\mu\nu}\phi$.

Exercise 10 *Show this.*

Thus from the spacetime point of view there are three independent modes labelled by $g_{\mu\nu}$, $b_{\mu\nu}$ and ϕ. Just as for the open string there is a gauge symmetry

$$|G_{\mu\nu}\rangle \to |G_{\mu\nu}\rangle + i\xi_\mu L_{-1}\tilde{a}^\mu_{-1}|0; p\rangle + i\zeta_\mu \tilde{L}_{-1}a^\mu_{-1}|0; p\rangle \tag{1.3.111}$$

where we have used the fact that $\xi_\mu L_{-1}\tilde{a}^\mu_{-1}|0; p\rangle$ and $\zeta_\mu \tilde{L}_{-1}a^\mu_{-1}|0; p\rangle$ are null states, provided that $p^2 = 0$. The proof of this is essentially the same as it was for the open string. We need only ensure that the level matching condition is satisfied, which is clear, and that $\tilde{L}_n L_{-1}\tilde{a}_{-1}|0; p\rangle = L_n\tilde{L}_{-1}a_{-1}|0; p\rangle = 0$ for $n > 0$. Thus we need only check that

$$L_n\tilde{L}_{-1}a^\mu_{-1}|0; p\rangle = \frac{1}{2}\tilde{L}_{-1}\sum_m \eta_{\lambda\rho}a^\lambda_{n+m}a^\rho_{-m}a^\mu_{-1}|0; p\rangle = 0 \tag{1.3.112}$$

Just as before the $n > 1$ terms will vanish automatically. So we need only check

$$\begin{aligned} L_1\tilde{L}_{-1}a^\mu_{-1}|0; p\rangle &= \frac{1}{2}\tilde{L}_{-1}\sum_m \eta_{\lambda\rho}a^\lambda_{1+m}a^\rho_{-m}a^\mu_{-1}|0; p\rangle \\ &= \tilde{L}_{-1}\eta_{\lambda\rho}a^\lambda_0 a^\rho_1 a^\mu_{-1}|0; p\rangle \\ &= \tilde{L}_{-1}\eta_{\lambda\rho}a^\lambda_0[a^\rho_1, a^\mu_{-1}]|0; p\rangle \\ &= \tilde{L}_{-1}a^\mu_0|0; p\rangle \\ &= \frac{\sqrt{\alpha'}}{2}\tilde{L}_{-1}p^\mu|0; p\rangle \end{aligned} \tag{1.3.113}$$

Similarly for $\tilde{L}_n L_{-1}\tilde{a}^\mu_{-1}|0; p\rangle$. Thus we also find that $p^\mu\xi_\mu = p^\mu\zeta_\mu = 0$. This of course is required to preserve the condition $p^\mu G_{\mu\nu} = p^\nu G_{\mu\nu} = 0$.

In terms of $G_{\mu\nu}$ this implies that

$$G_{\mu\nu} \to G_{\mu\nu} + i\sqrt{\frac{\alpha'}{2}}p_\mu\xi_\nu + i\sqrt{\frac{\alpha'}{2}}p_\nu\zeta_\mu \tag{1.3.114}$$

or, switching to coordinate space representations and the individual tensor modes, we find

$$\begin{aligned} g_{\mu\nu} &\to g_{\mu\nu} + \frac{1}{2}\sqrt{\frac{\alpha'}{2}}\partial_\mu(\xi_\nu + \zeta_\nu) + \frac{1}{2}\sqrt{\frac{\alpha'}{2}}\partial_\nu(\xi_\mu + \zeta_\mu) \\ b_{\mu\nu} &\to B_{\mu\nu} + \frac{1}{2}\sqrt{\frac{\alpha'}{2}}\partial_\mu(\xi_\nu - \zeta_\nu) - \frac{1}{2}\sqrt{\frac{\alpha'}{2}}\partial_\nu(\xi_\mu - \zeta_\mu) \\ \phi &\to \phi + 2\sqrt{\frac{\alpha'}{2}}\partial_\mu(\xi^\mu + \zeta^\mu) \end{aligned} \tag{1.3.115}$$

If we let $v_\mu = \frac{1}{2}\sqrt{\frac{\alpha'}{2}}(\xi_\mu + \zeta_\mu)$ and $\Lambda_\mu = \frac{1}{2}\sqrt{\frac{\alpha'}{2}}(\xi_\mu - \zeta_\mu)$ and use $\partial^\mu\xi_\mu = p^\mu\zeta_\mu = 0$ then we find

$$
\begin{aligned}
g_{\mu\nu} &\to g_{\mu\nu} + \partial_\mu v_\nu + \partial_\nu v_\mu \\
b_{\mu\nu} &\to b_{\mu\nu} + \partial_\mu \Lambda_\nu - \partial_\nu \Lambda_\mu \\
\phi &\to \phi
\end{aligned} \tag{1.3.116}
$$

The first term line gives the infinitesimal form of a diffeomorphism, $x^\mu \to x^\mu - v^\mu$ and thus we can identify $g_{\mu\nu}$ to be a metric tensor. The second line gives a generalization of and electromagnetic gauge transformation. The analogue of the gauge invariant field strength is

$$
H_{\lambda\mu\nu} = \partial_\lambda b_{\mu\nu} + \partial_\mu b_{\nu\lambda} + \partial_\nu b_{\lambda\mu} \tag{1.3.117}
$$

Thus the massless field content at level 1 consists of a graviton mode $g_{\mu\nu}$, an anti-symmetric tensor field $b_{\mu\nu}$ and a scalar ϕ, subject to the gauge transformations (1.3.115). Finally the massless condition $p^2 G_{\mu\nu} = 0$ leads to

$$
\begin{aligned}
\partial^2 g_{\mu\nu} &= 0 \\
\partial^2 b_{\mu\nu} &= 0 \\
\partial^2 \phi &= 0
\end{aligned} \tag{1.3.118}
$$

The conditions $p^\mu G_{\mu\nu} = p^\nu G_{\mu\nu} = 0$ now reduce to the linearized equations

$$
\begin{aligned}
\partial^\mu g_{\mu\nu} + \partial_\nu \phi &= 0 \\
\partial^\mu b_{\mu\nu} &= 0
\end{aligned} \tag{1.3.119}
$$

These equations can be viewed as gauge fixing conditions (in effect $\phi = \frac{1}{D} g_{\lambda\sigma} \eta^{\lambda\sigma}$). The fields $g_{\mu\nu},\ b_{\mu\nu}$ and ϕ are known as the graviton (metric), Kalb–Ramond (b-field) and dilaton respectively.

1.4 Light-Cone Gauge

So far we have quantized a string in flat D-dimensional spacetime. Apart from D we have the parameters a and α'. In fact α' is not a parameter, it is a dimensional quantity—it has the dimensions of length-squared—and simply sets the scale. What is important are unitless quantities such as $p^2\alpha'$. For example small momentum means $p^2\alpha' \ll 1$.

We are left with D and a but actually these are fixed: quantum consistency demands that $D = 26$ and $a = 1$. We have seen that things would go horribly wrong if $a \neq 1$.

The easiest way to see this is to introduce light-cone gauge. Recall that the action we started with had diffeomorphism symmetry. We used this symmetry to fix $\gamma_{\alpha\beta} = e^{2\rho}\eta_{\alpha\beta}$. However there is still a residual symmetry. In particular in terms of the coordinates $\sigma^\pm$ then under a transformation

$$
\sigma'^{+} = \sigma'^{+}(\sigma^+) \qquad \sigma'^{-} = \sigma'^{-}(\sigma^-) \tag{1.4.120}
$$

we see that $\gamma'_{\alpha\beta} = e^{2\rho'}\eta_{\alpha\beta}$ with

$$\rho' = \rho + \frac{1}{2}\ln\left(\frac{\partial\sigma^+}{\partial\sigma'^+}\frac{\partial\sigma^-}{\partial\sigma'^-}\right) \tag{1.4.121}$$

i.e. this preserves the conformal gauge. In terms of the worldsheet coordinates σ, τ we see that

$$\tau' = \frac{1}{2}(\sigma'^+ + \sigma'^-) \tag{1.4.122}$$

and since $\sigma'^{\pm}$ are arbitrary functions of $\sigma^{\pm}$ we see that any τ that solves the two-dimensional wave equation can be obtained by such a diffeomorphism. Therefore, without loss of generality, we can choose the worldsheet 'time' coordinate τ to be any of the spacetime coordinates (since these solve the two-dimensional wave-equation). Of course there are many choices but the usual one is to define

$$\hat{X}^+ = \frac{1}{2}(X^0 + X^{D-1}) \qquad \hat{X}^- = \frac{1}{2}(X^0 - X^{D-1}) \tag{1.4.123}$$

and then take

$$\hat{X}^+ = x^+ + \alpha' p^+\tau \tag{1.4.124}$$

This is called light cone gauge.

Next we evaluate the conformal symmetry constraints (1.3.15). We observe that in these coordinates the spacetime $\eta_{\mu\nu}$ is

$$\eta_{-+} = \eta_{+-} = -2 \qquad \eta_{ij} = \delta_{ij} \tag{1.4.125}$$

Thus we find that

$$\begin{aligned} T_{00} = T_{11} &= -2\alpha' p^+\dot{\hat{X}}^- + \frac{1}{2}\dot{X}^i\dot{X}^j\delta_{ij} + \frac{1}{2}X'^iX'^j\delta_{ij} = 0 \\ T_{01} = T_{10} &= -2\alpha' p^+\hat{X}'^- + \dot{X}^iX'^j\delta_{ij} = 0 \end{aligned} \tag{1.4.126}$$

where $i, j = 1, 2, 3, \ldots, D-2$. This allows one to explicitly solve for X^- in term of the mode expansions for X^i.

Problem Show that with our conventions

$$X^- = x^- + \alpha' p^-\tau + i\left(\sum_{n\neq 0}\frac{a_n^-}{n}e^{-in\sigma^+} + \frac{\tilde{a}_n^-}{n}e^{-in\sigma^-}\right) \tag{1.4.127}$$

where

$$a_n^- = \frac{1}{2p^+}\sum_m a^i_{n-m}a^j_m\delta_{ij} \tag{1.4.128}$$

and the massshell constraint is

$$-4\alpha' p^+ p^- + \alpha' p^i p^j \delta_{ij} + 2(N + \tilde{N}) = 0 \qquad (1.4.129)$$

with

$$N + \tilde{N} = \frac{1}{2}\delta_{ij} \sum_{n \neq 0} a^i_n a^j_{-n} + \tilde{a}^i_n \tilde{a}^j_{-n} \qquad (1.4.130)$$

To continue we note that in the quantum theory there is a normal ordering ambiguity in the definition of $N + \tilde{N}$ and we must include our constant a again into the definition. Hence we must take (temporarily putting in the :: symbols for normal ordering)

$$: N + \tilde{N} := \delta_{ij} \sum_{n=1}^{\infty} \alpha^i_{-n} \alpha^j_n + \tilde{\alpha}^i_{-n} \tilde{\alpha}^j_n \qquad (1.4.131)$$

However since we have dropped an infinite constant, the intercept a will now show up in the mass shell constraint as

$$-4\alpha' p^+ p^- + \alpha' p^i p^j \delta_{ij} + 2(N + \tilde{N} - 2a) = 0 \qquad (1.4.132)$$

Note that $-4p^+ p^- + p^i p^j \delta_{ij} = \eta_{\mu\nu} p^\mu p^\nu$ so this really just tells us that the mass of a state is

$$M^2 = \frac{2}{\alpha'}(N + \tilde{N} - 2a) \qquad (1.4.133)$$

Where we have dropped the :: to indicate normal ordering.

We still have a level matching condition for closed strings

$$N = \tilde{N} \qquad (1.4.134)$$

This arises because we only have one spacetime momentum p^μ (not separate ones for left and right moving modes).

Note that this breaks the $SO(1, D-1)$ symmetry of our flat target space since we choose X^0 and X^{D-1} whereas any pair will do (so long as one is timelike). Thus we will not see a manifest $SO(1, D-1)$ symmetry but just an $SO(D-2)$ symmetry from rotations of the $\tilde{X}^i$. However it is important to realize that the $SO(1, D-1)$ symmetry is not really broken, we have merely performed a kind of gauge fixing (recall there was this underlying gauge symmetry of the string spectrum). It is just no longer manifest.

1.4.1 $D = 26$, $a = 1$

On the other hand the benefit of this procedure is that the physical Hilbert space is manifestly postive definite because we remove the oscillators $a_n^0, \tilde{a}_n^0, a_n^{D-1}, \tilde{a}_n^{D-1}$. This is often a helpful way to determine the physical spectrum of the theory.

For example we can reconsider the low lying states that we constructed above. The ground states are unchanged as they do not involve any oscillators. For the open string we find the $D-2$ states at level one

$$|A_i\rangle = a_{-1}^i|0; p\rangle \tag{1.4.135}$$

These are the transverse components of a massless gauge field. For the closed string we find, at level one,

$$|G_{ij}\rangle = G_{ij}a_{-1}^i\tilde{a}_{-1}^j|0; p\rangle \tag{1.4.136}$$

These correspond to the physical components, in a certain gauge, of the metric, Kalb-Ramond field and dilaton Note however that there is no remnant at all of gauge symmetry which is a crucial feature of dynamics

Now formally a is given by

$$\begin{aligned} a &= -\frac{1}{2}\sum_{m=1}^{\infty}[a_m^i, a_{-m}^j]\delta_{ij} \\ &= -\frac{D-2}{2}\sum_{m=1}^{\infty} m \end{aligned} \tag{1.4.137}$$

This is divergent however it can be regularized in the following manner. We note that

$$a = -\frac{D-2}{2}\zeta(-1) \tag{1.4.138}$$

where $\zeta(s)$ is the Riemann ζ-function

$$\zeta(s) = \sum_{m=1}^{\infty}\frac{1}{m^s} \tag{1.4.139}$$

This is analytic for complex s with $\mathrm{Re}(s) > 1$. Thus it can be extended to a holomorphic function of the complex plane, with poles at a discrete number of points. Analytically continuing to $s = -1$ one finds $\zeta(-1) = -1/12$ and hence

$$a = \frac{D-2}{24} \tag{1.4.140}$$

We have seen that in order to have a sensible theory we must take $a = 1$ (otherwise there are no massless states or nice gauge invariances). Hence we must take $D = 26$.

This is not a very satisfactory derivation of the dimension of spacetime. A more convincing argument is the following. Light cone gauge is just a gauge. Therefore although the manifest spacetime Lorentz symmetry is no longer present there is still an $SO(1, D-1)$ Lorentz symmetry, even though only an $SO(D-2)$ subgroup is manifest in light cone gauge. In light cone gauge the spacetime Lorentz generators $M^{\mu}{}_{\nu}$ split into

$$M^{i'}{}_{j} \qquad M^{+}{}_{j}, \qquad M^{-}{}_{j}, \qquad M^{+}{}_{-} \tag{1.4.141}$$

The quantization procedure preserves $SO(D-2)$ so the commutators $[M^{i}{}_{j}, M^{k}{}_{l}]$ are as they should be. However problems can arise with $[M^{i}{}_{j}, M^{+}{}_{k}]$ *etc*. It is too lengthy a calculation to do here, but one can show that the full $SO(1, D-1)$ Lorentz symmetry, generated by the charges (1.3.26), is preserved in the quantum theory, i.e. once normal ordering is taken into account, if and only if $a = 1$ and $D = 26$. You are urged to read the Sect. 2.3 of [1] and Witten or Sect. 12.5 of [3] where this is shown more detail.

1.4.2 Partition Function

A useful concept is the notion of a partition function which 'counts' the physical states. Since light cone gauge only contains physical states this is most easily computed here.

Let us start with an open string and define

$$Z = \sum q^{L_0 - 1} \tag{1.4.142}$$

where the sum is over states (at zero momentum) and $q = e^{-2\pi t}$ is 'place-holder'. First note that for a string in flat spacetime L_0 is a sum of 24 independent free bosons. Thus

$$Z = (Z_1)^{24} \tag{1.4.143}$$

where

$$\begin{aligned} Z_1 &= \sum q^{\sum_l a_{-l} a_l - \frac{1}{24}} \\ &= \sum q^{-\frac{1}{24}} \prod_l q^{a_{-l} a_l} \end{aligned} \tag{1.4.144}$$

For a single boson we have the oscillators $a_{-1}, a_{-2}, \ldots$. Each oscillator a_{-l} can be used k times in which case $a_{-l} a_l$ contributes k to the exponent. We need to sum over all k and using $\sum_{k=0}^{\infty} q^{kl} = (1 - q^l)^{-1}$ we find

$$Z_1 = q^{-\frac{1}{24}} \prod_{l=1}^{\infty} (1 - q^l)^{-1} \tag{1.4.145}$$

and hence

$$Z = q^{-1} \prod_{l=1}^{\infty} (1 - q^l)^{-24} = \eta(t) \tag{1.4.146}$$

where $\eta(t)$ is known as the Dedekind eta-function. It can be extended to the upper half complex plane $\tau = \theta + it,\ t > 0$ and is known to posses the following property:

$$\eta(-\tau^{-1}) = \eta(\tau) \tag{1.4.147}$$

In particular it is invariant under $t \to 1/t$ when $\theta = 0$. This property is known as modular invariance. This is a crucial feature of strings (and requires that we have 24 physical oscillators—another important feature of $D=26$).

We can provide a physical interpretation of Z by noting that

$$\int_0^{\infty} dt e^{-2\pi t(L_0 - 1)} = \frac{1}{2\pi} \frac{1}{L_0 - 1} \tag{1.4.148}$$

and $1/(L_0 - 1)$ is the propagator. Thus Z has the interpretation of a vacuum one-loop diagram:

$$Z = \mathrm{Tr}\langle 0| \left(\frac{1}{2\pi} \frac{1}{L_0 - 1} \right) |0\rangle \tag{1.4.149}$$

(hence the restriction to zero momentum). The variable t arises in the Schwinger proper time formalism. The worldsheet of an open string is a cylinder of radius R and length L. By conformal invariance the only parameter that matters is $t = R/L$. In the large t limit the open string is relatively short compared to the size of the loop. In this case the important states that propagate around the loop are the light modes, corresponding to the IR limit of open strings. Indeed here we see, taking $q \to 0$

$$Z \sim q^{-1} + 24q^0 + \mathcal{O}(q) \tag{1.4.150}$$

here we see the tachyon dominants, followed by the 24 massless modes and then the massive spectrum gives ever vanishing corrections.

What about the $t \to 0$ limit? In this case the the cylinder has a very short radius compared to its length. This corresponds to the UV behaviour of the open string and the massive states dominate. Here we can use modular invariance to evaluate

$$\lim_{t \to 0} \eta(e^{-2\pi t}) = \lim_{t \to 0} \eta(e^{-2\pi/t}) = \lim_{\tilde{t} \to \infty} \eta(e^{-2\pi \tilde{t}}) \sim \tilde{q}^{-1} + 24\tilde{q}^0 + \mathcal{O}(\tilde{q}) \tag{1.4.151}$$

where $\tilde{q} = e^{-2\pi\tilde{t}}$. An alternative interpretation of such a diagram is that it can be viewed as a closed string of radius R propagating at tree-level along a distance L, i.e. no loops. Again Z dominated by the closed string tachyon and then the 24 (left-right symmetric) massless closed string modes.

This is one of the most important features of string theory. The UV description of open strings has a dual interpretation in terms of an IR propagation of closed strings and *vice-versa*.

Exercise 11 *Show that for a periodic fermion, where* $L_0 = \sum_l d_{-l}d_l + \frac{1}{24}$ *and* $\{d_n, d_m\} = n\delta_{n,-m}$, *one has*

$$Z_1 = q^{\frac{1}{24}} \prod_{l=1}^{\infty}(1 + q^l) \tag{1.4.152}$$

and for an anti-periodic fermion, where $L_0 = \sum_r b_{-r}b_r - \frac{1}{48}$, $\{b_r, b_s\} = r\delta_{r,-s}$ *and* $r, s \in \mathbf{Z} + \frac{1}{2}$, *one has*

$$Z_1 = q^{-\frac{1}{48}} \prod_{l=1}^{\infty}(1 + q^{l-\frac{1}{2}}) \tag{1.4.153}$$

1.5 Curved Spacetime and an Effective Action

1.5.1 Strings in Curved Spacetime

We have considered quantized strings propagating in flat spacetime. This lead to a spectrum of states that included the graviton as well as other modes. More generally a string should be allowed to propagate in a curved background with non-trivial values for the metric and other fields. Our ansatz will be to consider the most general two-dimensional action for the embedding coordinates X^μ coupled to two-dimensional gravity subject to the constraint of conformal invariance. This later condition is required so that the two-dimensional worldvolume metric decouples from the other fields. We will consider only closed strings in this section. The reason for this is that these days one views open strings as description soliton like objects, called Dp-branes, that naturally sit inside the closed string theory.

Before proceeding we note that

$$S_{EH} = \frac{1}{4\pi}\int d^2\sigma\sqrt{-\gamma}R = \chi \tag{1.5.1}$$

is a topological invariant called the Euler number, i.e. the integrand is locally a total derivative. Thus we could add the term S_{EH} to the action and not change the equations of motion.

With this in mind the most general action we can write down for a closed string is

$$S_{closed} = -\frac{1}{4\pi\alpha'}\int d^2\sigma \alpha' \sqrt{-\gamma}\phi(X)R + \sqrt{-\gamma}\gamma^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu g_{\mu\nu}(X)$$
$$+ \varepsilon^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu b_{\mu\nu}(X) \quad (1.5.2)$$

where ϕ is a scalar, $g_{\mu\nu}$ symmetric and $b_{\mu\nu}$ antisymmetric. These are precisely the correct degrees of freedom to be identified with the massless modes of the string. One can think of this worldsheet theory as two-dimensional quantum gravity coupled to some matter in the form of scalar fields. More generally one can think of and conformal field theory (with central charge equal to 26) as defining the action for a string.

Furthermore this action has the diffeomorphism symmetry $X^\mu \to X'^\mu(X)$

$$\partial_\alpha X'^\mu = \frac{\partial X'^\mu}{\partial X^\nu}\partial_\alpha X^\nu \qquad \phi' = \phi$$
$$g'_{\mu\nu} = \frac{\partial X^\lambda}{\partial X'^\mu}\frac{\partial X^\rho}{\partial X'^\nu} g_{\lambda\rho} \qquad b'_{\mu\nu} = \frac{\partial X^\lambda}{\partial X'^\mu}\frac{\partial X^\rho}{\partial X'^\nu} b_{\lambda\rho} \quad (1.5.3)$$

automatically built in. It also incorporates the b-field gauge symmetry

$$b'_{\mu\nu} = b_{\mu\nu} + \partial_\mu \lambda_\nu - \partial_\mu \lambda_\nu \quad (1.5.4)$$

however to see this we note that

$$\begin{aligned}\delta S_{closed} &= -\frac{1}{2\pi\alpha'}\int d^2\sigma \varepsilon^{\alpha\beta}\partial_\alpha X^\mu \partial_\beta X^\nu \partial_\mu \lambda_\nu \\ &= -\frac{1}{2\pi\alpha'}\int d^2\sigma \partial_\alpha(\varepsilon^{\alpha\beta}\partial_\beta X^\nu \lambda_\nu) \\ &= 0\end{aligned} \quad (1.5.5)$$

where we used the fact that $\varepsilon^{\alpha\beta}\partial_\alpha\partial_\beta X^\nu = 0$ in the second to last line and the fact that the worldsheet is a closed manifold in the last line, i.e. the periodic boundary conditions.

Notice something important. If the dilaton ϕ is constant then the first term in the action is a topological invariant, the Euler number. In the path integral formulation the partition function for the full theory is defined by summing over all worldsheet topologies

$$Z = \sum_{g=0}^{\infty}\int D\gamma DX e^{-S} \quad (1.5.6)$$

Here the path integral is over the worldsheet fields $\gamma_{\alpha\beta}$ and X^μ. Now each genus g worldsheet will appear suppressed by the factor $e^{-\phi\chi_g} = e^{-2\phi(g-1)}$. Thus $g_s = e^\phi$ can be thought of as the string coupling constant which counts which genus surface is contributing to a calculation. In particular for $g_s \to 0$ one can just consider the leading order term where the worldsheet is a sphere.

However if one wants to consider the splitting and joining of strings then one must take $g_s > 0$ and include higher genus surfaces. In particular the first non-trivial string interactions arise when the worldsheet is a torus. To see the analogy with quantum field theory note that a torus can be thought of as the worldvolume of a closed string that has gone around in a loop. Thus it is analogous to 1-loop processes in quantum field theory. Similarly higher genus surfaces incorporate higher loop processes. One of the great features of string theory is that each of these contributions is finite. So this defines a finite perturbative expansion of a quantum theory which includes gravity!

As stated above our general principle is the conformal invariance of the worldsheet theory, which ensures that the worldsheet metric $\gamma_{\alpha\beta}$ decouples. The action we just wrote down is conformal as a classical action. However this will not generically be the case in the quantum theory. Divergences in the quantum theory require regularization and renormalization and these effects will break conformal invariance by introducing an explicit scale: the renormalization group scale. It turns out that conformal invariance is more or less equivalent to finiteness of the quantum field theory. This restriction leads to equations of motions for the spacetime fields ϕ, $g_{\mu\nu}$ and $b_{\mu\nu}$ (which from the worldvolume point of view are just fancy coupling constants). It is beyond the scope of these lectures to show this but the constraints of conformal invariance at the one loop level give equations of motion

$$
\begin{aligned}
0 &= R_{\mu\nu} + \frac{1}{4} H_{\mu\lambda\rho} H_\nu^{\lambda\rho} - 2D_\mu D_\nu \phi + \mathscr{O}(\alpha') \\
0 &= D^\lambda H_{\lambda\mu\nu} - 2D^\lambda \phi H_{\lambda\mu\nu} + \mathscr{O}(\alpha') \\
0 &= 4D^2\phi + 4(D\phi)^2 - R - \frac{1}{12} H^2 + \mathscr{O}(\alpha')
\end{aligned}
\tag{1.5.7}
$$

where $H_{\mu\nu\lambda} = 3\partial_{[\mu} b_{\nu\lambda]}$. In general there will be corrections to these equations coming from all orders in perturbation theory, i.e. higher powers of α'. However such terms will be higher order spacetime derivatives and can be safely ignored at energy scales below the string scale.

1.5.2 A Spacetime Effective Action

A string propagating in spacetime has an infinite tower of massive excitations. However all but the lightest (massless) modes will be too heavy to observe in any experiment that we do. Thus in many cases one really just wants to consider the dynamics of the massless modes. This introduces the concept of an effective action. This is a very general concept (ubiquitous in quantum field theory) whereby we introduce an action for the light modes that we are interested in (below some scale M). The action is constructed so that it has all the correct symmetries of the full theory and its equations of motion reproduce the correct scattering amplitudes of the light modes that the full theory predicts. In general effective actions need not be renormalizable and they are not expected to be valid at energy scales above the scale M where the massive modes we've ignored can be excited and can no longer be ignored. Often one says that the massive modes have been integrated out. Meaning that one has performed the path integral over modes with momenta larger than M and is just left with a path integral over the low momentum modes.

In our case we have considered a string propagating in a curved spacetime that can be thought of as a background coming from a non-trivial configuration of its massless modes. In particular in our discussion we implicitly assumed that the massive modes were set to zero. The result was that quantum conformal invariance predicted the equations of motion (1.5.7). These are the on-shell conditions for a string to propagate in spacetime as derived in the full quantum theory. Note that they pick up an infinite series of α' corrections and also an infinite series of g_s corrections (where we allow the splitting an joining of strings). In other words, at lowest order in α' and g_s these are the equations of motion for the spacetime fields. Furthermore these equations of motion can be derived from the spacetime action

$$S_{effective} = -\frac{1}{2\alpha'^{12}} \int d^{26}x \sqrt{-g} e^{-2\phi} \left(R - 4(\partial\phi)^2 + \frac{1}{12} H_{\mu\nu\lambda} H^{\mu\nu\lambda} \right) + \dots \tag{1.5.8}$$

Exercise 12 *Show that the equations of motion of* (1.5.8) *are indeed* (1.5.7). *You may need to recall that* $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu}$ *and* $g^{\mu\nu}\delta R_{\mu\nu} = D_\mu D_\nu \delta g^{\mu\nu} - g_{\mu\nu} D^2 \delta g^{\mu\nu}$.

This is therefore the effective action for the massless modes of a closed string. It plays the same role that the free scalar equation played for the point particle (although $S_{effective}$ does not include the infinite tower of string states which isn't there for the point particle). The ellipsis denotes contributions from higher loops which will contain higher numbers of derivatives and which are suppressed by higher powers of α'. Note that string theory also predicts corrections to the effective action from string loops, that is from higher genus Riemann surfaces. These terms will come with factors of $e^{-2g\phi}$ where $g = 0, -1, -2, \dots$ and can be ignored if the string coupling $g_s = e^{\phi}$ is small.

1.6 Superstrings

In the final section let us try to extend the pervious sections to the superstring. Conceptually not much changes but there are several additional bells and whistles that need to be considered.

1.6.1 Type II Strings

The starting point for the superstring is include fermions ψ^μ on the worldsheet so as to construct a supersymmetric action

$$S = -\frac{1}{4\pi\alpha'}\int d^2\sigma\,\partial_\alpha X^\mu \partial_\beta X^\nu \eta_{\mu\nu}\eta^{\alpha\beta} + i\bar{\psi}^\mu\gamma^\alpha\partial_\alpha\psi^\nu\eta_{\mu\nu} \tag{1.6.9}$$

where $\bar{\psi} = \psi^T\gamma_0$ and γ^α are real 2×2 matrices that satisfy $\{\gamma_\alpha, \gamma_\beta\} = 2\eta_{\alpha\beta}$. A convenient choice is $\gamma^0 = i\sigma^2$ and $\gamma^1 = \sigma^1$. This action is also conformally invariant and in addition has the supersymmetry

$$\delta X^\mu = i\bar{\varepsilon}\psi^\mu, \qquad \delta\psi^\mu = \gamma^\alpha\partial_\alpha X^\mu\varepsilon \tag{1.6.10}$$

for any constant ε.

Exercise 13 *Show this.*

The mode expansion for the X^μ remains as before with the a_n^μ and $\tilde{a}_n^\mu$ oscillators. When we expand the fermionic fields we can allow for two types of boundary conditions (let us just consider boundary conditions consistent with a closed string where $\sigma \sim \sigma + 2\pi$ and $X^\mu(\tau,\sigma) = X^\mu(\tau,\sigma+2\pi)$:

$$\begin{aligned} \mathrm{R}: &\quad \psi^\mu(\tau,\sigma+2\pi) = \psi^\mu(\tau,\sigma) \\ NS: &\quad \psi^\mu(\tau,\sigma+2\pi) = -\psi^\mu(\tau,\sigma) \end{aligned} \tag{1.6.11}$$

these are known as the Ramond and Neveu-Schwarz sectors respectively. Thus we find

$$\begin{aligned} \mathrm{R}: &\quad \psi^\mu(\tau,\sigma+2\pi) = \sum_{n\in\mathbf{Z}} d_n e^{-in\sigma^+} + \tilde{d}_n e^{-in\sigma^-} \\ \mathrm{NS}: &\quad \psi^\mu(\tau,\sigma+2\pi) = \sum_{r\in\mathbf{Z}+\frac{1}{2}} b_r e^{-ir\sigma^+} + \tilde{b}_r e^{-ir\sigma^-} \end{aligned} \tag{1.6.12}$$

One finds that these satisfy the anti-commutation relations

$$\begin{aligned} \{d_m^\mu, d_n^\nu\} = \eta^{\mu\nu}\delta_{m,-n} &\qquad \{b_r^\mu, b_s^\nu\} = \eta^{\mu\nu}\delta_{r,-s} \\ \{\tilde{d}_m^\mu, \tilde{d}_n^\nu\} = \eta^{\mu\nu}\delta_{m,-n} &\qquad \{\tilde{b}_r^\mu, \tilde{b}_s^\nu\} = \eta^{\mu\nu}\delta_{r,-s} \end{aligned} \tag{1.6.13}$$

with all other anti-commutators vanishing.

One important consequence of supersymmetry is that the algebra of constraints generated by L_n is enhanced to a super-Virasoro algebra with odd generators G_r and F_n (depending on whether or not one is in the NS or R sector respectively). The super-Virasoro algebra turns out to be (see references [1–4])

$$
\begin{aligned}
[L_m,\ L_n] &= (m-n)L_{m+n} + \frac{D}{8}m(m^2-1)\delta_{m,-n} \\
[L_m, G_r] &= \left(\frac{m}{2}-r\right)G_{m+r} \\
\{G_r, G_s\} &= 2L_{r+s} + \frac{D}{2}\left(r^2-\frac{1}{4}\right)\delta_{r,-s}
\end{aligned}
\tag{1.6.14}
$$

in the NS sector and

$$
\begin{aligned}
[L_m,\ L_n] &= (m-n)L_{m+n} + \frac{D}{8}m^3\delta_{m,-n} \\
[L_m, F_n] &= \left(\frac{m}{2}-n\right)F_{m+n} \\
\{F_n, F_m\} &= 2L_{m+n} + \frac{D}{2}m^2\delta_{m,-n}
\end{aligned}
\tag{1.6.15}
$$

in the R sector. Here all operators are normal ordered. Just as before this only affects L_0 and F_0 however there is no associated intercept a for F_0 since it is fermionic (and in addition this is not allowed by the $\{F_0, F_0\}$ anti-commutator). Note that the fermionic generators are in effect the 'square-root' of L_n, as we expect in a supersymmetric theory. We won't go into more details here but we must impose the physical constraints for the positive modded generators. Just as L_0 gives a spacetime Klein–Gordon equation, F_0 gives a spacetime Dirac equation.

Let us compute the intercept a. As before we go to light-cone gauge where we fix two of the coordinates X^μ and their superpartners ψ^μ. We then compute the vacuum energy of the remaining $D-2$ bosonic and fermionic oscillators. The result depends on the boundary conditions we use. Noting that the sign of the fermionic contribution is opposite to that of a boson one finds

$$
\begin{aligned}
a_R &= -\frac{D-2}{2}\sum_{n=1}^{\infty} n + \frac{D-2}{2}\sum_{n=1}^{\infty} n \\
&= -\frac{D-2}{2}\left(-\frac{1}{12}+\frac{1}{12}\right) \\
&= 0
\end{aligned}
\tag{1.6.16}
$$

The vanishing of a_R is a direct consequence of the fact that there is a Bose–Fermi degeneracy in the R-sector. In particular each periodic fermion contributes $-\frac{1}{24}$ to a. In the NS sector we find

$$
\begin{aligned}
a_{NS} &= -\frac{D-2}{2}\sum_{n=1}^{\infty} n + \frac{D-2}{2}\sum_{r=0}^{\infty}\left(r+\frac{1}{2}\right) \\
&= -\frac{D-2}{2}\sum_{n=1}^{\infty} n + \frac{D-2}{4}\sum_{n=odd}^{\infty} n
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{D-2}{2}\sum_{n=1}^{\infty} n + \frac{D-2}{4}\left(\sum_{n=1}^{\infty} n - \sum_{n=even}^{\infty} n\right) \\
&= -\frac{D-2}{2}\sum_{n=1}^{\infty} n + \frac{D-2}{4}\left(\sum_{n=1}^{\infty} n - \sum_{m=1}^{\infty} 2m\right) \\
&= -\frac{D-2}{2}\sum_{n=1}^{\infty} n - \frac{D-2}{4}\sum_{n=1}^{\infty} n \\
&= (D-2)\left(\frac{1}{24} + \frac{1}{48}\right) \\
&= \frac{D-2}{16}
\end{aligned}
\tag{1.6.17}
$$

Note that this shows that each anti-periodic fermion contributes $\frac{1}{48}$ to a. Having determined the incepts we can now go out of Light cone gauge and consider the covariant theory.

Let us now look at the lightest states. There is a different ground state for each sector which we denote by $|R; p\rangle$ and $|NS; p\rangle$ where p^μ labels the spacetime momentum. As before we assume that these states are annihilated by any oscillator with positive frequency.

We see that $|R; p\rangle$ is massless and hence all the higher level states created from it by the action of a creation operator will be massive with a mass of order the string scale. However the Ramond ground state $|R; p\rangle$ is degenerate. In particular we see that there are fermion zero-modes d_0^μ which satisfy $\{d_0^\mu, d_0^\nu\} = \eta^{\mu\nu}$, $\mu, \nu = 0, \ldots, D-1$ in light cone gauge. This is a Clifford algebra and it is known that there is a unique representation and it is $2^{[\frac{D}{2}]}$-dimensional. Thus the Ramond ground state is in fact a spinor with $2^{[\frac{D}{2}]}$ independent components.

Let us look at the Neveu–Schwarz ground state $|NS, p\rangle$. It is clear that since $a_{NS} > 0$ this state is a tachyon. We can then consider the higher level states (for simplicity we just consider open strings)

$$
\begin{aligned}
&a_{-1}^\mu|NS, p\rangle \qquad M^2 = 1 - \frac{D-2}{16} \\
&b_{-\frac{1}{2}}^\mu|NS, p\rangle \qquad M^2 = \frac{1}{2} - \frac{D-2}{16}
\end{aligned}
$$

Thus the next lightest state is $b_{-\frac{1}{2}}^\mu|NS, p\rangle$ and its mass-squared is $M^2 = -\frac{D-10}{16}$. Thus if $D<10$ then these states are also tachyonic. However as before the magic (that is gauge symmetries from null states) happens when these states are massless, i.e. $D=10$. In this case the states $a_{-1}^\mu|NS, p\rangle$ are massive. Thus we take $D=10$ and $a_{NS} = 1/2$. Indeed as before this is forced upon us if we want the $SO(1, D-1)$ Lorentz symmetry of spacetime to be preserved in the quantum theory.

Nevertheless we are still left with some bad features. For one the Neveu–Schwarz ground state is still a tachyon. There is also another puzzling feature: $|NS, p\rangle$ is a spacetime scalar and hence it must be a boson. We can then construct the spacetime vector $b^{\mu}_{-\frac{1}{2}}|NS, p\rangle$. From the spacetime point of view this state should be a boson since it transforms under Lorentz transformations as a vector. However it is created from $|NS, p\rangle$ by a fermionic operator and thus will obey Fermi-statistics. This is contradictory.

The solution to both these problems is to project out the odd states and in particular $|NS, p\rangle$. This is known as the GSO projection. More specifically we declare that $|NS, p\rangle$ is a fermionic state. Mathematically we introduce the operator $(-1)^F$ which acts as $(-1)^F|NS, p\rangle = -|NS, p\rangle$ and $\{\psi^{\mu}, (-1)^F\} = 0$, $[X^{\mu}, (-1)^F] = 0$. We then project out all fermionic states, i.e. states in the eigenspace $(-1)^F = -1$. Thus $|NS, p\rangle$ and $a^{\mu}_{-1}|NS, p\rangle$ are removed from the spectrum but the massless states $b^{\mu}_{-\frac{1}{2}}|NS, p\rangle$ remain.

Let us now consider the Ramond sector states. We already saw that the ground state here is massless but degenerate. Indeed it is a spinor of $SO(1,9)$, that is to say it can be represented by a vector in the 32-dimensional vector space that furnishes a representation of the Clifford algebra relation $\{d^{\mu}_0, d^{\nu}_0\} = \eta^{\mu\nu}$, $\mu, \nu = 0, \ldots, 9$. We need to discuss how $(-1)^F$ acts here. There is a natural candidate where we take $(-1)^F = \pm\Gamma_{11} = \pm\Gamma_0\Gamma_1 \ldots \Gamma_9$, the chirality operator in the 10-dimensional Clifford algebra. Thus after the GSO projection $|R, p\rangle$ is a chiral spinor with 16 independent components. More generally in the Ramond sector we project out states with $(-1)^F = -1$. The GSO projection is also required to ensure modular invariance.

In the Ramond sector of the open superstring either choice of sign is equivalent to the other, it is just a convention. Thus for the open superstring the lightest states are massless and consist of a spacetime vector (and hence a boson) $b^{\mu}_{-\frac{1}{2}}|NS, p\rangle$ along with a spacetime fermion $|R, p\rangle$ which can be identified with a chiral spinor. Note that there is a Bose-Fermi degeneracy: on-shell, and gauged fixed we find 8 bosonic and 8 fermionic states (Why?—you can see this in lightcone gauge).

Let us consider closed strings. Here the states are essentially obtained by taking a tensor product of left and right moving modes and hence there are four possibilities:

$$
\begin{aligned}
&|NS\rangle_L \otimes |NS\rangle_R \\
&|R\rangle_L \otimes |R\rangle_R \\
&|NS\rangle_L \otimes |R\rangle_R \\
&|R\rangle_L \otimes |NS\rangle_L
\end{aligned}
\tag{1.6.18}
$$

In this case the relative sign taken in the GSO projection is important. There are two choices: either we chose the same chirality projector for the left and right moving modes or the opposite. This leads to two distinction theories known as the type IIB and type IIA superstring respectively. The states one find are of the form

$$
\begin{aligned}
&|NS\rangle_L \otimes |NS\rangle_R \\
&|R+\rangle_L \otimes |R-\rangle_R \\
&|NS\rangle_L \otimes |R-\rangle_R \\
&|R+\rangle_L \otimes |NS\rangle_L
\end{aligned} \tag{1.6.19}
$$

for type IIA and

$$
\begin{aligned}
&|NS\rangle_L \otimes |NS\rangle_R \\
&|R+\rangle_L \otimes |R+\rangle_R \\
&|NS\rangle_L \otimes |R+\rangle_R \\
&|R+\rangle_L \otimes |NS\rangle_L
\end{aligned} \tag{1.6.20}
$$

for type IIB. Here the $\pm$ sign corresponds to the different choice of GSO projector for the left and right moving modes.

The spacetime bosons come from either the NS-NS or R-R sectors whereas the spacetime fermions from the NS-R or R-NS sectors. One sees that in the type IIA theory there are fermionic states with both spacetime chiralities but in the type IIB theory only one chirality appears.

Let us look more closely at the massless bosonic states. The NS-NS sector is essentially the same as the spectrum of the bosonic string only now they are created from the vacuum by $b^\mu_{-\frac{1}{2}}$ and $\tilde{b}^\mu_{-\frac{1}{2}}$ rather than a^μ_{-1} and $\tilde{a}^\mu_{-1}$. In particular we still find a graviton, Kalb-Ramond field and a dilaton. This sector is universal to all closed string theories.

However we also have R-R fields. These arise as a tensor product of a left and right spinor ground state. As such they form a ‘bi-spinor’:

$$
F_{\alpha\beta} = |R\pm\rangle_{L\alpha} \otimes |R\pm\rangle_{R\beta} \tag{1.6.21}
$$

Any bi-spinor can be expanded in terms of the associated Γ-matrices:

$$
F_{\alpha\beta} = \sum_{p=0}^{10} F_{\mu_1\ldots\mu_p} (\Gamma^{\mu_1\ldots\mu_p}\Gamma^0)_{\alpha\beta} \tag{1.6.22}
$$

Here we have used the fact that $\{1, \Gamma^\mu, \Gamma^{\mu_1\mu_2}, \ldots, \Gamma^{\mu_1\ldots\mu_{10}}\}$ form a basis of 32×32 matrices and used $C^{-1} = \Gamma^0$ to lower the spinor index. Next we note that

$$
\Gamma^{11}\Gamma^{\mu_1\ldots\mu_p} = \frac{1}{(10-p)!}\varepsilon^{\mu_1\ldots\mu_p\nu_1\ldots\nu_{10-p}}\Gamma_{\nu_1\ldots\nu_{10-p}} \tag{1.6.23}
$$

Using the GSO projection on the left movers implies that $(\Gamma^{11})_\gamma{}^\alpha F_{\alpha\beta} = F_{\gamma\beta}$ and hence we see that

$$F^{\mu_1\ldots\mu_p} = \frac{1}{(10-p)!}\varepsilon^{\mu_1\ldots\mu_p\nu_1\ldots\nu_{10-p}}F_{\nu_1\ldots\nu_{10-p}} \tag{1.6.24}$$

This implies that only the fields with $p \leq 5$ are independent of each other. In addition $F_{\mu_1\ldots\mu_5}$ is self-dual. Finally the GSO projection on the right movers tells us that $F_{\alpha\gamma}(\Gamma^{11})^{\gamma}{}_{\beta} = \pm F_{\alpha\beta}$ where the sign is—for type IIA and + for type IIB. This implies that $p = even$ for type IIA and $p = odd$ for type IIB. The physical state conditions, in particular the vanishing of F_0 and $\tilde{F}_0$, imply that $\partial_{[\mu_{p+1}}F_{\mu_1\ldots\mu_p]} = 0$ and $\partial^{\mu_1}F_{\mu_1\ldots\mu_p} = 0$.

We motivated superstrings by considering a worldsheet action that was supersymmetric. However it turns out that, after the GSO projection, these theories also have spacetime supersymmetry with 32 supersymmetry generators, the maximum possible. In particular the massless fermionic states arising from the NS-R and R-NS sectors give two gravitini and a dilatino.

1.6.2 Type I and Heterotic String

There are three other possibilities. For example one can introduce open strings. Since open strings can combine into a closed string this theory must also contain closed strings but the presence of open strings leads to $SO(32)$ gauge fields in spacetime. This is known as the type I string. It is further complicated by the fact that the worldsheets of the strings are not oriented. The resulting theory has half as much spacetime supersymmetry as the type II theories. Indeed these days the type I string is generally viewed as an 'orientifold' of the type IIB string in the presence of so-called D9-branes. It is also thought to be dual to the heterotic $SO(32)$ string.

A more bizarre construction is to exploit the fact the left and right moving modes sectors of the string worldsheet do not talk to each other (in a closed string). Thus one could take the left moving modes of a superstring living in 10 dimensions and tensor them with the right moving modes of a bosonic string, which live in 26 dimensions. Remarkably this can be made to work and leads to two types of string theories known as the heterotic strings. These theories contain $E_8 \times E_8$ or $SO(32)$ spacetime gauge fields.

Thus the right moving sector contains 16 extra bosons. A fact about two-dimensions is that a right moving boson is the same as a pair of right moving fermions (since the Lorentz group in two dimension splits into two commuting, Abelian, parts that act on left and right movers respectively). This is known as bosonization (or sometimes fermionization, depending on your point of view). Since a right moving fermion is more natural than a right moving boson we will work with 10 scalars X^{μ} and left-moving fermions ψ^{μ}_{-}, $\mu = 0, 1, \ldots, 9$ along with 32 right moving fermions λ^{A}_{+}, $A = 1, \ldots, 32$. In this case left and right moving means:

$$\gamma_{01}\psi^{\mu}_{-} = -\psi^{\mu}_{-} \qquad \gamma_{01}\lambda^{A}_{+} = \lambda^{A}_{+} \tag{1.6.25}$$

The worldsheet action of a heterotic string is now given by

$$S = -\frac{1}{4\pi\alpha'}\int d^2\sigma\,\partial_\alpha X^\mu\partial^\beta X^\nu\eta_{\mu\nu} + i\bar{\psi}^\mu_-\gamma^\alpha\partial_\alpha\psi^\nu_-\eta_{\mu\nu} \tag{1.6.26}$$

$$+\,i\bar{\lambda}^A_+\gamma^\alpha\partial_\alpha\lambda^B_+\delta_{AB} \tag{1.6.27}$$

This has (1,0) supersymmetry:

Exercise 14 *Show that this action is invariant under*

$$\begin{aligned}\delta X^\mu &= i\bar{\varepsilon}_+\psi^\mu_- \\ \delta\psi^\mu_+ &= \gamma^\alpha\partial_\alpha X^\mu\varepsilon_+ \\ \delta\lambda^A_- &= 0\end{aligned} \tag{1.6.28}$$

provided that $\gamma_{01}\varepsilon_+ = \varepsilon_+$.

Exercise 15 *Show that the action can be written as*

$$S = -\frac{1}{4\pi\alpha'}\int d^2\sigma\,\partial_\alpha X^\mu\partial^\beta X^\nu\eta_{\mu\nu} + i(\psi^\mu_-)^T(\partial_\tau - \partial_\sigma)\psi^\nu_-\eta_{\mu\nu} \tag{1.6.29}$$

$$+\,i(\lambda^A_+)^T(\partial_\tau + \partial_\sigma)\lambda^B_+\delta_{AB} \tag{1.6.30}$$

So that ψ^μ_- *and* λ^A_+ *are indeed left and right-moving respectively.*

Quantization proceeds much as before, but with all the bells and whistles turned on. The scalars are expanded in terms left and right moving oscillators a^μ_n and $\tilde{a}^\mu_n$. The ψ^μ_- have NS and R sectors with left moving oscillators b^μ_r and d^μ_n. And λ^A_+ has an expansion in terms of right moving oscillators $\tilde{b}^A_r$ and $\tilde{d}^A_n$ for NS and R sectors respectively. In the left moving sector we have $a_{NS} = 1/2$ and $a_R = 0$, just as for the type II superstrings. In the right moving sector we have (going to light cone gauge removes two X^μ fields but none of the λ^A_+ fields)

$$\begin{aligned}\tilde{a}_{NS} &= 8\cdot\frac{1}{24} + 32\cdot\frac{1}{48} = 1 \\ \tilde{a}_R &= 8\cdot\frac{1}{24} - 32\cdot\frac{1}{24} = -1\end{aligned} \tag{1.6.31}$$

In particular we see that the right moving Ramond vacuum is massive.

Again the GSO projection is need to give modular invariance and to get rid of the tachyons. Let us look at the massless modes. For the left moving sector again we must take states of the form $b^\mu_{-\frac{1}{2}}|NS\rangle_L$ and $|R\rangle_L$, where again $|R\rangle_L$ is a degenerate spinor ground state with 8 physical states. However in the right moving sector we need only consider the NS states of the form $\tilde{a}^\mu_{-1}|NS\rangle_R$ and $b^A_{-\frac{1}{2}}b^B_{-\frac{1}{2}}|NS\rangle_R$.

Looking at the massless spacetime bosons we find the metric, dilaton and Kalb-Ramond field from $b^{\mu}_{-\frac{1}{2}}|NS\rangle_L \otimes \tilde{a}^{\mu}_{-1}|NS\rangle_R$. However we also obtain a vector state $b^{\mu}_{-\frac{1}{2}}|NS\rangle_L \otimes b^{A}_{-\frac{1}{2}}b^{B}_{-\frac{1}{2}}|NS\rangle_R$. This vector state has index structure $A_{\mu}{}^{AB}$ and can indeed be identified with a 10-dimensional gauge field. The fermionic states then give gravitini, dilatinio and gauginos. The resulting theory has 16 spacetime supersymmetries: half of the maximum of 32 that the type II theories enjoy.

Finally modular invariance and anomaly cancelation (the spacetime spectrum is chiral and for a general gauge group has anomalies) fixes the possible gauge groups to be either $E_8 \times E_8$ or $SO(32)$.

1.6.3 The Spacetime Effective Action

The superstrings have a spacetime supersymmetry and include gravity. Therefore their low energy effective actions are those of a supergravity. Such theories are so tightly constrained by their symmetries that, at least to lowest order in derivatives, their action is unique and known. In particular the bosonic section of these theories is given by

$$S_{IIA} = \frac{1}{\alpha'^4}\int d^{10}x\sqrt{-g}\left(e^{-2\phi}(R + 4(\partial\phi)^2 - \frac{1}{12}H_3^2) - \frac{1}{4}F_2^2 - \frac{1}{48}F_4^2\right) + \dots$$
$$S_{IIB} = \frac{1}{\alpha'^4}\int d^{10}x\sqrt{-g}\left(e^{-2\phi}(R + 4(\partial\phi)^2 - \frac{1}{12}H_3^2) - \frac{1}{2}F_1^2 - \frac{1}{12}F_3^2 - \frac{1}{240}F_5^2\right) + \dots$$

where the ellipsis denotes additional terms (known as Chern–Simons terms) and the subscript $n = 1, 2, 3, 4, 5$ indicates the number of anti-symmetric indices of the field strength $F_n = F_{\mu_1\dots\mu_n}$. Note that in the S_{IIB} case there is field strength $F_\mu = \partial_\mu a$ which can be thought of as arising from an additional scalar. In addition the equation of motion that arises from S_{IIB} must be supplemented by the constraint that the five-index field strength $F_{\mu_1\mu_2\mu_3\mu_4\mu_5}$ is self-dual:

$$F_{\mu_1\mu_2\mu_3\mu_4\mu_5} = \frac{1}{5!}\sqrt{-g}\varepsilon_{\mu_1\mu_2\mu_3\mu_4\mu_5\nu_1\nu_2\nu_3\nu_4\nu_5}F^{\nu_1\nu_2\nu_3\nu_4\nu_5} \quad (1.6.32)$$

We can also construct (in limited detail) the effective action for the heterotic and type I superstrings. These are fixed by supersymmetry and gauge symmetry to be of the form

$$S_I = \frac{1}{\alpha'^4}\int d^{10}x\sqrt{-g}e^{-2\phi}\left(R + 4(\partial\phi)^2 - \frac{1}{12}H_3^2 - \frac{1}{4}\mathrm{tr}(F)^2 + \dots\right) \quad (1.6.33)$$

where again the ellipsis denotes fermionic and Green–Schwarz terms that are crucial for anomaly cancelation.

When compactified on a circle the bosonic string admits a new duality known as T-duality. In the superstring case one finds that type IIA string theory on a circle of radius R is equivalent to type IIB string theory on a circle of radius α'/R. However one finds more remarkable dualities. It turns out that the type IIB supergravity has a symmetry $\phi \leftrightarrow -\phi$.[1] From the point of view of the string theory this is suggests a duality between strongly coupled strings with g_s large and weakly coupled stings with g_s small. This self-duality of the type IIB string is known as S-duality.

What happens in the strong coupling limit, $g_s \to \infty$ of the type IIA superstring? Well is it conjectured that $\sqrt{\alpha'}e^{2\phi/3}$ can be interpreted as the radius of an extra, eleventh, dimension. There is a unique supergravity theory in eleven dimensions and indeed the type IIA string effective action comes from dimensional reduction of this theory on a circle. However there is now a great deal of evidence that the whole of type IIA string theory arises as an expansion of an eleven-dimensional theory about zero-radius (in on of its dimensions). This theory is known as M-theory and is rather poorly understood. However it's existence does seem be justified. The lowest order term is in a derivative expansion is fixed by supersymmetry to be

$$S_M = \frac{1}{\kappa^9}\int d^{11}x\sqrt{-g}(R - \frac{1}{48}G_4^2) + \ldots \tag{1.6.34}$$

where again the ellipsis denotes Chern-Simons and fermionic terms. One also finds the heterotic $E_8 \times E_8$ string by compactification of M-theory on a line interval.

Furthermore it promises to be very powerful as it controls not only the strong coupling limit of the type IIA string but, as a consequence of duality, the strong coupling limit of all the five known string theories. Thus one no longer thinks of there being five separate string theories but instead one unique theory, M-theory, which contains five different perturbative descriptions depending on what one considers to be a small parameter.

References

1. Green, M.B., Schwarz, J.H., Witten, E.: Superstring Theory, vol. 1 and 2. University Press, Cambridge (1987) (Cambridge Monographs On Mathematical Physics)
2. Polchinski, J.: String Theory, vol. 1 and 2, p. 558. University Press, Cambridge (1998)
3. Zwiebach, B.: A First Course in String Theory, p. 558. University Press, Cambridge (2004)
4. Becker, K., Becker, M., Schwarz, J.H.: String Theory and M-Theory: A Modern Introduction, p. 739. Cambridge University Press, Cambridge (2004)

[1] This is simplifying things if the R-R-scalar a is not zero but a more general statement is true in that case.

Chapter 2
D-Branes and Orientifolds

Ralph Blumenhagen

2.1 The Free Boson with Boundaries

2.1.1 Boundary Conditions

We start by discussing the Boundary Conformal Field Theory of the free boson theory in order to illustrate the appearance of boundaries from a Lagrangian and geometrical point of view.[1]

Conditions for the Fields

The two-dimensional action for a free boson $X(\tau, \sigma)$ is given by

$$\mathscr{S} = \frac{1}{4\pi} \int d\sigma \; d\tau \; \Big(\big(\partial_\sigma X\big)^2 + \big(\partial_\tau X\big)^2 \Big). \tag{2.1}$$

Note that we fixed the overall normalisation constant and we slightly changed our notation such that $\tau \in (-\infty, +\infty)$ denotes the two-dimensional time coordinate and $\sigma \in [0, \pi]$ is the coordinate parametrising the distance between the boundaries.

The variation of the action (2.1) is obtained in the usual way, but now with the boundary terms taken into account. More specifically, we compute the variation as follows

R. Blumenhagen (✉)
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut),
Föhringer Ring 6, 80805 München, Germany
e-mail: blumenha@mpp.mpg.de

[1] This lecture is based on one chapter in the lecture notes [1]. Some relevant general references can be found in there. For a good guide through the vast literature we refer to the review article [2].

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_2,

$$
\begin{aligned}
\delta_X \mathscr{S} &= \frac{1}{\pi} \int d\sigma \; d\tau \; \Big(\big(\partial_\sigma X\big) \; \big(\partial_\sigma \delta X\big) + \big(\partial_\tau X\big) \; \big(\partial_\tau \delta X\big) \; \Big) \\
&= \frac{1}{\pi} \int d\sigma \; d\tau \; \Big(-\big(\partial_\sigma^2 + \partial_\tau^2\big) X \cdot \delta X + \partial_\tau \big(\partial_\tau X \cdot \delta X\big) + \partial_\sigma \big(\partial_\sigma X \cdot \delta X\big)\Big).
\end{aligned}
\tag{2.2}
$$

The equation of motion is obtained by requiring this expression to vanish for all variations δX. The vanishing of the first term in the last line leads to $\Box X = 0$ which we already obtained previously. The remaining two terms can be written as follows

$$
\begin{aligned}
&\frac{1}{\pi} \int d\sigma \; d\tau \; \Big(\; \partial_\tau \big(\partial_\tau X \cdot \delta X\big) + \partial_\sigma \big(\partial_\sigma X \cdot \delta X\big)\Big) \\
&= \frac{1}{\pi} \int d\sigma \; d\tau \; \nabla \cdot \Big(\nabla X \; \delta X \; \Big) \\
&= \frac{1}{\pi} \int_{\mathscr{B}} dl_{\mathscr{B}} \; \big(\nabla X \cdot \mathbf{n}\big) \; \delta X
\end{aligned}
$$

where we introduced $\nabla = (\partial_\tau, \partial_\sigma)^T$ and used Stokes theorem to rewrite the integral $\int d\sigma d\tau$ as an integral over the boundary $\mathscr{B}$. Furthermore, $dl_{\mathscr{B}}$ denotes the line element along the boundary and $\mathbf{n}$ is a unit vector normal to $\mathscr{B}$. In our case, the boundary is specified by $\sigma = 0$ and $\sigma = \pi$ so that $\mathbf{n} = (0, \pm 1)^T$ as well as $dl_{\mathscr{B}} = d\tau$. The vanishing of the last two terms in (2.2) can therefore be expressed as

$$
0 = \frac{1}{\pi} \int d\tau \; \big(\partial_\sigma X\big) \; \delta X \; \Big|_{\sigma=0}^{\sigma=\pi} .
$$

This equation allows for two different solutions and hence for two different boundary conditions. The first possibility is a Neumann boundary condition given by $\partial_\sigma X|_{\sigma=0,\pi} = 0$. The second possibility is a Dirichlet condition $\delta X|_{\sigma=0,\pi} = 0$ for all τ which implies $\partial_\tau X|_{\sigma=0,\pi} = 0$. In summary, the two different boundary condition for the free boson theory read as follows

$$
\boxed{
\begin{array}{rll}
\partial_\sigma X|_{\sigma=0,\pi} &= 0 & \text{Neumann condition,} \\
\delta X|_{\sigma=0,\pi} &= 0 = \partial_\tau X|_{\sigma=0,\pi} & \text{Dirichlet condition.}
\end{array}
}
\tag{2.3}
$$

Remark

Let us remark that in string theory, a hypersurface in space-time where open strings can end is called a D-brane. In order to explain this point, let us consider a theory of N free bosons $X^\mu(\tau, \sigma)$ with $\mu = 0, \ldots, N-1$ which describe the motion of a string in an N-dimensional space-time. We organise the fields in the following way

$$
\Bigg(\underbrace{X^0, X^1, \ldots, X^{r-1}}_{\text{Neumann conditions}}, \; \underbrace{X^r, \ldots, X^{N-1}}_{\text{Dirichlet conditions}} \Bigg) ,
$$

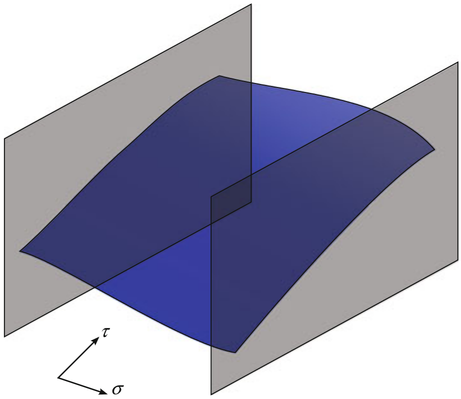


Fig. 2.1 Two-dimensional surface with boundaries which can be interpreted as an open string world-sheet stretched between two D-branes

where r denotes the number of bosons with Neumann boundary conditions leaving $(N - r)$ bosons with Dirichlet conditions.

Let us now focus on one endpoint of the open string, say at $\sigma = 0$. A Dirichlet boundary condition for X^μ reads $\delta X^\mu|_{\sigma=0} = 0$ which means that the endpoint of the open string is fixed to a particular value $x_0^\mu = \text{const}$. However, in case of Neumann boundary conditions, there is no restriction on the position of the string endpoint which can therefore take any value. Clearly, since the string moves in time, there are Neumann conditions for the time coordinate X^0. Then, the r-dimensional hypersurface in space-time described by $X^\mu = x_0^\mu = \text{const.}$ for $\mu = r, \ldots, N-1$ is called a D$(r-1)$-brane where D stands for Dirichlet.

As an example, take $N = 3$ and consider Fig. 2.1 where we see a world-sheet of an open string stretched between two D1-branes.

Conditions for the Laurent Modes

Above, we have considered the BCFT in terms of the real variables (τ, σ) which was convenient in order to arrive at (2.3). However, for more advanced studies a description in terms of complex variables is very useful. Similarly as before, a mapping from the infinite strip described by the real variables (τ, σ) to the complex upper half plane H^+ is achieved by $z = \exp(\tau + i\sigma)$. Note in particular, as illustrated in Fig. 2.2, the boundary $\sigma = 0, \pi$ is mapped to the real axis $z = \bar{z}$.

Having this map in mind, we can express the boundary conditions (2.3) for the field $X(\sigma, \tau)$ in terms of the corresponding Laurent modes. Recalling that $j(z) = i\,\partial X(z, \bar{z})$, we find

$$\partial_\sigma X = i\big(\partial - \bar{\partial}\big)X = j(z) - \bar{j}(\bar{z}) = \sum_{n\in\mathbb{Z}} \big(j_n\, z^{-n-1} - \bar{j}_n\, \bar{z}^{-n-1} \big),$$

$$i \cdot \partial_\tau X = i\big(\partial + \bar{\partial}\big)X = j(z) + \bar{j}(\bar{z}) = \sum_{n\in\mathbb{Z}} \big(j_n\, z^{-n-1} + \bar{j}_n\, \bar{z}^{-n-1} \big),$$

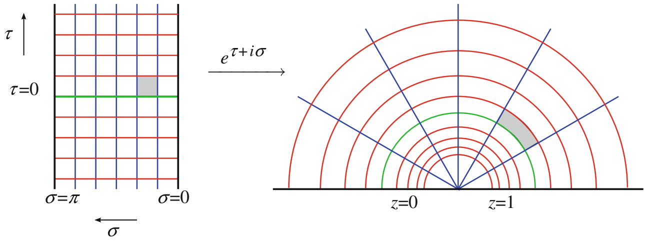


Fig. 2.2 Illustration of the map $z = \exp(\tau + i\sigma)$ from the infinite strip to the complex upper half plane H^+

where we used $\partial = \frac{1}{2}(\partial_0 - i\partial_1)$ and $\overline{\partial} = \frac{1}{2}(\partial_0 + i\partial_1)$. For transforming the right-hand side of these equations as $z \mapsto e^w$ with $w = \tau + i\sigma$, we employ that $j(z)$ is a primary field of conformal dimension $h = 1$. In particular, recalling that conformal transformations act on primary fields of dimension $(h, \overline{h})$ as

$$\phi(z, \overline{z}) \mapsto \phi'(z, \overline{z}) = \left(\frac{\partial f}{\partial z}\right)^h \left(\frac{\partial \overline{f}}{\partial \overline{z}}\right)^{\overline{h}} \phi\left(f(z), \overline{f}(\overline{z})\right), \tag{2.4}$$

we have $j(z) = \left(\frac{\partial z}{\partial w}\right)^1 j(w) = z\, j(w)$ leading to

$$\begin{aligned} \partial_\sigma X &= \sum_{n\in\mathbb{Z}} \left(j_n\, e^{-n(\tau+i\sigma)} - \overline{j}_n\, e^{-n(\tau-i\sigma)} \right), \\ i \cdot \partial_\tau X &= \sum_{n\in\mathbb{Z}} \left(j_n\, e^{-n(\tau+i\sigma)} + \overline{j}_n\, e^{-n(\tau-i\sigma)} \right). \end{aligned} \tag{2.5}$$

The Neumann as well as the Dirichlet boundary conditions at $\sigma = 0$ are then easily obtained as

$$\begin{aligned} \partial_\sigma X \big|_{\sigma=0} &= \sum_{n\in\mathbb{Z}} \left(j_n - \overline{j}_n\right) e^{-n\,\tau} = 0, \\ \partial_\tau X \big|_{\sigma=0} &= \sum_{n\in\mathbb{Z}} \left(j_n + \overline{j}_n\right) e^{-n\,\tau} = 0. \end{aligned}$$

Since for generic τ the summands above are linearly independent, these two equations are respectively solved by $j_n \pm \overline{j}_n = 0$ for all n. In summary, we note that boundaries introduce relations between the chiral and anti-chiral modes of the conformal fields which read

$$\boxed{\begin{array}{lll} j_n - \overline{j}_n = 0 & & \text{Neumann condition,} \\ j_n + \overline{j}_n = 0, & (\pi_0 = 0) & \text{Dirichlet condition.} \end{array}} \tag{2.6}$$

From a string theory point of view, (2.6) implies that an open string has only half the degrees of freedom of a closed string.

A computation of the center of mass for the open string gives

$$\pi_0 = \frac{1}{2}\, j_0 = \frac{1}{2}\, \bar{j}_0. \tag{2.7}$$

In view of (2.6), we thus see that there are no restrictions on π_0 for Neumann boundary conditions and so the endpoints of the string are free to move along the D-brane. For Dirichlet conditions on the other hand, we have $\pi_0 = 0$ implying that the endpoints are fixed.

Combined Boundary Condition

In the previous paragraph, we have considered the boundary at $\sigma = 0$. Let us now turn to the other boundary at $\sigma = \pi$. Performing the same steps as before, we see that Neumann–Neumann as well as Dirichlet–Dirichlet conditions are characterised by the constraints found in (2.6).

However, mixed boundary conditions, e.g. Neumann-Dirichlet, require a modification. In particular, $j_n - \bar{j}_n = 0$ at $\sigma = 0$ and $j_n + \bar{j}_n e^{-2in\sigma} = 0$ at $\sigma = \pi$ can only be solved for $n \in \mathbb{Z} + \frac{1}{2}$. All possible combinations of boundary conditions are then summarised as

$$\begin{aligned}
& j_n - \bar{j}_n = 0, && n \in \mathbb{Z} && \text{Neumann–Neumann,} \\
& j_n - \bar{j}_n = 0, && n \in \mathbb{Z} + \frac{1}{2} && \text{Neumann–Dirichlet,} \\
& j_n + \bar{j}_n = 0, && n \in \mathbb{Z} + \frac{1}{2} && \text{Dirichlet–Neumann,} \\
& j_n + \bar{j}_n = 0, && n \in \mathbb{Z} && \text{Dirichlet–Dirichlet.}
\end{aligned}$$

Solutions to the Boundary Condition

Next, let us determine the solutions to the boundary conditions stated above. First, we integrate equations (2.5) to obtain $X(\tau, \sigma)$ in the closed sector

$$X\big(\tau, \sigma\big) = x_0 - i\,\big(\tau + i\sigma\big)\, j_0 - i\,\big(\tau - i\sigma\big)\, \bar{j}_0 + \sum_{n \neq 0} \frac{i}{n} \Big(j_n e^{-n(\tau + i\sigma)} + \bar{j}_n e^{-n(\tau - i\sigma)} \Big) \tag{2.8}$$

where x_0 is an integration constant. We then implement the boundary conditions to project onto the open sector. For the Neumann–Neumann case we find

$$X^{(\mathrm{N,N})}\big(\tau, \sigma\big) = x_0 - 2i\tau j_0 + 2i \sum_{n \neq 0} \frac{j_n}{n}\, e^{-n\,\tau} \cos\big(n\sigma\big),$$

and for the Dirichlet–Dirichlet case we obtain along the same lines

$$X^{(\mathrm{D},\mathrm{D})}\big(\tau,\sigma\big) = x_0 + 2\sigma j_0 + 2\sum_{n\neq 0} \frac{j_n}{n}\, e^{-n\,\tau}\,\sin\big(n\sigma\big).$$

Having arrived at this solution, we can become more concrete about the Dirichlet–Dirichlet boundary conditions. We impose that $X(\tau,\sigma=0)=x_0^a$ and $X(\tau,\sigma=\pi)=x_0^b$, which means that the endpoints of the string are fixed at positions x_0^a and x_0^b. Using the explicit solution for $X^{(\mathrm{D},\mathrm{D})}(\tau,\sigma)$, we obtain the relation

$$\boxed{j_0 = \frac{x_0^b - x_0^a}{2\pi}}\,. \tag{2.9}$$

Finally, for completeness, the solutions for the case of mixed Neumann–Dirichlet boundary conditions read as follows

$$X^{(\mathrm{N},\mathrm{D})}\big(\tau,\sigma\big) = x_0 + 2\,i \sum_{n\in\mathbb{Z}+\frac{1}{2}} \frac{j_n}{n}\, e^{-n\,\tau}\cos\big(n\sigma\big),$$
$$X^{(\mathrm{D},\mathrm{N})}\big(\tau,\sigma\big) = x_0 + 2 \sum_{n\in\mathbb{Z}+\frac{1}{2}} \frac{j_n}{n}\, e^{-n\,\tau}\sin\big(n\sigma\big)\;.$$

Conformal Symmetry

Let us remark that equations (2.6) apply to the Laurent modes of the two $U(1)$ currents $j(z)$ and $\bar{j}(\bar{z})$ of the free boson theory leaving only a diagonal $U(1)$ symmetry. However, in addition there is always the conformal symmetry generated by the energy-momentum tensor. Since boundaries in general break certain symmetries, we expect also restrictions on the Laurent modes of energy-momentum tensor.

Indeed, recalling that $T(z)$ and $\overline{T}(\bar{z})$ can be expressed in terms of the currents $j(z)$ and $\bar{j}(\bar{z})$ in the following way

$$T(z) = \frac{1}{2}\,N\big(jj\big)(z), \qquad \overline{T}(\bar{z}) = \frac{1}{2}\,N\big(\bar{j}\,\bar{j}\big)(\bar{z}),$$

we find that the Neumann as well as the Dirichlet boundary conditions (2.6) imply for $L_n = \frac{1}{2}N(jj)_n$ that

$$\boxed{L_n - \overline{L}_n = 0}\,. \tag{2.10}$$

Let us emphasise that this condition can be expressed as $T(z) = \overline{T}(\bar{z})$ which in particular means the central charges of the holomorphic and anti-holomorphic theories have to be equal, i.e. $c=\bar{c}$. For string theory, this observation has the immediate implication that boundaries, that is D-branes, can only be defined for the Type II Superstring Theories, as opposed to the heterotic string theories.

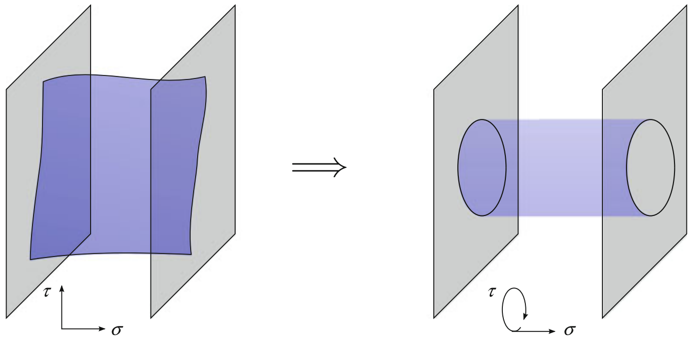


Fig. 2.3 Illustration how the cylinder partition function is obtained from the infinite strip by cutting out a finite piece and identifying the ends

2.1.2 Partition Function

Definition

Let us now consider the one-loop partition function for BCFTs. To do so, we first review the construction for the case without boundaries and then compare to the present situation.

- The one-loop partition function for CFTs without boundaries is defined as follows. We start from a theory defined on the infinite cylinder described by (τ, σ) where σ is periodic and $\tau \in (-\infty, +\infty)$. Next, we impose periodicity conditions also on the time coordinate τ yielding the topology of a torus. The partition function is then determined as

$$\mathscr{Z}(\tau, \overline{\tau}) = \mathrm{Tr}_{\mathscr{H}}\left(q^{L_0 - \frac{c}{24}} \overline{q}^{\bar{L}_0 - \frac{\bar{c}}{24}}\right). \tag{2.11}$$

- In the present case, the space coordinate σ is not periodic and thus we start from a theory defined on the infinite strip given by $\sigma \in [0, \pi]$ and $\tau \in (-\infty, +\infty)$. For the definition of the one-loop partition function, we again make the time coordinate τ periodic leaving us with the topology of a cylinder instead of a torus. This is illustrated in Fig. 2.3.
- Similarly to the modular parameter of the torus, there is a modular parameter t with $0 \leq t < \infty$ parametrising different cylinders. The inequivalent cylinders are described by $\{(\tau, \sigma) : 0 \leq \sigma \leq \pi, 0 \leq \tau \leq t\}$.

For the partition function, we need to determine the operator generating translations in time circling the cylinder once along the τ direction. Because boundaries lead to an identification of the left- and right-moving sector as required by (2.10), we see that this operator is the Hamiltonian say in the open sector

$$H_{\mathrm{open}} = (L_{\mathrm{cyl.}})_0 = L_0 - \frac{c}{24},$$

which we inferred from the closed sector Hamiltonian $H_{\text{closed}} = (L_{\text{cyl.}})_0 + (\overline{L}_{\text{cyl.}})_0$. In analogy to the case of the torus partition function, we then define the cylinder partition function as $\mathscr{Z} = \text{Tr}\exp(-2\pi t H_{\text{open}})$ which can be brought into the following form

$$\mathscr{Z}^{\mathscr{C}}(t) = \text{Tr}_{\mathscr{H}_{\mathscr{B}}}\Big(q^{L_0 - \frac{c}{24}} \Big) \qquad \text{where} \qquad q = e^{-2\pi t}.$$

Here, the superscript $\mathscr{C}$ on $\mathscr{Z}$ indicates the cylinder partition function and $\mathscr{H}_{\mathscr{B}}$ denotes the Hilbert space of all states satisfying one of the boundary conditions (2.6). Clearly, from a string theory point of view, this is just the Hilbert space of an open string.

Free Boson I: Cylinder Partition Function (Loop-Channel)

We close this section by determining the cylinder partition function for the free boson. For the free boson, the Laurent modes of the energy-momentum tensor are written using the modes of the current $j(z) = i\,\partial X(z)$. In particular, we have

$$L_0 = \frac{1}{2}\, j_0\, j_0 + \sum_{k\geq 1} j_{-k}\, j_k.$$

Since the current $j(z)$ is a field of conformal dimension one, we find that $j_n|0\rangle = 0$ for $n > -1$ and that states in the Hilbert space have the following form

$$\big|n_1, n_2, n_3, \ldots\big\rangle = j_{-1}^{n_1}\, j_{-2}^{n_2}\, j_{-3}^{n_3} \ldots \big|0\big\rangle \qquad \text{with} \qquad n_i \geq 0 \tag{2.12}$$

and $n_i \in \mathbb{Z}$. The current algebra for the Laurent modes reads

$$\big[j_m, j_n\big] = m\, \delta_{m,-n}.$$

Next, let us compute the action of L_0 on a state (2.12). Clearly, j_0 commutes with all j_{-k} and let is first assume that it annihilates the vacuum. For the other terms we calculate

$$\big[j_{-k} j_k,\ j_{-k}^{n_k}\big] = n_k\, k\, j_{-k}^{n_k}, \tag{2.13}$$

and so we find for the zero Laurent mode of the energy-momentum tensor that

$$L_0\, \big|n_1, n_2, n_3, \ldots\big\rangle = \sum_{k\geq 1} j_{-1}^{n_1}\, j_{-2}^{n_2} \ldots \big(j_{-k} j_k\big)\, j_{-k}^{n_k} \ldots \big|0\big\rangle = \sum_{k\geq 1} k\, n_k\, \big|n_1, n_2, n_3, \ldots\big\rangle.$$

We will utilize this last result in the calculation of the partition function where for simplicity we only focus on the holomorphic part. We compute

$$
\begin{aligned}
&\mathrm{Tr}\left(q^{L_0-\frac{c}{24}}\right)\\
&= q^{-\frac{1}{24}} \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty} \ldots \langle n_1, n_2, n_3, \ldots| \sum_{p=0}^{\infty} \frac{1}{p!}(2\pi\tau)^p \, (L_0)^p \, |n_1, n_2, n_3, \ldots\rangle\\
&= q^{-\frac{1}{24}} \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty} \ldots \langle n_1, n_2, n_3, \ldots| \sum_{p=0}^{\infty} \frac{1}{p!}(2\pi\tau)^p \, \Big(\sum_{k=1}^{\infty} k\, n_k\Big)^p \, |n_1, n_2, n_3, \ldots\rangle\\
&= q^{-\frac{1}{24}} \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\sum_{n_3=0}^{\infty} \ldots \left(q^{1\cdot n_1}\cdot q^{2\cdot n_2}\cdot q^{3\cdot n_3}\cdot\;\ldots\right)\\
&= q^{-\frac{1}{24}} \left(\sum_{n_1=0}^{\infty} q^{1\,n_1}\right)\cdot\left(\sum_{n_2=0}^{\infty} q^{2\,n_2}\right)\cdot\left(\sum_{n_3=0}^{\infty} q^{3\,n_3}\right)\cdot\;\ldots\\
&= q^{-\frac{1}{24}} \prod_{k=1}^{\infty}\sum_{n_k=0}^{\infty} q^{k\,n_k} = q^{-\frac{1}{24}} \prod_{k=1}^{\infty}\frac{1}{1-q^k}
\end{aligned}
$$

where in the last step we employed the result for the infinite geometric series and the ellipsis indicate that the structure extends to infinity. We then define the Dedekind η-function as

$$
\boxed{\eta(\tau) = q^{\frac{1}{24}} \prod_{n=1}^{\infty}\left(1-q^n\right)} \tag{2.14}
$$

so that

$$
\mathrm{Tr}_{\mathscr{H}_{\mathscr{B}}}\left(\, q^{L_0-\frac{c}{24}}\,\right)\Big|_{\text{without } j_0} = \frac{1}{\eta\;(it)}.
$$

However, recall that we have assumed the action of j_0 on the vacuum to vanish which is in general not applicable. Taking into account the effect of j_0, we now study the three different cases of boundary conditions in turn.

- For the case of Neumann–Neumann boundary conditions, the momentum mode $\pi_0 = \frac{1}{2}\, j_0$ is unconstrained and in principle contributes to the trace. Since it is a continuous variable, the sum is replaced by an integral

$$
\mathrm{Tr}_{\mathscr{H}_{\mathscr{B}}}\left(q^{\frac{1}{2}j_0^2}\right) = \sum_{n_0} \langle n_0|\, e^{-\pi t\, j_0^2}\, |n_0\rangle = \sum_{n_0} e^{-\pi t\, n_0^2} \quad\longrightarrow\quad \int_{-\infty}^{\infty} d\pi_0\, e^{-4\pi t\,\pi_0^2},
$$

  where we utilised $n_0 = 2\pi_0$. Evaluating this Gaussian integral leads to the following additional factor for the partition function

$$
\frac{1}{2\sqrt{t}}\,. \tag{2.15}
$$

- For the Dirichlet–Dirichlet case, we have seen in equation (2.9) that j_0 is related to the positions of the string endpoints. Therefore, we have a contribution to the partition function of the form

$$q^{\frac{1}{2} j_0^2} = \exp\left(-2\pi t \, \frac{1}{2}\left(\frac{x_0^b - x_0^a}{2\pi}\right)^2\right) = \exp\left(-\frac{t}{4\pi}\left(x_0^b - x_0^a\right)^2\right).$$

- Finally, for the case of mixed Neumann–Dirichlet boundary conditions, the Laurent modes j_n take half-integer values for n. We do not present a detailed calculation for this case. We just mention that there is a twisted sector where the Laurent modes j_n also take half-integer values for n. It is then possible to extract $\mathrm{Tr}_{n\in\mathbb{Z}+\frac{1}{2}}\left(q^{L_0-\frac{c}{24}}\right)$ giving us the partition function in the present case.

In summary, the cylinder partition functions for the example of the free boson read

$$\begin{aligned}
\mathcal{Z}^{\mathcal{C}\,(\mathrm{D,D})}_{\mathrm{bos.}}(t) &= \exp\left(-\frac{t}{4\pi}\left(x_0^b - x_0^a\right)^2\right)\frac{1}{\eta\,(it)},\\
\mathcal{Z}^{\mathcal{C}\,(\mathrm{N,N})}_{\mathrm{bos.}}(t) &= \frac{1}{2\sqrt{t}}\,\frac{1}{\eta\,(it)},\\
\mathcal{Z}^{\mathcal{C}\,(\mathrm{mixed})}_{\mathrm{bos.}}(t) &= \sqrt{\frac{\eta\,(it)}{\vartheta_4(it)}}.
\end{aligned} \tag{2.16}$$

2.2 Boundary States for the Free Boson

In the last section, we have described the boundaries for the free boson CFT implicitly via the boundary conditions for the fields. However, in an abstract CFT usually there is no Lagrangian formulation available and no boundary terms will arise from a variational principle. Therefore, to proceed, we need a more inherent formulation of a boundary.

In the following, we first illustrate the construction of so-called boundary states for the example of the free boson and in the next section, we generalise the structure to Rational Conformal Field Theories with boundaries.

2.2.1 Boundary Conditions

Boundary States

Let us start with the following observation. As it is illustrated in Fig. 2.4, by interchanging τ and σ, we can interpret the cylinder partition function of the Boundary Conformal Field Theory on the left-hand side as a tree-level amplitude of the underlying theory shown on the right-hand side. From a string theory point of view, the

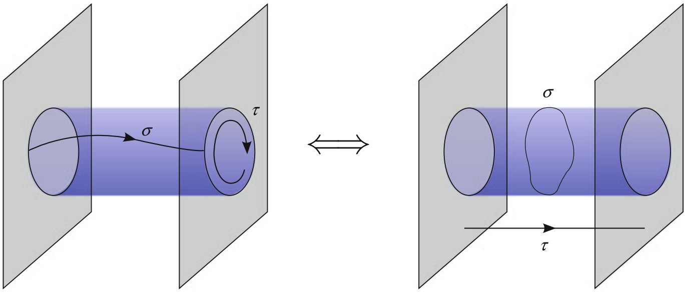


Fig. 2.4 Illustration of world-sheet duality relating the cylinder amplitude in the open and closed sector

tree-level amplitude describes the emission of a closed string at boundary A which propagates to boundary B and is absorbed there. Thus, a boundary can be interpreted as an object, which couples to closed strings. Note that in order to simplify our notation, we call the sector of the BCFT *open* and the sector of the underlying CFT *closed*. The relation above then reads

$$(\sigma, \tau)_{\text{open}} \quad \longleftrightarrow \quad (\tau, \sigma)_{\text{closed}}, \tag{2.17}$$

which in string theory is known as the world-sheet duality between open and closed strings.

The boundary for the closed sector can be described by a coherent state in the Hilbert space $\mathscr{H} \otimes \overline{\mathscr{H}}$ which takes the general form

$$\big|B\big\rangle = \sum_{i,\bar{j}\in\mathscr{H}\otimes\overline{\mathscr{H}}} \alpha_{i\bar{j}} \, \big|i, \bar{j}\big\rangle.$$

Here $i, \bar{j}$ label the states in the holomorphic and anti-holomorphic sector of $\mathscr{H} \otimes \overline{\mathscr{H}}$, and the coefficients $\alpha_{i\bar{j}}$ encode the *strength* of how the closed string mode $|i, \bar{j}\rangle$ couples to the boundary $|B\rangle$. Such a coherent state is called a *boundary state* and provides the CFT description of a D-brane in string theory.

Boundary Conditions

Let us now translate the boundary conditions (2.3) into the picture of boundary states. By using relation (2.17), we readily obtain

$$\boxed{\begin{array}{ll} \partial_\tau X_{\text{closed}}\big|_{\tau=0} \, \big|B_{\text{N}}\big\rangle = 0 & \quad \text{Neumann condition,} \\ \partial_\sigma X_{\text{closed}}\big|_{\tau=0} \, \big|B_{\text{D}}\big\rangle = 0 & \quad \text{Dirichlet condition.} \end{array}} \tag{2.18}$$

Next, for the free boson theory we would like to express the boundary conditions (2.18) of a boundary state in terms of the Laurent modes. To do so, we recall (2.5) and set $\tau = 0$ to obtain

$$
\begin{aligned}
i \cdot \partial_\tau X_{\text{closed}}\big|_{\tau=0} &= \sum_{n\in\mathbb{Z}} \big(j_n\, e^{-in\sigma} + \bar{j}_n\, e^{+in\sigma} \big), \\
\partial_\sigma X_{\text{closed}}\big|_{\tau=0} &= \sum_{n\in\mathbb{Z}} \big(j_n\, e^{-in\sigma} - \bar{j}_n\, e^{+in\sigma} \big).
\end{aligned}
\tag{2.19}
$$

We then relabel $n \to -n$ in the second term of each line and observe again that for generic σ, the summands are linearly independent. Therefore, the boundary conditions (2.18) expressed in terms of the Laurent modes read

$$
\boxed{
\begin{aligned}
&\big(j_n + \bar{j}_{-n}\big)\, \big|B_{\text{N}}\big\rangle = 0, \quad \big(\, \pi_0\, |B_{\text{N}}\rangle = 0 \,\big) && \text{Neumann condition,} \\
&\big(j_n - \bar{j}_{-n}\big)\, \big|B_{\text{D}}\big\rangle = 0 && \text{Dirichlet condition,}
\end{aligned}
}
\tag{2.20}
$$

for each n. Such conditions relating the chiral and anti-chiral modes acting on the boundary state are called *gluing conditions*. Note that for the case of Neumann boundary conditions, in the string theory picture the relation $\pi_0 = 0$ means that there is no momentum transfer through the boundary. On the other hand, for Dirichlet conditions there is no restriction on π_0.

Solutions to the Gluing Conditions

Next, we are going to state the solutions for the gluing conditions for the example of the free boson and verify them thereafter. For now, let us ignore the constraints on j_0. We will come back to this issue later.

The boundary states for Neumann and Dirichlet conditions in terms of the Laurent modes j_n and $\bar{j}_n$ read

$$
\boxed{
\begin{aligned}
&\big|B_{\text{N}}\big\rangle = \frac{1}{\mathcal{N}_{\text{N}}} \exp\!\bigg(- \sum_{k=1}^{\infty} \frac{1}{k}\, j_{-k}\, \bar{j}_{-k} \bigg) \big|\, 0\big\rangle && \text{Neumann condition,} \\
&\big|B_{\text{D}}\big\rangle = \frac{1}{\mathcal{N}_{\text{D}}} \exp\!\bigg(+ \sum_{k=1}^{\infty} \frac{1}{k}\, j_{-k}\, \bar{j}_{-k} \bigg) \big|\, 0\big\rangle && \text{Dirichlet condition,}
\end{aligned}
}
\tag{2.21}
$$

where $\mathcal{N}_{\text{N}}$ and $\mathcal{N}_{\text{D}}$ are normalisation constants to be fixed later. One possibility to verify the boundary states is to straightforwardly evaluate the gluing conditions (2.20) for the solutions (2.21) explicitly. However, in order to highlight the underlying structure, we will take a slightly different approach.

Construction of Boundary States

In the following, we focus on a boundary state with Neumann conditions but comment on the Dirichlet case at the end. To start, we rewrite the Neumann boundary state in (2.21) as

$$\begin{aligned}
\big|B_{\rm N}\big\rangle &= \frac{1}{\mathcal{N}_{\rm N}}\, \exp\!\bigg(-\sum_{k=1}^{\infty}\frac{1}{k}\, j_{-k}\,\bar{j}_{-k}\bigg)\big|0\big\rangle \\
&= \frac{1}{\mathcal{N}_{\rm N}} \prod_{k=1}^{\infty}\sum_{m=0}^{\infty} \frac{1}{\sqrt{m!}}\left(\frac{j_{-k}}{\sqrt{k}}\right)^{m}|0\rangle \otimes \frac{1}{\sqrt{m!}}\left(\frac{-\bar{j}_{-k}}{\sqrt{k}}\right)^{m}|\bar{0}\rangle \\
&= \frac{1}{\mathcal{N}_{\rm N}} \sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\cdots\prod_{k=1}^{\infty} \frac{1}{\sqrt{m_k!}}\left(\frac{j_{-k}}{\sqrt{k}}\right)^{m_k}|0\rangle \otimes \frac{1}{\sqrt{m_k!}}\left(\frac{-\bar{j}_{-k}}{\sqrt{k}}\right)^{m_k}|\bar{0}\rangle,
\end{aligned} \tag{2.22}$$

where we first have written the sum in the exponential as a product and then we expressed the exponential as an infinite series. Next, we note that the following states form a complete orthonormal basis for all states constructed out of the Laurent modes j_{-k}

$$\big|\mathbf{m}\big\rangle = \big|m_1, m_2, \ldots\big\rangle = \prod_{k=1}^{\infty} \frac{1}{\sqrt{m_k!}}\left(\frac{j_{-k}}{\sqrt{k}}\right)^{m_k}|0\rangle. \tag{2.23}$$

The orthonormal property can be seen by computing

$$\big\langle \mathbf{n} \mid \mathbf{m}\big\rangle = \prod_{k=1}^{\infty} \frac{1}{\sqrt{n_k!\, m_k!}}\, \frac{1}{\sqrt{k}^{\,n_k+m_k}}\, \langle 0|\, j_{+k}^{n_k}\, j_{-k}^{m_k}\, |0\rangle_k = \prod_{k=1}^{\infty} \delta_{n_k,m_k},$$

where we used that

$$\langle 0|\, j_{+k}^{n}\, j_{-k}^{m}\, |0\rangle = k\, n\, \langle 0|\, j_{+k}^{n-1}\, j_{-k}^{m-1}\, |0\rangle = \delta_{m,n}\, k^n\, n!.$$

We now introduce an operator U mapping the chiral Hilbert space to its charge conjugate $U : \mathcal{H} \to \mathcal{H}^+$ and similarly for the anti-chiral sector. In particular, the action of U reads

$$U\, j_k\, U^{-1} = -j_k = -\big(j_{-k}\big)^{\dagger}, \qquad U\, \bar{j}_k\, U^{-1} = -\bar{j}_k = -\big(\bar{j}_{-k}\big)^{\dagger}, \qquad U\, c\, U^{-1} = c^*,$$

where c is a constant and * denotes complex conjugation. In the present example, the ground state $|0\rangle$ is non-degenerate and is left invariant by U.[2] Knowing these properties, we can show that U is anti-unitary. For this purpose, we expand a general state as $|a\rangle = \sum_{\mathbf{m}} A_{\mathbf{m}} |\mathbf{m}\rangle$ and compute

$$\begin{aligned}
U\, \big|a\big\rangle &= \sum_{\mathbf{m}}\, U\, A_{\mathbf{m}}\, U^{-1} \prod_{k=1}^{\infty} \frac{1}{\sqrt{m_k!}}\left(\frac{U\, j_{-k}\, U^{-1}}{\sqrt{k}}\right)^{m_k} U\, \big|0\big\rangle \\
&= \sum_{\mathbf{m}} A^*_{\mathbf{m}} \prod_{k=1}^{\infty} (-1)^{m_k} \big|\mathbf{m}\big\rangle,
\end{aligned} \tag{2.24}$$

where $\mathbf{m}$ denotes the multi-index $\{m_1, m_2, \ldots\}$. By using that $|\mathbf{m}\rangle$ and $|\mathbf{n}\rangle$ form an orthonormal basis, we can now show that U is anti-unitary

2 For degenerate ground states a non-trivial action on the ground state might need to be defined.

$$\left\langle\, Ub \,\middle|\, Ua \,\right\rangle = \sum_{\mathbf{n},\mathbf{m}} \langle \mathbf{n} \mid B_{\mathbf{n}} \prod_{k=1}^{\infty} (-1)^{n_k+m_k} A^*_{\mathbf{m}} \,|\mathbf{m}\rangle = \sum_{\mathbf{m}} A^*_{\mathbf{m}}\, B_{\mathbf{m}} = \left\langle\, a \,\middle|\, b \,\right\rangle.$$

After introducing an orthonormal basis and the anti-unitary operator U, we now express (2.22) in a more general way which will simplify and generalise the following calculations

$$\boxed{|B\rangle = \frac{1}{\mathscr{N}} \sum_{\mathbf{m}} |\mathbf{m}\rangle \otimes |U\, \overline{\mathbf{m}}\rangle}.$$

Verification of the Gluing Conditions

In order to verify the gluing conditions (2.20) for Neumann boundary states, we note that these have to be satisfied also when an arbitrary state $\langle\, \overline{a} \mid \otimes \langle\, b \mid$ is multiplied from the left. We then calculate

$$\begin{aligned}\langle\, \overline{a} \mid \otimes \langle\, b \mid j_n + \overline{j}_{-n}\; |B\rangle &= \frac{1}{\mathscr{N}} \sum_{\mathbf{m}} \langle\, \overline{a} \mid \otimes \langle\, b \mid j_n + \overline{j}_{-n}\; |\mathbf{m}\rangle \otimes |U\, \overline{\mathbf{m}}\rangle \\ &= \frac{1}{\mathscr{N}} \sum_{\mathbf{m}} \langle\, b \mid j_n\; |\mathbf{m}\rangle \langle\, \overline{a}\; |U\, \overline{\mathbf{m}}\rangle + \langle\, b\; |\mathbf{m}\rangle \langle\, \overline{a} \mid \overline{j}_{-n}\; |U\, \overline{\mathbf{m}}\rangle.\end{aligned}$$

Next, due to the identifications on the boundary, the holomorphic and the anti-holomorphic algebra are identical. We can therefore replace matrix elements in the anti-holomorphic sector by those in the holomorphic sector. Using finally the anti-unitarity of U and that $\sum_{\mathbf{m}} |\mathbf{m}\rangle\langle\mathbf{m}| = \mathbb{1}$, we find

$$\begin{aligned}&\langle\overline{a} \mid \otimes \langle b \mid j_n + \overline{j}_{-n}\; |B\rangle \\ &\qquad= \frac{1}{\mathscr{N}} \sum_{\mathbf{m}} \langle\, b \mid j_n\; |\mathbf{m}\rangle \langle\, a\; |U\, \mathbf{m}\rangle + \langle\, b\; |\mathbf{m}\rangle \langle\, a \mid j_{-n}\; |U\, \mathbf{m}\rangle \\ &\qquad= \frac{1}{\mathscr{N}} \sum_{\mathbf{m}} \langle\, b \mid j_n\; |\mathbf{m}\rangle \langle\, \mathbf{m}\; |U^{-1}a\rangle + \langle\, b\; |\mathbf{m}\rangle \langle\, \mathbf{m} \mid (-j_n)\; |U^{-1}a\rangle \\ &\qquad= \frac{1}{\mathscr{N}} \Big(\langle\, b \mid j_n\; |U^{-1}a\rangle - \langle\, b \mid j_n\; |U^{-1}a\rangle \Big) = 0.\end{aligned}$$

Therefore, we have verified that the Neumann boundary state in (2.21) is indeed a solution to the corresponding gluing condition in (2.20).

For the case of Dirichlet boundary conditions, the action of U on the Laurent modes j_n and $\overline{j}_n$ is chosen with a + sign while we still require U to be anti-unitary, i.e. $U\; c\; U^{-1} = c^*$. The calculation is then very similar to the Neumann case presented here. Note furthermore, the construction of boundary states and the verification of the gluing conditions is also applicable for more general CFTs, for instance RCFTs, which we will consider in Sect. 2.3.

Momentum Dependence of Boundary States

In (2.7) we gave the result for the center of mass for an open string. This differs from the closed string case by a factor $\frac{1}{2}$ due to the fact that open strings have by convention length π while closed string have length 2π. In the following, the relation between j_0, $\bar{j}_0$ and π_0 should be clear from the context, but let us summarise that

$$\big(\pi_0\big)_{\text{closed}} = j_0 = \bar{j}_0, \qquad \big(\pi_0\big)_{\text{open}} = \frac{1}{2}\, j_0 = \frac{1}{2}\, \bar{j}_0. \tag{2.25}$$

From a string theory point of view, in addition to the boundary conditions (2.20) there is a further natural constraint on a boundary state with Dirichlet conditions. Namely, the closed string at time $\tau = 0$ is located at the boundary at position x_0^a. We therefore impose

$$X_{\text{closed}}\big(\tau = 0, \sigma\big)\ \big|B_{\text{D}}\big\rangle = x_0^a\ \big|B_{\text{D}}\big\rangle$$

and similarly for $\tau = \pi$. An easy way to realise this constraint is to perform a Fourier transformation from momentum space $|B_{\text{D}}, \pi_0\rangle$ to the position space. Concretely, this reads

$$\big|B_{\text{D}}, x_0^a\big\rangle = \int d\pi_0\ e^{i\pi_0 x_0^a}\ \big|B_{\text{D}}, \pi_0\big\rangle.$$

For the boundary state with Neumann conditions, we have $\pi_0 = 0$ and in position space, there is no definite value for x_0. We thus omit this label.

Conformal Symmetry

In studying the example of the free boson, we have expressed all important quantities in terms of the $U(1)$ current modes j_n and $\bar{j}_n$. However, in more general CFTs such additional symmetries may not be present but the conformal symmetry generated by the energy-momentum tensors always is. In view of generalisations of our present example, let us therefore determine the boundary conditions of the boundary states in terms of the Laurent modes L_n and $\overline{L}_n$.

Mainly guided by the final result, let us compute the following expression by employing that $T(z) = \frac{1}{2}N(jj)(z)$ which implies $L_n = \frac{1}{2}\sum_{k>-1} j_{n-k} j_k + \frac{1}{2}\sum_{k\leq -1} j_k j_{n-k}$

$$\begin{aligned}
&\left(L_n-\overline{L}_{-n}\right)\,\left|B_{\rm N,D}\right\rangle\\
&=\frac{1}{2}\left(\sum_{k>-1}\left(j_{n-k}j_k-\overline{j}_{-n-k}\overline{j}_k\right)+\sum_{k\le -1}\left(j_kj_{n-k}-\overline{j}_k\overline{j}_{-n-k}\right)\right)\left|B_{\rm N,D}\right\rangle\\
&=\frac{1}{2}\left(j_nj_0-\overline{j}_{-n}\overline{j}_0+\sum_{k\ge 1}\left(j_{n-k}j_k-\overline{j}_{-n-k}\overline{j}_k+j_{-k}j_{n+k}-\overline{j}_{-k}\overline{j}_{-n+k}\right)\right)\left|B_{\rm N,D}\right\rangle.
\end{aligned}$$

Note that here we changed the summation index $k\to -k$ in the second sum. Next, we recall (2.20) and $j_0=\overline{j}_0$ to observe that the terms involving j_0 and $\overline{j}_0$ vanish when applied to $|B_{\rm N,D}\rangle$. The remaining terms can be rewritten as

$$\begin{aligned}
&\frac{1}{2}\sum_{k\ge 1}\Big(j_{n-k}\big(j_k\pm\overline{j}_{-k}\big)\mp j_{n-k}\overline{j}_{-k}\quad\mp\overline{j}_{-n-k}\big(j_{-k}\pm\overline{j}_k\big)\pm\overline{j}_{-n-k}j_{-k}\\
&+j_{-k}\big(j_{n+k}\pm\overline{j}_{-n-k}\big)\mp j_{-k}\overline{j}_{-n-k}\quad\mp\overline{j}_{-k}\big(j_{n-k}\pm\overline{j}_{-n+k}\big)\pm\overline{j}_{-k}j_{n-k}\Big)\left|B_{\rm N,D}\right\rangle.
\end{aligned}$$

By again employing the boundary conditions (2.20), we see that half of these terms vanish when acting on the boundary state while the other half cancels among themselves. In summary, we have shown that

$$\left(L_n-\overline{L}_{-n}\right)\,\left|B_{\rm N,D}\right\rangle=0\ .$$

2.2.2 Tree-Level Amplitudes

Cylinder Diagram in General

We now turn to the cylinder diagram which we compute in the closed sector. Referring again to Fig. 2.4, in string theory we can interpret this diagram as a closed string which is emitted at the boundary A, propagating via the closed sector Hamiltonian $H_{\rm closed}=L_0+\overline{L}_0-\frac{c+\overline{c}}{24}$ for a time $\tau=l$ until it reaches the boundary B where it gets absorbed. In analogy to Quantum Mechanics, such an amplitude is given by the overlap

$$\tilde{\mathscr{Z}}^{\mathscr{C}}(l)=\langle\Theta B|\ e^{-2\pi l\left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\ |B\rangle\ , \tag{2.26}$$

where the tilde indicates that the computation is performed in the closed sector (or at tree-level) and l is the length of the cylinder connecting the two boundaries.

Let us now explain the notation $\langle\Theta B|$. This bra-vector is understood as the hermitian conjugate of the ket-vector $|\Theta B\rangle$. Furthermore, we have introduced the CPT operator Θ which acts as charge conjugation (C), parity transformation (P) $\sigma\mapsto -\sigma$ and time reversal (T) $\tau\mapsto -\tau$ for the two-dimensional CFT. The reason for considering this operator can roughly be explained by the fact that the orientation of the boundary a closed string is emitted at is opposite to the orientation of the boundary where the closed string gets absorbed. For the momentum dependence of a boundary state $|B, \pi_0\rangle$, this implies in particular that

$$\langle\, \pi_0^a \mid \pi_0^b \,\rangle = \delta\big(\pi_0^a + \pi_0^b\big). \tag{2.27}$$

Without a detailed derivation, we finally note that the theory of the free boson is CPT invariant and so the action of Θ on the boundary states (2.21) of the free boson theory (and on ordinary numbers $c \in \mathbb{C}$) reads

$$\Theta\ \big|B, \pi_0\big\rangle = \frac{1}{\mathcal{N}^*}\ \big|B, \pi_0\big\rangle, \qquad \Theta\ c\ \Theta^{-1} = c^*, \tag{2.28}$$

where * denotes complex conjugation.

Free Boson II: Cylinder Diagram (Tree-Channel)

Let us now be more concrete and compute the overlap of two boundary states (2.26) for the example of the free boson. To do so, we note that for the free boson CFT we have $c = \overline{c} = 1$ and that

$$L_0 = \frac{1}{2}\, j_0\, j_0 + \sum_{k\geq 1} j_{-k}\, j_k,$$

and similarly for $\overline{L}_0$. Next, we perform the following calculation in order to evaluate (2.26). In particular, we use $j_{-k} j_k\, j_{-k}^{m_k}|0\rangle = m_k\ k\, j_{-k}^{m_k}|0\rangle$ to find

$$\begin{aligned} q^{\,\sum_{k\geq 1} j_{-k} j_k}\ \big|\mathbf{m}\big\rangle &= \prod_{k=1}^{\infty} \sum_{p=0}^{\infty} \frac{\big(-2\pi i \tau\big)^p}{p!}\ \big(j_{-k} j_k\big)^p \prod_{l=1}^{\infty} \frac{1}{\sqrt{m_l!}} \left(\frac{j_{-l}}{\sqrt{l}}\right)^{m_l} |0\rangle \\ &= \prod_{k=1}^{\infty} \sum_{p=0}^{\infty} \frac{\big(-2\pi i \tau\big)^p}{p!}\ \big(m_k\ k\big)^p \prod_{l=1}^{\infty} \frac{1}{\sqrt{m_l!}} \left(\frac{j_{-l}}{\sqrt{l}}\right)^{m_l} |0\rangle \\ &= \prod_{k=1}^{\infty} q^{\,m_k\, k}\ \big|\mathbf{m}\big\rangle. \end{aligned} \tag{2.29}$$

The cylinder diagram for the three possible combinations of boundary conditions is then computed as follows.

- For the case of Neumann–Neumann boundary conditions, we have $j_0|B_{\rm N}\rangle = \bar{j}_0|B_{\rm N}\rangle = 0$ and so the momentum contribution vanishes. For the remaining part, we calculate using (2.29) and (2.24)

$$
\begin{aligned}
\widetilde{\mathscr{Z}}^{\mathscr{C}\,(\mathrm{N,N})}_{\mathrm{bos.}}(l) &= \frac{e^{-2\pi l\left(-\frac{2}{24}\right)}}{\mathscr{N}_{\mathrm{N}}^2} \sum_{\mathbf{m}} \langle \mathbf{m}|\; e^{-2\pi l \sum_{k\geq 1} j_{-k} j_k}\; |\mathbf{m}\rangle \times \\
&\qquad \times \left\langle U\overline{\mathbf{m}}\right|\; e^{-2\pi l \sum_{k\geq 1} \bar{j}_{-k}\bar{j}_k}\; \left|U\overline{\mathbf{m}}\right\rangle \\
&= \frac{e^{-2\pi l\left(-\frac{2}{24}\right)}}{\mathscr{N}_{\mathrm{N}}^2} \sum_{\mathbf{m}} \prod_{k=1}^{\infty} e^{-2\pi l\, m_k k}\, (-1)^{\sum_{l=1}^{\infty} m_l}\, e^{-2\pi l\, m_k k}\, (-1)^{\sum_{l=1}^{\infty} m_l} \\
&= \frac{e^{\frac{\pi l}{6}}}{\mathscr{N}_{\mathrm{N}}^2} \prod_{k=1}^{\infty} \sum_{m_k=0}^{\infty} \left(e^{-4\pi l\, k}\right)^{m_k} = \frac{1}{\mathscr{N}_{\mathrm{N}}^2}\, e^{\frac{\pi l}{6}} \prod_{k=1}^{\infty} \frac{1}{1-e^{-4\pi l\, k}}
\end{aligned}
$$

where in the last step we performed a summation of the geometric series. Let us emphasise that due to the action of the CPT operator Θ shown in (2.28), $\mathscr{N}^2$ is just the square of $\mathscr{N}$ and not the absolute value squared. Then, with $q = e^{2\pi i\tau}$ and $\tau = 2il$ we find that the cylinder diagram for Neumann–Neumann boundary conditions is expressed as

$$
\widetilde{\mathscr{Z}}^{\mathscr{C}\,(\mathrm{N,N})}_{\mathrm{bos.}}(l) = \frac{1}{\mathscr{N}_{\mathrm{N}}^2}\, \frac{1}{\eta(2il)}. \tag{2.30}
$$

- Next, we consider the case of Dirichlet–Dirichlet boundary conditions. Noting that U now acts trivially on the basis states, we see that apart from the momentum contribution the calculation is similar to the case with Neumann–Neumann conditions. However, for the momentum dependence we compute using (2.27) and (2.28)

$$
\begin{aligned}
&\int_{-\infty}^{\infty} d\pi_0^a\, d\pi_0^b\, e^{+ix_0^a\pi_0^a}\, e^{+ix_0^b\pi_0^b}\, \left\langle \pi_0^a \right| e^{-2\pi l (j_0)^2} \left|\pi_0^b\right\rangle \\
&= \int_{-\infty}^{\infty} d\pi_0^a\, d\pi_0^b\, e^{+ix_0^a\pi_0^a}\, e^{+ix_0^b\pi_0^b}\, e^{-2\pi l\left(\pi_0^b\right)^2}\, \delta\left(\pi_0^a+\pi_0^b\right) \\
&= \int_{-\infty}^{\infty} d\pi_0^a\, e^{-2\pi l\left(\pi_0^a + i\,\frac{x_0^b-x_0^a}{4\pi l}\right)^2}\, e^{-\frac{\left(x_0^b-x_0^a\right)^2}{8\pi l}} = \frac{1}{\sqrt{2l}}\, e^{-\frac{\left(x_0^b-x_0^a\right)^2}{8\pi l}}
\end{aligned}
$$

where we completed a perfect square and performed the Gaussian integration. In order to arrive at the result above, we also employed that in the closed sector $\pi_0 = j_0 = \bar{j}_0$. The cylinder diagram with Dirichlet–Dirichlet boundary conditions therefore reads

$$
\widetilde{\mathscr{Z}}^{\mathscr{C}\,(\mathrm{D,D})}_{\mathrm{bos.}}(l) = \frac{1}{\mathscr{N}_{\mathrm{D}}^2}\, \exp\left(-\,\frac{\left(x_0^b-x_0^a\right)^2}{8\pi l}\right) \frac{1}{\sqrt{2l}}\, \frac{1}{\eta(2il)}.
$$

- Finally, for mixed Neumann–Dirichlet conditions, the boundary state satisfies $j_0|B_{\mathrm{D}}\rangle = \bar{j}_0|B_{\mathrm{D}}\rangle = \pi_0|B_{\mathrm{D}}\rangle$ which leads us to

$$
\int d\pi_0\, e^{i\pi_0 x_0} \langle \pi_0 = 0|\; e^{-2\pi l\, j_0^2}\; |\pi_0\rangle = \int d\pi_0\, e^{i\pi_0 x_0} e^{-2\pi l\, \pi_0^2}\, \delta(\pi_0) = 1.
$$

In the anti-holomorphic sector of the Dirichlet boundary state, the action of U on the basis states $|\overline{\mathbf{m}}\rangle$ is trivial and so we obtain a single factor of $(-1)^{\sum_k m_k}$. For the full cylinder diagram, this implies

$$\tilde{\mathscr{Z}}^{\mathscr{C}\,(\text{mixed})}_{\text{bos.}}(l) = \frac{e^{\frac{\pi l}{6}}}{\mathscr{N}_{\text{N}}\mathscr{N}_{\text{D}}} \prod_{k=1}^{\infty} \sum_{m_k=0}^{\infty} \left(-e^{-4\pi l\, k}\right)^{m_k} = \frac{e^{\frac{\pi l}{6}}}{\mathscr{N}_{\text{N}}\mathscr{N}_{\text{D}}} \prod_{k=1}^{\infty} \frac{1}{1+e^{-4\pi l\, k}}.$$

One defines the $\vartheta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}(\tau, z)$-functions as

$$\vartheta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}(\tau, z) = \sum_{n\in\mathbb{Z}} q^{\frac{1}{2}(n+\alpha)^2}\, e^{2\pi i(n+\alpha)(z+\beta)},$$

which can be shown to also have a representation as an infinite product

$$\frac{\vartheta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}(\tau, z)}{\eta(\tau)} = e^{2\pi i\alpha(z+\beta)}\, q^{\frac{\alpha^2}{2}-\frac{1}{24}} \prod_{n=1}^{\infty}\left(1+q^{n+\alpha-\frac{1}{2}}\, e^{2\pi i(z+\beta)}\right) \times \left(1+q^{n-\alpha-\frac{1}{2}}\, e^{-2\pi i(z+\beta)}\right).$$

In particular one can write

$$\vartheta_2(\tau) \equiv \vartheta\begin{bmatrix}1/2\\ 0\end{bmatrix}(\tau, 0) = \sum_{n=-\infty}^{\infty} q^{(n+\frac{1}{2})^2} \tag{2.31}$$

so that we can express the cylinder diagram for mixed boundary conditions as

$$\tilde{\mathscr{Z}}^{\mathscr{C}\,(\text{mixed})}_{\text{bos.}}(l) = \frac{\sqrt{2}}{\mathscr{N}_{\text{N}}\mathscr{N}_{\text{D}}} \sqrt{\frac{\eta\,(2il)}{\vartheta_2(2il)}}.$$

Loop-Channel—Tree-Channel Equivalence

Let us come back to Fig. 2.4. As it is illustrated there and motivated at the beginning of this section, we expect the cylinder diagram in the closed and open sector to be related. More specifically, this relation is established by $(\sigma, \tau)_{\text{open}} \leftrightarrow (\tau, \sigma)_{\text{closed}}$ where σ is the world-sheet space coordinate and τ is world-sheet time. However, this mapping does not change the cylinder, in particular, it does not change the modular parameter τ. In the open sector, the cylinder has length $\frac{1}{2}$ and circumference t when measured in units of 2π, while in the closed sector we have length l and circumference 1. Then, the modular parameter in the open and closed sector are

$$\tau_{\text{open}} = \frac{\alpha_2}{\alpha_1} = \frac{it}{1/2} = 2it, \qquad \tau_{\text{closed}} = \frac{\alpha_2}{\alpha_1} = \frac{i}{l}.$$

As we have emphasised, the modular parameters in the open and closed sector have to be equal which leads us to the relation

$$\boxed{t = \frac{1}{2l}}\,.$$

This is the formal expression for the pictorial *loop-channel–tree-channel equivalence* of the cylinder diagram illustrated in Fig. 2.4.

We now verify this relation for the example of the free boson explicitly which will allow us to fix the normalisation constants $\mathcal{N}_\mathrm{D}$ and $\mathcal{N}_\mathrm{N}$ of the boundary states. Recalling the cylinder partition function (2.16) in the open sector, we compute

$$\mathcal{Z}^{\mathcal{C}(\mathrm{N,N})}_{\mathrm{bos.}}\big(t\big) = \frac{1}{2\sqrt{t}}\,\frac{1}{\eta(it)} \quad \xrightarrow{t=\frac{1}{2l}} \quad \sqrt{\frac{l}{2}}\,\frac{1}{\eta\left(-\frac{1}{2il}\right)} = \frac{1}{2\;\eta(2il)} = \frac{\mathcal{N}_\mathrm{N}^2}{2}\;\tilde{\mathcal{Z}}^{\mathcal{C}(\mathrm{N,N})}_{\mathrm{bos.}}\big(l\big),$$

where we used the modular properties of the Dedekind η function

$$\eta\left(-\tau^{-1}\right) = \sqrt{-i\tau}\,\eta(\tau). \tag{2.32}$$

Therefore, requiring the results in the loop- and tree-channel to be related, we can fix

$$\boxed{\mathcal{N}_\mathrm{N} = \sqrt{2}}\,. \tag{2.33}$$

Next, for Dirichlet–Dirichlet boundary conditions, we find

$$\begin{aligned}\mathcal{Z}^{\mathcal{C}(\mathrm{D,D})}_{\mathrm{bos.}}\big(t\big) &= \exp\left(-\frac{t}{4\pi}\left(x_0^b - x_0^a\right)^2\right)\frac{1}{\eta\;(it)}\\ &\xrightarrow{t=\frac{1}{2l}} \;\exp\left(-\frac{1}{8\pi l}\left(x_0^b - x_0^a\right)^2\right)\frac{1}{\eta\left(-\frac{1}{2il}\right)} = \mathcal{N}_\mathrm{D}^2\;\tilde{\mathcal{Z}}^{\mathcal{C}(\mathrm{D,D})}_{\mathrm{bos.}}\big(l\big),\end{aligned}$$

which allows us to fix the normalisation constant as

$$\boxed{\mathcal{N}_\mathrm{D} = 1}\,.$$

Finally, the loop-channel–tree-channel equivalence for mixed Neumann–Dirichlet boundary conditions can be verified along similar lines. This discussion shows that indeed the cylinder partition function for the free boson in the open and closed sector are related via a modular transformation, more concretely via a modular *S*-transformation.

Summary and Remark

Let us now briefly summarise our findings of this section and close with some remarks.

- By performing the so-called world-sheet duality $(\sigma,\tau)_{\text{open}} \leftrightarrow (\tau,\sigma)_{\text{closed}}$, we translated the Neumann and Dirichlet boundary conditions from the open sector to the closed sector. In string theory, the boundary in the closed sector is interpreted as an object which absorbs or emits closed strings.
- Working out the boundary conditions in terms of the Laurent modes of the free boson theory, we obtained the gluing conditions

$$\left(j_n \pm \bar{j}_{-n} \right) \left| B_{\text{N,D}} \right\rangle = 0$$

  which imply that the two $U(1)$ symmetries generated by $j(z)$ and $\bar{j}(\bar{z})$ are broken to a diagonal $U(1)$.
- For the example of the free boson theory, we stated the solution $|B\rangle$ to the gluing conditions and verified them. Along the way, we also outlined the idea for constructing boundary states for more general theories.
- The cylinder amplitude in the closed sector (tree-level) is computed from the overlap of two boundary states

$$\tilde{\mathscr{Z}}^{\mathscr{C}}(l) = \langle \Theta B|\; e^{-2\pi l \left(L_0 + \overline{L}_0 - \frac{c+\bar{c}}{24} \right)}\; |B\rangle \, .$$

  We performed this calculation for the free boson and checked that it is related to the cylinder partition function in the open sector via world-sheet duality. In particular, this transformation is a modular S-transformation.
- Finally, the BCFT also has to preserve the conformal symmetry generated by $T(z)$. The boundary states respect this symmetry in the sense that the following conditions have to be satisfied

$$\left(L_n - \overline{L}_{-n} \right) \left| B_{\text{N,D}} \right\rangle = 0 \, ,$$

  which we checked for the example of the free boson theory.
- Very similarly, one can generalise the concept of boundaries and boundary states to the CFT of a free fermion which is very important for applications in Superstring Theory.

  As we mentioned already, in string theory boundary states are called D-branes to emphasise the space-time point of view of such objects. They are higher dimensional generalisations of strings and membranes, and indeed they play a very important role in understanding the non-perturbative sector of string theory. It was one of the big insights at the end of the last millennium that such higher dimensional objects are naturally contained in string theory (which started as a theory of only one-dimensional objects) and gave rise to various surprising dualities, the most famous surely being the celebrated AdS/CFT correspondence.

2.3 Boundary States for RCFTs

After having studied the Boundary CFT of the free boson in great detail, let us now generalise our findings to theories without a Lagrangian description. In particular, we focus on RCFTs and we will formulate the corresponding Boundary RCFT just in terms of gluing conditions for the theory on the sphere.

Boundary Conditions

We consider Rational Conformal Field Theories with chiral and anti-chiral symmetry algebras $\mathscr{A}$ respectively $\overline{\mathscr{A}}$. For the theory on the sphere the Hilbert space splits into irreducible representations of $\mathscr{A} \otimes \overline{\mathscr{A}}$ as

$$\mathscr{H} = \bigoplus_{i,\bar{j}} M_{i\bar{j}} \; \mathscr{H}_i \otimes \overline{\mathscr{H}}_{\bar{j}}$$

where $M_{i\bar{j}}$ are the same multiplicities of the highest weight representation appearing in the modular invariant torus partition function. Note that for the case of RCFTs we are considering, there is only a finite number of irreducible representations and that the modular invariant torus partition function is given by a combination of chiral and anti-chiral characters as follows

$$\mathscr{Z}(\tau, \overline{\tau}) = \sum_{i,\bar{j}} M_{i\bar{j}} \; \chi_i(\tau) \; \overline{\chi}_{\bar{j}}(\overline{\tau}).$$

Generalising the results from the free boson theory, we state without derivation that a boundary state $|B\rangle$ in the RCFT preserving the symmetry algebra $\mathscr{A} = \overline{\mathscr{A}}$ has to satisfy the following gluing conditions

$$\boxed{\begin{aligned} \left(L_n - \overline{L}_{-n} \right) \big|B\big\rangle = 0 \quad & \text{conformal symmetry}, \\ \left(W_n^i - (-1)^{h^i} \; \overline{W}_{-n}^i \right) \big|B\big\rangle = 0 \quad & \text{extended symmetries}, \end{aligned}} \tag{2.34}$$

where W_n^i is the holomorphic Laurent mode of the extended symmetry generator W^i with conformal weight $h^i = h(W^i)$, and $\overline{W}^i$ denotes the generator in the anti-holomorphic sector. However, the condition for the extended symmetries can be relaxed, so that also Dirichlet boundary conditions similar to the example of a free boson are included

$$\left(W_n^i - (-1)^{h^i} \; \Omega\big(\overline{W}_{-n}^i\big)\right) \big|B\big\rangle = 0,$$

where $\Omega : \mathscr{A} \rightarrow \mathscr{A}$ is an automorphism of the chiral algebra $\mathscr{A}$. Such an automorphism Ω is also called a *gluing automorphism* and for our example of the free boson with Dirichlet boundary conditions, it simply is $\Omega : \overline{W}_n \mapsto -\overline{W}_n$.

Ishibashi States

Let us introduce the charge conjugation matrix C which maps highest weight representations i to their charge conjugate i^+. Denoting then the Hilbert space built upon the charge conjugate represention by $\mathcal{H}_i^+$, we can state the important result of Ishibashi:

> For $\overline{\mathcal{A}} = \mathcal{A}$ and $\overline{\mathcal{H}}_i = \mathcal{H}_i^+$, to each highest weight representation ϕ_i of $\mathcal{A}$ one can associate an up to a constant unique state $|\mathcal{B}_i\rangle\rangle$ such that the gluing conditions are satisfied.

Note that since the CFTs we are considering are rational, there is only a finite number of highest weight states and thus only a finite number of such so-called Ishibashi states $|\mathcal{B}_i\rangle\rangle$.

We now construct the Ishibashi states in analogy to the boundary states of the free boson. Denoting by $|\phi_i, \mathbf{m}\rangle$ an orthonormal basis for $\mathcal{H}_i$, the Ishibashi states are written as

$$\big|\mathcal{B}_i\rangle\rangle = \sum_{\mathbf{m}} \big|\phi_i, \mathbf{m}\rangle \otimes U|\overline{\phi}_i, \overline{\mathbf{m}}\rangle, \tag{2.35}$$

where $U : \overline{\mathcal{H}} \to \overline{\mathcal{H}}^+$ is an anti-unitary operator acting on the symmetry generators $\overline{W}^i$ as follows

$$U\, \overline{W}^i_n\, U^{-1} = (-1)^{h^i}\, \big(\overline{W}^i_{-n}\big)^\dagger.$$

The proof that the Ishibashi states are solutions to the gluing conditions (2.34) is completely analogous to the example of the free boson and so we will not present it here.

The Cardy Condition

For later purpose, let us now compute the following overlap of two Ishibashi states

$$\langle\langle\mathcal{B}_j\big|\, e^{-2\pi l\, \left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\, \big|\mathcal{B}_i\rangle\rangle. \tag{2.36}$$

Utilising the gluing conditions for the conformal symmetry generator (2.34), we see that we can replace $\overline{L}_0$ by L_0 and $\overline{c}$ by c. Next, because the Hilbert spaces of two different HWRs ϕ_i and ϕ_j are independent of each other, the overlap above is only nonzero for $i = j^+$. Note that here we have written the charge conjugate j^+ of the highest weight ϕ_j because the hermitian conjugation also acts as charge conjugation. We then obtain

$$\begin{aligned}\langle\langle\mathcal{B}_j\big|\, e^{-2\pi l\, \left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\, \big|\mathcal{B}_i\rangle\rangle &= \delta_{ij^+}\, \langle\langle\mathcal{B}_i\big|\, e^{2\pi i\,(2il)\, \left(L_0-\frac{c}{24}\right)}\, \big|\mathcal{B}_i\rangle\rangle \\ &= \delta_{ij^+}\, \mathrm{Tr}_{\mathcal{H}_i}\Big(q^{L_0-\frac{c}{24}}\Big) \\ &= \delta_{ij^+}\, \chi_i\big(2il\big)\end{aligned} \tag{2.37}$$

with χ_i the character defined as

$$\chi_i(\tau) := Tr_{\mathscr{H}_i}\left(q^{L_0-\frac{c}{24}}\right) \tag{2.38}$$

over the Hilbert space $\mathscr{H}_i$ built on the highest weigth state ϕ_i. Performing a modular S-transformation for this overlap, by the same reasoning as for the free boson, we expect to obtain a partition function in the boundary sector. However, because the S-transform of a character $\chi_i(2il)$ in general does not give non-negative integer coefficients in the loop-channel, it is not clear whether to interpret such a quantity as a partition function counting states of a given excitation level.

As it turns out, the Ishibashi states are not the boundary states itself but only building blocks guaranteed to satisfy the gluing conditions. A true boundary state in general can be expressed as a linear combination of Ishibashi states in the following way

$$\big|B_\alpha\big\rangle = \sum_i B^i_\alpha \, \big|\mathscr{B}_i\big\rangle\!\big\rangle. \tag{2.39}$$

The complex coefficients B^i_α in (2.39) are called reflection coefficients and are very constrained by the so-called Cardy condition. This condition essentially ensures the loop-channel–tree-channel equivalence. Indeed, using relation (2.37) and choosing normalisations such that the action of the CPT operator Θ introduced in (2.28) reads

$$\Theta\,\big|B_\alpha\big\rangle = \sum_i \big(B^i_\alpha\big)^* \,\big|\mathscr{B}_{i^+}\big\rangle\!\big\rangle, \tag{2.40}$$

the cylinder amplitude between two boundary states of the form (2.39) can be expressed as follows

$$\begin{aligned}\widetilde{\mathscr{Z}}_{\alpha\beta}(l) &= \langle\, \Theta B_\alpha |\, e^{-2\pi l\,\left(L_0+\overline{L}_0-\frac{c+\bar c}{24}\right)}\, |B_\beta\rangle \\ &= \sum_{i,j} B^j_\alpha\, B^i_\beta\, \big\langle\!\big\langle \mathscr{B}_{j^+}\big|\, e^{-2\pi l\,\left(L_0+\overline{L}_0-\frac{c+\bar c}{24}\right)}\, \big|\mathscr{B}_i\big\rangle\!\big\rangle \\ &= \sum_i B^i_\alpha\, B^i_\beta\, \chi_i\big(2il\big).\end{aligned}$$

Performing a modular S-transformation $l \mapsto \frac{1}{2t}$ on the characters χ_i, this closed sector cylinder diagram is transformed to the following expression in the open sector

$$\widetilde{\mathscr{Z}}_{\alpha\beta}\big(l\big) \to \widetilde{\mathscr{Z}}_{\alpha\beta}\big(\frac{1}{2t}\big) = \sum_{i,j} B^i_\alpha\, B^i_\beta\, S_{ij}\, \chi_j\big(it\big) = \sum_j n^j_{\alpha\beta}\, \chi_j\big(it\big) = \mathscr{Z}_{\alpha\beta}(t),$$

where S_{ij} is the modular S-matrix and where we introduced the new coefficients $n^i_{\alpha\beta}$. Now, the Cardy condition is the requirement that this expression can be interpreted as a partition function in the open sector. That is, for all pairs of boundary states $|B_\alpha\rangle$ and $|B_\beta\rangle$ in a RCFT the following combinations have to be non-negative integers

$$\boxed{n^j_{\alpha\beta} = \sum_i B^i_\alpha\, B^i_\beta\, S_{ij} \quad \in \quad \mathbb{Z}^+_0}\,.$$

Construction of Boundary States

The Cardy condition just illustrated is very reminiscent of the Verlinde formula, where a similar combination of complex numbers leads to non-negative fusion rule coefficients. For the case of a *charge conjugate* modular invariant partition function, that is when the characters $\chi_i(\tau)$ are combined with $\overline{\chi}_{i^+}(\overline{\tau})$ as $\mathscr{Z} = \sum_i \chi_i(\tau)\ \overline{\chi}_{i^+}(\overline{\tau})$, we can construct a generic solution to the Cardy condition by choosing the reflection coefficients in the following way

$$\boxed{B^i_\alpha = \frac{S_{\alpha\, i}}{\sqrt{S_{0\, i}}}}\,.$$

Note, for each highest weight representation ϕ_i in the RCFT, there not only exists an Ishibashi state but also a boundary state, i.e. the index α in $|B_\alpha\rangle$ also runs from one to the number of HWRs. Employing then the Verlinde formula

$$N^k_{ij} = \sum_n \frac{S_{in} S_{jn} S^{kn}}{S_{0n}} \tag{2.41}$$

and denoting the non-negative, integer fusion coefficients by $N^\alpha_{j\beta}$, we find that the Cardy condition for the coefficients $n^j_{\alpha\beta}$ is always satisfied

$$n^j_{\alpha\beta} = \sum_i \frac{S_{\alpha i}\ S_{\beta i}\ S_{ij}}{S_{0i}} = \sum_i \frac{S_{\alpha i}\ S_{\beta i}\ S^*_{ij+}}{S_{0i}} = N^{j+}_{\alpha\beta} \quad \in \quad \mathbb{Z}^+_0\,.$$

Note that here we employed $S^*_{ij} = S_{ij^+}$ which is verified by noting that $S^{-1} = S^*$ as well as that $S^2 = C$ with C the charge conjugation matrix $C_{ij} = \delta_{ij^+}$.

2.4 CFTs on Non-orientable Surfaces

Up to this point, we have studied Conformal Field Theories defined on the Riemann sphere respectively the complex plane, and on the torus. For Boundary CFTs, the corresponding surfaces are the upper half-plane and the cylinder. We note that all these surfaces are orientable, that is an orientation can be chosen globally.

However, in string theory it is necessary to also define CFTs on non-orientable surfaces. One such surface is the so-called crosscap $\mathbb{RP}^2$ which can be viewed as the two-sphere with opposite points identified. Other non-orientable surfaces are the Möbius strip and the Klein bottle, and a summary of all surfaces relevant for the following is shown in Fig. 2.5.

Orientifolds

Before formulating CFTs on non-orientable surfaces, let us briefly explain the string theory origin of such theories. Recalling the action for a free boson (2.1), we observe that this theory has a discrete symmetry denoted as Ω which takes the form

$$\Omega \ : \ X\big(\tau,\sigma\big) \mapsto \widetilde{X}\big(\tau,\sigma\big) = X\big(\tau,-\sigma\big), \tag{2.42}$$

with τ and σ again world-sheet time and space coordinates. To see that the action (2.1) is invariant under Ω, observe that

$$\begin{aligned} \Omega\ \big(\partial_\sigma X\big)(\tau,\sigma)\ \Omega^{-1} &= -\big(\partial_\sigma X\big)(\tau,-\sigma), \\ \Omega\ \big(\partial_\tau X\big)(\tau,\sigma)\ \Omega^{-1} &= +\big(\partial_\tau X\big)(\tau,-\sigma). \end{aligned} \tag{2.43}$$

Next, let us note that from the mapping (2.42), we see that Ω acts as a world-sheet parity operator. In the string theory picture, this means that Ω changes the orientation of a closed string. As with any other symmetry, we can study the quotient of the original theory by the symmetry. Since Ω changes orientation, in analogy to orbifolds, such a quotient is called an *orientifold*.

The Example of the Free Boson in More Detail

Let us further elaborate on the action of the orientifold projection Ω for the free boson. We first note that $-\sigma$ has to be interpreted properly because we normalised the world-sheet space coordinate as $\sigma \in [0,2\pi)$ for the closed sector and as $\sigma \in [0,\pi]$ in the open sector. The correct identification for $-\sigma$ then reads

$$-\sigma_{\text{closed}} \sim 2\pi - \sigma_{\text{closed}}, \qquad -\sigma_{\text{open}} \sim \pi - \sigma_{\text{open}}.$$

Next, we consider the free boson in the closed sector and express $\partial_\sigma X$ in (2.43) in terms of the Laurent modes j_n and $\bar{j}_n$ using (2.5)

$$\begin{aligned} \Omega\ \left(\partial_\sigma X\right)(\tau,\sigma)\ \Omega^{-1} &= -\left(\partial_\sigma X\right)(\tau,-\sigma) \\ &\sum_{n\in\mathbb{Z}}\Big(\ \Omega\ j_n\ \Omega^{-1}\ e^{-n(\tau+i\sigma)} - \Omega\ \bar{j}_n\ \Omega^{-1}\ e^{-n(\tau-i\sigma)}\Big) \\ &= \sum_{n\in\mathbb{Z}}\Big(-j_n\ e^{-n(\tau+i(2\pi-\sigma))} + \bar{j}_n\ e^{-n(\tau-i(2\pi-\sigma))}\Big)\,. \end{aligned} \tag{2.44}$$

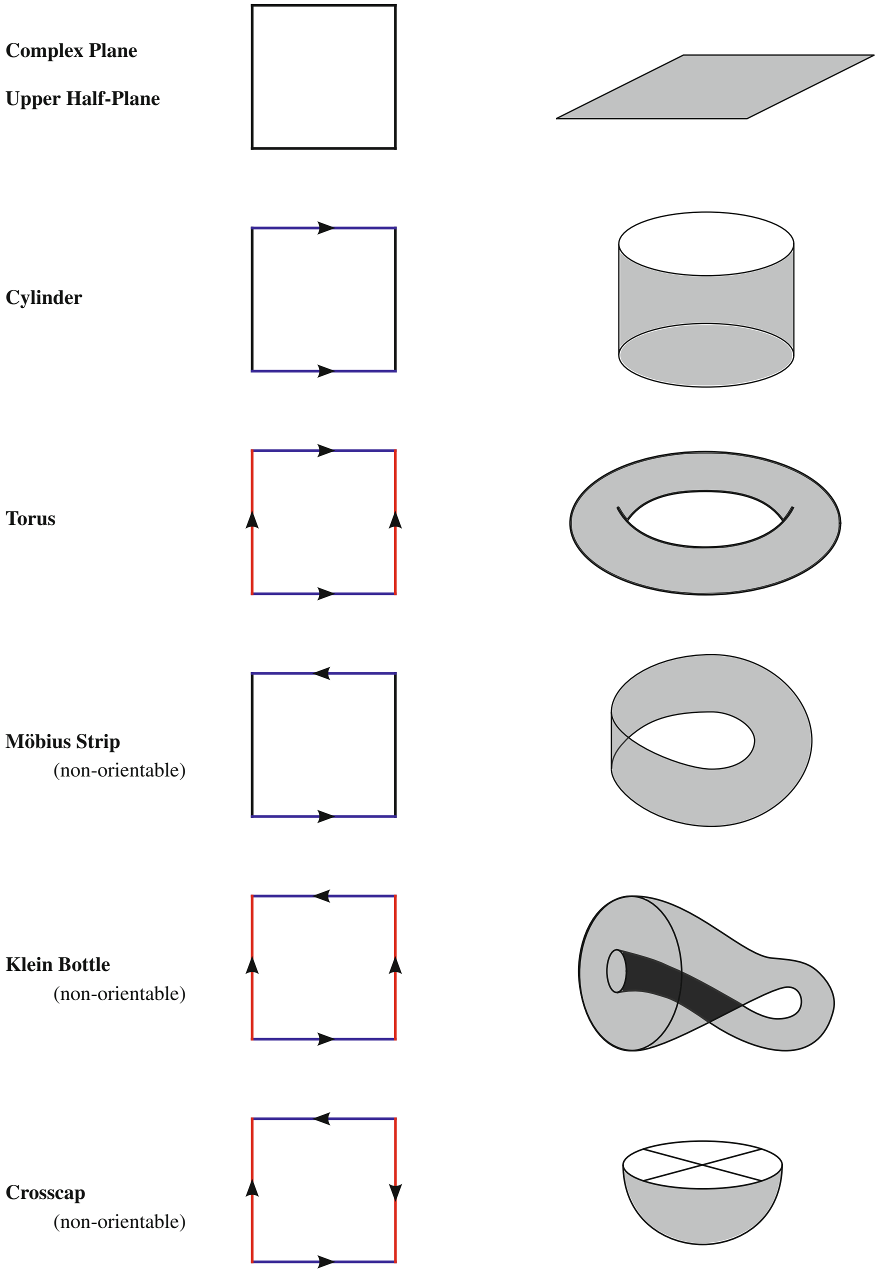


Fig. 2.5 Two-dimensional orientable and non-orientable surfaces. On the *left-hand* side, the fundamental domain can be found and it is indicated how opposite edges are identified leading to the surfaces illustrated on the *right-hand* side. Note that for the identification of opposite edges the orientation given by the arrows is crucial

From this relation we can determine the action of Ω on the modes in the closed sector as follows

$$\boxed{\Omega\, j_n\, \Omega^{-1} = \bar{j}_n, \quad \Omega\, \bar{j}_n\, \Omega^{-1} = j_n.} \tag{2.45}$$

For the open sector, we have to replace 2π on the right-hand side in (2.44) by π which leads to an additional factor of $(-1)^n$. Using then the boundary conditions of an open string (2.6) which relate the Laurent modes as $j_n = \pm\bar{j}_n$, we obtain the action of Ω in the open sector as

$$\boxed{\Omega\, j_n\, \Omega^{-1} = \pm(-1)^n\, j_n} \tag{2.46}$$

where the two signs correspond to Neumann–Neumann respectively Dirichlet–Dirichlet boundary conditions. For the case of mixed boundary conditions, we recall that the Laurent modes have labels $n \in \mathbb{Z} + \frac{1}{2}$ and we note that Ω interchanges the endpoints of an open string as well as the boundary conditions. In particular, we find

$$\Omega\, j_n^{(\mathrm{N,D})}\, \Omega^{-1} = -(-1)^n\, j_n^{(\mathrm{D,N})}, \qquad \Omega\, j_n^{(\mathrm{D,N})}\, \Omega^{-1} = +(-1)^n\, j_n^{(\mathrm{N,D})}. \tag{2.47}$$

Partition Function: Klein Bottle

Let us now consider partition functions for general orientifold theories. We start with the usual form of a modular invariant partition function in a CFT

$$\mathscr{Z}(\tau,\overline{\tau}) = \mathrm{Tr}_{\mathscr{H}\times\overline{\mathscr{H}}}\Big(q^{L_0-\frac{c}{24}}\ \overline{q}^{\overline{L}_0-\frac{c}{24}}\Big), \tag{2.48}$$

where we indicated the trace over the combined Hilbert space $\mathscr{H} \times \overline{\mathscr{H}}$ explicitly. Next, we generalise our findings from the example of the free boson and define the action of the world-sheet parity operator Ω on the Hilbert space as follows

$$\Omega \mid i, \bar{j}\, \rangle = \pm \mid \Omega(j), \overline{\Omega(i)}\, \rangle\ , \tag{2.49}$$

where i denotes a state in the holomorphic sector of the theory and $\bar{j}$ stands for the anti-holomorphic sector. The two different signs originate from the two possibilities of Ω acting on the vacuum $|0\rangle$ compatible with the requirement that $\Omega^2 = \mathbb{1}$. The simplest choice for $\Omega(i)$ is $\Omega(i) = i$, but also more general $\mathbb{Z}_2$ involutions are possible, for instance $\Omega(i) = i^+$ where $^+$ denotes charge conjugation.

In order to obtain the partition function we project the entire Hilbert space $\mathscr{H} \times \overline{\mathscr{H}}$ onto those states which are invariant under Ω, i.e. we introduce the projection operator $\frac{1}{2}(1+\Omega)$ into the partition function (2.48). We therefore obtain

$$\begin{aligned}\mathscr{Z}^{\Omega}(\tau,\overline{\tau}) &= \mathrm{Tr}_{\mathscr{H}\times\overline{\mathscr{H}}}\left(\frac{1+\Omega}{2}\ q^{L_0-\frac{c}{24}}\ \overline{q}^{\overline{L}_0-\frac{c}{24}}\right)\\ &= \frac{1}{2}\,\mathscr{Z}(\tau,\overline{\tau}) + \frac{1}{2}\,\mathrm{Tr}_{\mathscr{H}\times\overline{\mathscr{H}}}\Big(\ \Omega\ \ q^{L_0-\frac{c}{24}}\ \overline{q}^{\overline{L}_0-\frac{c}{24}}\Big).\end{aligned}$$

The first term is just one-half of the torus partition function which we already studied. Let us therefore turn to the second term

$$\mathscr{Z}^{\mathscr{K}}(\tau,\overline{\tau}) = \mathrm{Tr}_{\mathscr{H}\times\overline{\mathscr{H}}}\Big(\,\Omega\; q^{L_0-\frac{c}{24}}\;\overline{q}^{\overline{L}_0-\frac{\overline{c}}{24}}\Big). \tag{2.50}$$

The insertion of Ω into the trace has the effect that by looping once around the direction τ of a torus, the closed string comes back to itself up to the action of Ω, that is up to a change of orientation. Geometrically, such a diagram is not a torus but a Klein bottle illustrated in Fig. 2.5. This is also the reason for the superscript $\mathscr{K}$ of the partition function and for its name: the Klein bottle partition function.

We will now specify the action of Ω as $\Omega(i) = i$ and $\Omega\;|0\rangle = +|0\rangle$ in order to make (2.50) more explicit. For this choice we obtain

$$\big\langle\, i,\overline{j}\;\big|\;\Omega\;\big|\;i,\overline{j}\,\big\rangle = \big\langle\, i,\overline{j}\;\big|\;j,\overline{i}\,\big\rangle = \delta_{ij}, \tag{2.51}$$

where we used Eq. (2.49). Therefore, only left-right symmetric states $|i,\overline{i}\rangle$ contribute to the trace in (2.50) and we can simplify the partition function as follows

$$\begin{aligned}
\mathscr{Z}^{\mathscr{K}}(\tau,\overline{\tau}) &= \mathrm{Tr}_{\mathscr{H}\times\overline{\mathscr{H}}}\Big(\,\Omega\; q^{L_0-\frac{c}{24}}\;\overline{q}^{\overline{L}_0-\frac{\overline{c}}{24}}\Big)\\
&= \sum_{i,\overline{j}}\big\langle\, i,\overline{j}\;\big|\;\Omega\; q^{L_0-\frac{c}{24}}\;\Omega^{-1}\;\Omega\;\overline{q}^{\overline{L}_0-\frac{\overline{c}}{24}}\;\Omega^{-1}\;\Omega\;\big|\;i,\overline{j}\,\big\rangle\\
&= \sum_{i}\big\langle\, i,\overline{i}\;\big|\;\Omega\; q^{L_0-\frac{c}{24}}\;\Omega^{-1}\;\Omega\;\overline{q}^{\overline{L}_0-\frac{\overline{c}}{24}}\;\Omega^{-1}\;\big|\;i,\overline{i}\,\big\rangle
\end{aligned}$$

where we employed (2.51). Since only the diagonal subset will contribute to the trace, we see from this expression that effectively L_0 and $\overline{L}_0$ as well as c and $\overline{c}$ can be identified. Observing finally that $q\overline{q} = e^{-4\pi\tau_2}$, we arrive at

$$\mathscr{Z}^{\mathscr{K}}(\tau,\overline{\tau}) = \sum_{i}\big\langle\, i,\overline{i}\;\big|\;\big(q\,\overline{q}\big)^{L_0-\frac{c}{24}}\;\big|\;i,\overline{i}\,\big\rangle = \mathrm{Tr}_{\mathscr{H}_{\mathrm{sym.}}}\Big(\,e^{-4\pi t\left(L_0-\frac{c}{24}\right)}\,\Big), \tag{2.52}$$

with $t = \tau_2$ and $\mathscr{H}_{\mathrm{sym.}}$ denoting the states $\big|\;i,\overline{i}\;\big\rangle$ in the Hilbert space which are combined in a left-right symmetric way.

Free Boson III: Klein Bottle Partition Function (Loop-Channel)

Let us now determine the Klein bottle partition function for the example of the free boson. As it is evident from (2.52), this partition function is the character of the free boson theory with modular parameter $\tau = 2it$. However, for the momentum contribution, we need to perform a calculation similar to the one in the open sector shown on page 57. In particular, from (2.52) we extract the j_0 part, replace the sum by an integral and compute

$$\mathrm{Tr}_{\mathscr{H}_{\mathrm{sym.}}}\Big(\,e^{-4\pi t\,\frac{1}{2}\,j_0^2}\,\Big)\quad\longrightarrow\quad\int_{-\infty}^{+\infty} d\pi_0\; e^{-4\pi t\,\frac{1}{2}\,\pi_0^2} = \frac{1}{\sqrt{2t}}\,,$$

where we observed that in the closed sector $j_0 = \pi_0$. Combining this result with the character of the free boson theory, we obtain the following expression for the full Klein bottle partition function

$$\mathscr{Z}^{\mathscr{K}}_{\text{bos.}}(\tau, \overline{\tau}) = \frac{1}{\sqrt{2t}} \, \frac{1}{\eta\big(2it\big)} \; . \tag{2.53}$$

Partition Function: Möbius Strip

After having studied CFTs on non-orientable surfaces in the closed sector, let us now turn to the open sector. Again, the partition function has to be projected onto states invariant under the orientifold action Ω. Following the same steps as for the closed sector, we find

$$\mathscr{Z}^{\Omega}(t) = \text{Tr}_{\mathscr{H}_B}\left(\frac{1+\Omega}{2} \, e^{-2\pi t (L_0 - \frac{c}{24})} \right) = \frac{1}{2}\, \mathscr{Z}^{\mathscr{C}}(t) + \frac{1}{2}\, \text{Tr}_{\mathscr{H}_B}\Big(\Omega \, e^{-2\pi t (L_0 - \frac{c}{24})} \Big).$$

The first term is the cylinder amplitude, but the second term

$$\mathscr{Z}^{\mathscr{M}}(t) = \text{Tr}_{\mathscr{H}_\mathscr{B}}\Big(\Omega \, e^{-2\pi t (L_0 - \frac{c}{24})} \Big) \tag{2.54}$$

describes an open string whose orientation changes when looping along the t direction. The geometry of such a surface is that of a Möbius strip also shown in Fig. 2.5. The corresponding partition function is called the Möbius strip partition function and hence the superscript $\mathscr{M}$.

Free Boson IV: Möbius Strip Partition Function (Loop-Channel)

We now calculate the Möbius strip partition function for the free boson. The Hilbert space of the the free boson is spanned by states of the form

$$|n_1, n_2, n_3, \ldots\rangle = j_{-1}^{n_1} j_{-2}^{n_2} j_{-3}^{n_3} \ldots |0\rangle, \qquad \text{with}\, n_i \geq 0\,, \tag{2.55}$$

where j_n are the modes of the current $j(z)$. Recalling then the mapping (2.46), we see that the action of Ω on a state in the Hilbert space is

$$\Omega \, \big|n_1, n_2, n_3, \ldots\big\rangle = \prod_{k=1}^{\infty} (\pm 1)^{n_k} \, (-1)^{k\, n_k} \, \big|n_1, n_2, n_3, \ldots\big\rangle .$$

Taking the action of Ω into account we arrive at

$$\begin{aligned} \text{Tr}_{\mathscr{H}_\mathscr{B}}\Big(\Omega \, q^{L_0 - \frac{c}{24}}\Big)\Big|_{\text{without } j_0} &= q^{-\frac{1}{24}} \prod_{k=1}^{\infty} \sum_{n_k=0}^{\infty} (\pm 1)^{n_k} \, (-1)^{k\, n_k} \, q^{k\, n_k} \\ &= e^{\frac{\pi i}{24}} \, \big(-q\big)^{-\frac{1}{24}} \prod_{k=1}^{\infty} \frac{1}{1 \mp \big(-q\big)^k} \, . \end{aligned} \tag{2.56}$$

We also note that $-q$ with modular parameter τ can be expressed as $+q$ with modular parameter $\tau + \frac{1}{2}$.

For Neumann–Neumann boundary conditions, i.e. for the upper sign in the expression above, we employ the definition of the Dedekind η-function (2.14). However, since the momentum π_0 is unconstrained, we compute

$$\mathrm{Tr}_{\mathscr{H}_{\mathscr{B}}}\Big(\Omega\, e^{-2\pi t\, \frac{1}{2}\, j_0^2} \Big) \quad \longrightarrow \quad \int_{-\infty}^{+\infty} d\pi_0\, e^{-2\pi t\, \frac{1}{2}\, (2\pi_0)^2} = \frac{1}{2\sqrt{t}}\,,$$

where we used that j_0 is invariant under Ω as well as that in the open sector $j_0 = 2\pi_0$. The full Möbius strip partition function in the Neumann–Neumann sector then reads

$$\mathscr{Z}^{\mathscr{M}\,(\mathrm{N,N})}_{\mathrm{bos.}}(t) = e^{\frac{\pi i}{24}}\, \frac{1}{2\sqrt{t}}\, \frac{1}{\eta\big(\frac{1}{2} + it\big)}\,. \tag{2.57}$$

For Dirichlet–Dirichlet conditions, that means the lower sign in (2.56), we find for instance from (2.46) that $j_0 = 0$ so that there is no additional factor from the momentum integration. Recalling the expression for the ϑ_2-function from equation (2.31) we obtain

$$\mathscr{Z}^{\mathscr{M}\,(\mathrm{D,D})}_{\mathrm{bos.}}(t) = e^{\frac{\pi i}{24}}\, \sqrt{2}\, \sqrt{\frac{\eta\big(\frac{1}{2} + it\big)}{\vartheta_2\big(\frac{1}{2} + it\big)}}\,. \tag{2.58}$$

For mixed boundary conditions, the Möbius strip partition function vanishes as Ω exchanges Neumann–Dirichlet with Dirichlet–Neumann conditions and so there is no contribution to the trace.

Loop-Channel–Tree-Channel Equivalence

For the cylinder partition function, we have seen that the result in the open and closed sector are related via a modular S-transformation. One might therefore suspect that this equivalence between partition functions and overlaps of boundary states can also be found for non-orientable surfaces.

This is indeed the case which we illustrate in Fig. 2.6 for the Klein bottle partition function.

1. The fundamental domain of the Klein bottle shown in Fig. 2.6a is that of a torus up to a change of orientation. However, as opposed to the torus, the modular parameter of the Klein bottle is purely imaginary.
2. In Fig. 2.6b, the fundamental domain is halved and the identification of segments and points is indicated explicitly by arrows and symbols.
3. Next, we shift one half of the fundamental domain as shown in Fig. 2.6c.
4. In Fig. 2.6d, the shifted part has been flipped and the appropriate edges have been identified.

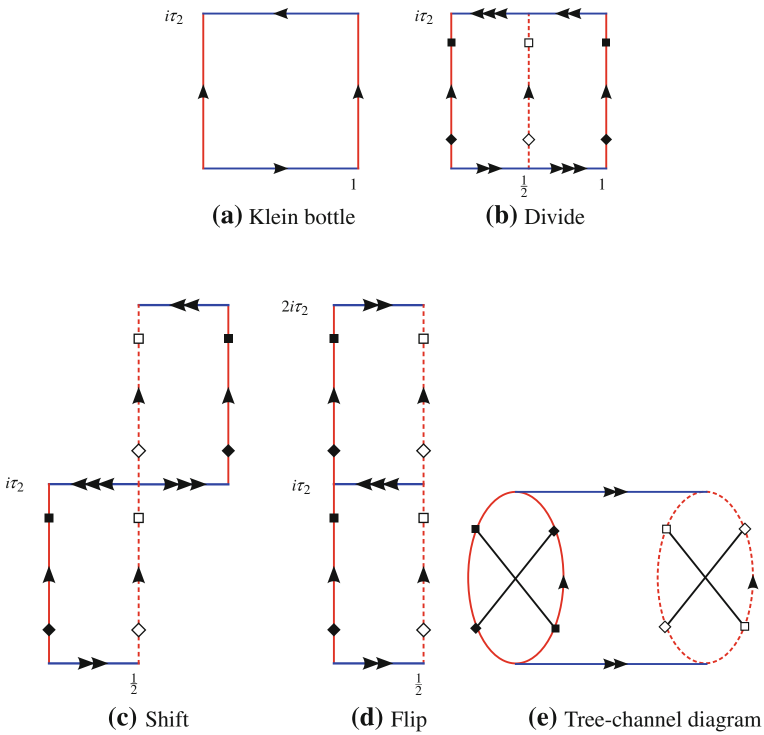


Fig. 2.6a–e Transformation of the fundamental domain of the Klein bottle to a tree-channel diagram between two crosscaps

5. A fundamental domain of this form can be interpreted as a cylinder between two crosscaps as illustrated in 2.6e.

Analogous to the cylinder diagram (2.26), we expect now that the Klein bottle amplitude can be computed as the overlap of two so-called *crosscap* states $|C\rangle$ in the following way

$$\widetilde{\mathcal{Z}}^{\mathcal{K}}(l) = \langle \Theta\; C|\, e^{-2\pi l\left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\,|C\rangle\,. \tag{2.59}$$

Considering then again Fig. 2.6d we find the modular parameter in the tree- and loop-channel as

$$\tau_{\text{open}} = \frac{\alpha_2}{\alpha_1} = \frac{2it}{\frac{1}{2}} = 4it, \qquad \tau_{\text{closed}} = \frac{\alpha_2}{\alpha_1} = \frac{i}{l},$$

and because of the tree-channel–loop-channel equivalence, they have to be equal. This implies that the length of the cylinder in Fig. 2.6e and equation (2.59) can be expressed as $l = \frac{1}{4t}$. We will elaborate on these crosscap states in more detail in the next section.

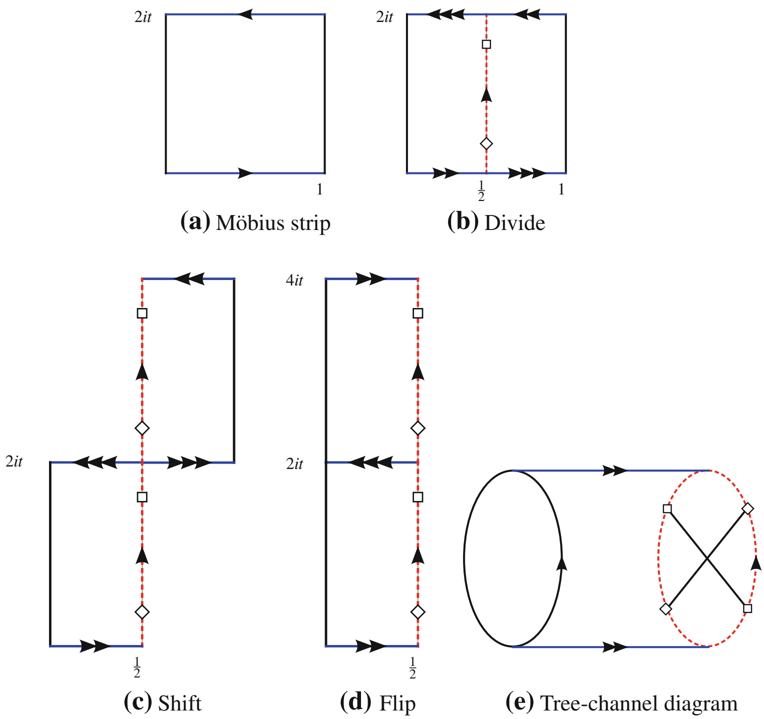


Fig. 2.7a–e Transformation of the fundamental domain of the Möbius strip to a tree-channel diagram between an ordinary boundary and a crosscap

For the Möbius strip amplitude, we can apply the same cuts and shifts as for the Klein bottle amplitude. As it is illustrated in Fig. 2.7, the resulting tree-channel diagram is a cylinder between an ordinary boundary and a crosscap. We thus expect that in the tree-channel, we can calculate the Möbius strip in the following way

$$\widetilde{\mathscr{Z}}^{\mathscr{M}}(l) = \langle \Theta\ C|\, e^{-2\pi l \left(L_0 + \overline{L}_0 - \frac{c+\bar{c}}{24}\right)}\, |B\rangle\,. \tag{2.60}$$

Finally, for the modular parameters in the tree- and loop-channel, we obtain

$$\tau_{\text{open}} = \frac{\alpha_2}{\alpha_1} = \frac{4it}{\frac{1}{2}} = 8it, \quad \tau_{\text{closed}} = \frac{\alpha_2}{\alpha_1} = \frac{i}{l},$$

which leads us to $l = \frac{1}{8t}$.

Remarks

- A summary of the various loop-channel and tree-channel expressions together with their modular parameters can be found in Table 2.1.

Table 2.1 Summary of loop-channel–tree-channel relations

	Loop-channel	$\tau = \ldots$	Tree-channel	$l = \ldots$
Torus	$\mathrm{Tr}_{\mathscr{H}\times\overline{\mathscr{H}}}\left(q^{L_0-\frac{c}{24}}\ \overline{q}^{\overline{L}_0-\frac{\overline{c}}{24}}\right)$	$\tau = \tau_1 + i\tau_2$		
Cylinder	$\mathrm{Tr}_{\mathscr{H}_{\mathscr{B}}}\left(q^{L_0-\frac{c}{24}}\right)$	$\tau = it$	$\langle\Theta\ B\|\, e^{-2\pi l\left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\,\|B\rangle$	$l = \frac{1}{2t}$
Klein bottle	$\mathrm{Tr}_{\mathscr{H}_{\text{sym.}}}\left(q^{L_0-\frac{c}{24}}\right)$	$\tau = 2it$	$\langle\Theta\ C\|\, e^{-2\pi l\left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\,\|C\rangle$	$l = \frac{1}{4t}$
Möbius strip	$\mathrm{Tr}_{\mathscr{H}_{\mathscr{B}}}\left(\Omega\ q^{L_0-\frac{c}{24}}\right)$	$\tau = it$	$\langle\Theta\ B\|\, e^{-2\pi l\left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)}\,\|C\rangle$	$l = \frac{1}{8t}$

- Almost all Ω projected CFTs in the closed sector are inconsistent and require the introduction of appropriate boundaries with corresponding boundary states. In string theory, these conditions are known as the tadpole cancellation conditions which we will discuss in the final Sect. 2.7.

2.5 Crosscap States for the Free Boson

Similarly to boundary states which describe the coupling of the closed sector of a CFT to a boundary, for orientifold theories there should exist a coherent state describing the coupling of the closed sector to the crosscap. In particular, analogous to the observation that a world-sheet boundary defines (or is confined to) a space-time D-brane, we say that a world-sheet crosscap defines (or is confined to) a space-time orientifold plane.

In this section, we will discuss crosscap states for the example of the free boson, and in the next section we are going to generalise the appearing structure to RCFTs.

Crosscap Conditions

We start our study of crosscap states by recalling the transformation of the Klein bottle respectively Möbius strip amplitude from the open to the closed sector shown in Figs. 2.6 and 2.7. There, we encountered a new type of boundary, the so-called crosscap, where opposite points are identified. For the construction of the crosscap state, we will employ this geometric intuition, however, later we also compute the tree-channel Klein bottle and Möbius strip amplitudes to check that they are indeed related via a modular transformation to the result in the loop-channel.

As it is illustrated in Fig. 2.8, in an appropriate coordinate system on a crosscap, we observe that points x on a circle are identified with $-x$. Parametrising this circle by $\sigma \in [0, 2\pi)$, we see that the identification $x \sim -x$ corresponds to $\sigma \sim \sigma + \pi$. For a closed string on a crosscap, we thus infer that the field X at (τ, σ) should be identified with the field X at $(\tau, \sigma + \pi)$. More concretely, this reads

$$X\big(\tau, \sigma\big)\ \big|C\big\rangle = X\big(\tau, \sigma + \pi\big)\ \big|C\big\rangle, \tag{2.61}$$

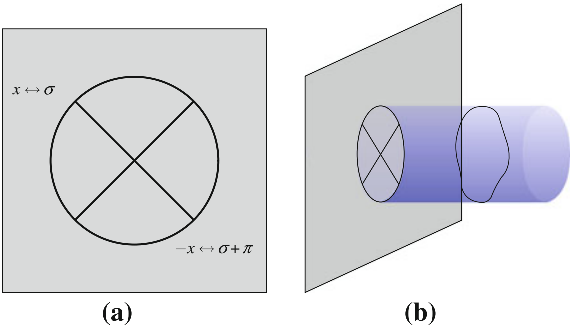


Fig. 2.8 Illustration of how points are identified on a crosscap, and how a closed string couples to a crosscap. **a** Identification of points on a crosscap. **b** Closed string at a cross-cap

and for the derivatives with respect to τ and σ, we impose

$$\begin{aligned}\big(\partial_\sigma X\big)(\tau,\sigma)\;\big|C\big\rangle &= +\big(\partial_\sigma X\big)(\tau,\sigma+\pi)\;\big|C\big\rangle,\\ \big(\partial_\tau X\big)(\tau,\sigma)\;\big|C\big\rangle &= -\big(\partial_\tau X\big)(\tau,\sigma+\pi)\;\big|C\big\rangle.\end{aligned} \tag{2.62}$$

Let us now choose coordinates such that $\tau = 0$ describes the field $X(\tau,\sigma)$ at the crosscap $|C\rangle$. Using then the Laurent mode expansions (2.5) as well as (2.62) with $\tau = 0$, we obtain that

$$\begin{aligned}\big(j_n - \bar{j}_{-n}\big)\;\big|C\big\rangle &= +(-1)^n\big(j_n - \bar{j}_{-n}\big)\;\big|C\big\rangle,\\ \big(j_n + \bar{j}_{-n}\big)\;\big|C\big\rangle &= -(-1)^n\big(j_n + \bar{j}_{-n}\big)\;\big|C\big\rangle,\end{aligned}$$

where, similarly as in the computation for the boundary states, we performed a change in the summation index $n \to -n$. By adding or subtracting these two expressions, we arrive at the gluing conditions for crosscap states

$$\boxed{\big(j_n + (-1)^n\,\bar{j}_{-n}\big)\;\big|C_{\mathrm{O1}}\big\rangle = 0}\,. \tag{2.63}$$

Note that we added the label O1 which stands for *orientifold one-plane*. The reason is that by inserting the expansion (2.8) of $X(\tau,\sigma)$ into (2.61), we see that the center of mass coordinate x_0 of the closed string is unconstrained. In the target space, the location of the crosscap is called an orientifold plane which in the present case fills out one dimension because there is no constraint on x_0. This explains the notation above.

Construction of Crosscap States

Apart from the factor $(-1)^n$, the gluing conditions (2.63) are very similar to those of a boundary state (2.20) with Neumann conditions. The solution to the gluing conditions is therefore also similar to the Neumann boundary state and reads

$$\boxed{\big|C_{\rm O1}\big\rangle = \frac{\kappa}{\sqrt{2}}\, \exp\!\bigg(-\sum_{k=1}^{\infty} \frac{(-1)^k}{k}\, j_{-k}\, \bar{j}_{-k}\bigg)\big|\, 0\,\big\rangle} \tag{2.64}$$

where we employed (2.33) and allowed for a relative normalisation factor κ between the boundary state with Neumann conditions $|B_{\rm N}\rangle$ and the crosscap state $|C_{\rm O1}\rangle$.

The proof that (2.64) is a solution to the gluing conditions (2.63) is analogous to the one shown on page 62. Note in particular, the crosscap state can be written as

$$\big|C_{\rm O1}\big\rangle = \frac{\kappa}{\sqrt{2}} \sum_{\mathbf{m}} |\mathbf{m}\rangle \otimes |U\overline{\mathbf{m}}\rangle \tag{2.65}$$

with the anti-unitary operator U acting in the following way

$$U\, \bar{j}_n\, U^{-1} = -(-1)^n\, \big(\bar{j}_{-n}\big)^{\dagger}. \tag{2.66}$$

Remark

Let us make the following remark. In equation (2.42), we have chosen a specific orientifold action Ω for the fields $X(\tau,\sigma)$ which leaves the action (2.1) invariant. However, we can also accompany Ω by another operation, for instance $\mathscr{R} : X(\tau,\sigma) \mapsto -X(\tau,\sigma)$, which also leaves (2.1) invariant. The combined action then reads

$$\Omega\, \mathscr{R} :\; X\big(\tau,\sigma\big) \mapsto \widetilde{X}\big(\tau,\sigma\big) = -X\big(\tau,-\sigma\big).$$

Note that this orientifold action describes a different theory and that there is no direct relation to the results obtained previously.

Performing the same steps as before, we arrive at the following expressions for the combined action $\Omega\mathscr{R}$ on the Laurent modes j_n and $\bar{j}_n$

closed sector $\quad \Omega\, \mathscr{R}\, j_n \big(\Omega\, \mathscr{R}\big)^{-1} = -\bar{j}_n, \quad \Omega\, \mathscr{R}\, \bar{j}_n \big(\Omega\, \mathscr{R}\big)^{-1} = -j_n,$

open sector $\quad \Omega\, \mathscr{R}\, j_n \big(\Omega\, \mathscr{R}\big)^{-1} = \mp(-1)^n\, j_n.$

For the action of $\mathscr{R}$ on the states, we find

$$\mathscr{R}\, \big|\mathbf{m}\big\rangle = \big(-1\big)^{\sum_k m_k}\, \big|\mathbf{m}\big\rangle,$$

which results in additional factors of (-1) in various loop-channel amplitudes. Concerning the construction of crosscap states, also the identification (2.61) receives an factor of (-1) which results in gluing conditions of the form

$$\big(j_n - (-1)^n\, \bar{j}_{-n}\big)\, \big|C_{\rm O0}\big\rangle = 0,$$

which is similar to the Dirichlet conditions for boundary states. The notation O0 indicates that the orientifold plane does not extend in one dimension but is only a point.

And indeed, using the expansion (2.8) of $X(\tau,\sigma)$ for $X(\tau,\sigma)|C\rangle = -X(\tau,\sigma)|C\rangle$, we see that the center of mass coordinate x_0 is constrained to $x_0 = 0$. Finally, we note that the solution to the gluing conditions in the present case reads

$$\big|C_{\mathrm{O}0}\big\rangle = \kappa\, \exp\Bigg(+ \sum_{k=1}^{\infty} \frac{(-1)^k}{k}\, j_{-k}\, \bar{j}_{-k} \Bigg) \big|\, 0\, \big\rangle .$$

After this remark about a different possibility for an orientifold projection, let us continue our studies with our original choice (2.42) which leads to O1 crosscap states $|C_{\mathrm{O}1}\rangle$.

Free Boson V: Klein Bottle Amplitude (Tree-Channel)

As we have argued in the previous section, from the overlap of two crosscap states we can compute the Klein bottle amplitude (2.59) in the closed sector, that is in the tree-channel. In order to do so, we recall the crosscap state (2.65) with the action of U given in (2.66). Noting for a basis state (2.23) that

$$U\,\big|\mathbf{m}\big\rangle = \prod_{k=1}^{\infty} (-1)^{m_k}\, (-1)^{m_k k}\, \big|\mathbf{m}\big\rangle, \tag{2.67}$$

and following the same calculation as on page 65 for the overlap of two boundary states in the Neumann–Neumann sector, we obtain

$$\tilde{\mathcal{Z}}^{\mathcal{K}\,(\mathrm{O}1,\mathrm{O}1)}_{\mathrm{bos.}}(l) = \big\langle \Theta\, C_{\mathrm{O}1}\big|\, e^{-2\pi l (L_0 + \bar{L}_0 - \frac{c+\bar{c}}{24})}\, \big|C_{\mathrm{O}1}\big\rangle = \frac{\kappa^2}{2\,\eta(2il)}. \tag{2.68}$$

Note that Θ is again the CPT operator introduced in equation (2.28) which, in particular, acts as complex conjugation on numbers. Finally, recalling from Table 2.1 the relation $l = \frac{1}{4t}$ between the tree-channel and loop-channel modular parameters, we find the loop-channel amplitude to be of the form

$$\tilde{\mathcal{Z}}^{\mathcal{K}\,(\mathrm{O}1,\mathrm{O}1)}_{\mathrm{bos.}}(l) = \frac{\kappa^2}{2\,\eta(2il)} \xrightarrow{\; l=\frac{1}{4t} \;} \frac{\kappa^2}{2\,\eta\left(-\frac{1}{2it}\right)} = \frac{\kappa^2}{2\,\sqrt{2t}}\, \frac{1}{\eta(2it)}\, ,$$

where we employed the modular property of the Dedekind η-function from equation (2.32). By comparing with the loop-channel result (2.53), we can now fix

$$\boxed{\kappa = \sqrt{2}}\, .$$

Free Boson VI: Möbius-Strip Amplitude (Tree-Channel)

Eventually, we compute the overlap of a crosscap state and a boundary state giving the tree-level Möbius strip amplitude. Employing equation (2.67) and performing a similar calculation as on page 65, we find for the Möbius strip diagram in the Neumann sector that

$$\begin{aligned}\tilde{\mathcal{Z}}^{\mathcal{M}(\mathrm{O1,N})}_{\mathrm{bos.}}(l) &= \left\langle \Theta\; C_{\mathrm{O1}} \right| e^{-2\pi l (L_0+\overline{L}_0-\frac{c+\overline{c}}{24})} \left| B_N \right\rangle \\ &= \frac{1}{\sqrt{2}}\, e^{\frac{\pi l}{6}} \prod_{k=1}^{\infty} \frac{1}{1-\left(-e^{-4\pi l}\right)^k} \\ &= \frac{1}{\sqrt{2}}\, e^{\frac{\pi i}{24}}\, \frac{1}{\eta\left(\frac{1}{2}+2il\right)} \end{aligned} \tag{2.69}$$

where we expressed (-1) as $e^{\pi i}$ and absorbed the additional factor into the definition of the modular parameter. The computation of the Möbius strip amplitude in the Dirichlet sector is very similar to the Neumann sector. We find

$$\begin{aligned}\tilde{\mathcal{Z}}^{\mathcal{M}(\mathrm{O1,D})}_{\mathrm{bos.}}(l) &= \left\langle \Theta\; C_{\mathrm{O1}} \right| e^{-2\pi l (L_0+\overline{L}_0-\frac{c+\overline{c}}{24})} \left| B_D \right\rangle \\ &= e^{\frac{\pi l}{6}} \prod_{k=1}^{\infty} \frac{1}{1+\left(-e^{-4\pi l}\right)^k} \\ &= \sqrt{2}\, e^{\frac{\pi i}{24}} \sqrt{\frac{\eta\left(\frac{1}{2}+2il\right)}{\vartheta_2\left(\frac{1}{2}+2il\right)}} \end{aligned}$$

where we used again the definition of the ϑ-functions. The momentum integration in this sector is trivial since j_0 acting on the crosscap state vanishes. This is again similar to the computation of the cylinder amplitude for mixed boundary conditions shown on page 67.

Modular transformations

After having computed the tree-channel Möbius strip amplitudes, we would like to transform these results to the loop-channel via the relation $l = \frac{1}{8t}$. However, by comparing with the loop-channel results (2.57) and (2.58), we see that this cannot be achieved by a modular S-transformation. Instead, we have to perform the following combination of T- and S-transformations

$$\mathcal{P} = TST^2\, S. \tag{2.70}$$

For the η-function with shifted argument, this transformation reads

$$\begin{aligned}
\eta\left(\frac{1}{2}+2il\right) \quad &\xrightarrow{S} \quad \eta\left(-\frac{1}{\frac{1}{2}+2il}\right)\sqrt{\frac{i}{\frac{1}{2}+2il}} \\
&\xrightarrow{T^2} \quad \eta\left(+\frac{4il}{\frac{1}{2}+2il}\right)\sqrt{\frac{i}{\frac{1}{2}+2il}}\; e^{-\frac{\pi i}{6}} \\
&\xrightarrow{S} \quad \eta\left(-\frac{\frac{1}{2}+2il}{4il}\right)\sqrt{\frac{i}{\frac{1}{2}+2il}}\sqrt{\frac{\frac{1}{2}+2il}{4l}}\; e^{-\frac{\pi i}{6}} \\
&\xrightarrow{T} \quad \eta\left(\frac{1}{2}+\frac{i}{8l}\right)\frac{1}{\sqrt{4l}}\,\sqrt{i}\, e^{-\frac{\pi i}{6}}\, e^{-\frac{\pi i}{12}} \\
&\qquad = \eta\left(\frac{1}{2}+\frac{i}{8l}\right)\frac{1}{\sqrt{4l}}
\end{aligned}$$

where in the last step we employed that $\sqrt{i} = e^{\frac{\pi i}{4}}$. For the Möbius strip amplitude with Neumann boundary conditions, we then compute the transformation from the tree-channel to the loop-channel as follows

$$\tilde{\mathscr{Z}}^{\mathscr{M}(\mathrm{O1,N})}_{\mathrm{bos.}}(l) = \frac{e^{\frac{\pi i}{24}}}{\sqrt{2}}\,\frac{1}{\eta\left(\frac{1}{2}+2il\right)} \quad \xrightarrow[l=\frac{1}{8t}]{\mathscr{P}} \quad e^{\frac{\pi i}{24}}\,\frac{1}{2\sqrt{t}}\,\frac{1}{\eta\left(\frac{1}{2}+it\right)}.$$

By comparing with the loop-channel result (2.57), we have verified the loop-channel—tree-channel equivalence for the Möbius strip amplitude in the Neumann sector.

In passing, we note that the Möbius strip loop- and tree-channel amplitudes for the Dirichlet sector are also related via a modular $\mathscr{P}$-transformation. In the same manner as above, one can then establish the loop-channel—tree-channel equivalence.

New Characters

In the last paragraph of this section, let us introduce a more general notation for the Möbius strip characters. We define *hatted* characters $\widehat{\chi}(\tau)$ in terms of the usual characters $\chi(\tau)$ as follows

$$\widehat{\chi}(\tau) = e^{-\pi i\,(h-\frac{c}{24})}\,\chi\left(\tau+\frac{1}{2}\right). \tag{2.71}$$

The action of the $\mathscr{P}$-transformation (2.70) for the new characters $\widehat{\chi}(\tau)$ can be deduced as follows. From the mapping of the modular parameter $\tau = 2il$ under the combination of S- and T-transformations

$$2il \quad \xrightarrow{T^{\frac{1}{2}}} \quad 2il+\frac{1}{2} \quad \xrightarrow{TST^2S} \quad \frac{i}{8l}+\frac{1}{2} \quad \xrightarrow{T^{-\frac{1}{2}}} \quad \frac{i}{8l},$$

we can infer the transformation of the hatted characters $\widehat{\chi}(\tau)$ as

$$\widehat{\chi}_i\big(\frac{i}{8l}\big) = \sum_j P_{ij}\, \widehat{\chi}_j\big(2il\big) \qquad \text{with} \qquad P = T^{\frac{1}{2}}\, S\, T^2\, S\, T^{\frac{1}{2}},$$

where $T^{\frac{1}{2}}$ is defined as the square root of the entries in the diagonal matrix

$$T_{ij} = \delta_{ij} e^{2\pi i (h_i - \frac{c}{24})}. \tag{2.72}$$

Note that the P-transformation corresponds to the S-transformation of the usual characters, in particular, P realises the loop-channel–tree-channel equivalence.

Finally, using some properties of the S-matrix

$$S^\dagger S = S S^\dagger = \mathbb{1}, \qquad S^T = S \tag{2.73}$$

as well as the relation $S^2 = (ST)^3 = C$ with C the charge conjugation matrix

$$\mathscr{P}^2 = C, \qquad P^2 = C, \qquad PP^\dagger = P^\dagger P = \mathbb{1}, \qquad P^T = P. \tag{2.74}$$

2.6 Crosscap States for RCFTs

Let us now generalise the construction of crosscap states to Conformal Field Theories without a Lagrangian description. In particular, we focus on RCFTs and we mainly state the general structure without explicit derivation.

Construction of Crosscap States

The crosscap gluing conditions for the generators of a symmetry algebra $\mathscr{A} \otimes \overline{\mathscr{A}}$ are in analogy to the conditions (2.63) for the example of the free boson and read

$$\begin{aligned} \big(\, L_n - (-1)^n\, \overline{L}_{-n} \,\big)\, \big|C\big\rangle = 0 \quad & \text{conformal symmetry,} \\ \big(\, W_n^i - (-1)^n\, (-1)^{h^i}\, \overline{W}^i_{-n} \,\big)\, \big|C\big\rangle = 0 \quad & \text{extended symmetries,} \end{aligned} \tag{2.75}$$

with again $h^i = h(W^i)$. For $\mathscr{A} = \overline{\mathscr{A}}$ and $\overline{\mathscr{H}}_i = \mathscr{H}_i^+$, we can define crosscap Ishibashi states $|\mathscr{C}_i\rangle\rangle$ satisfying the crosscap gluing conditions. A crosscap state $|C\rangle$ can then be expressed as a linear combination of the crosscap Ishibashi states in the following way

$$\big|C\big\rangle = \sum_i \varGamma^i\; |\, \mathscr{C}_i\rangle\rangle\ . \tag{2.76}$$

In fact, the crosscap Ishibashi states and the boundary Ishibashi states are related via

$$|\,\mathscr{C}_i\rangle\rangle = e^{\pi\, i(L_0-h(\phi_i))}\;|\mathscr{B}_i\rangle\rangle\;. \tag{2.77}$$

Indeed, knowing that the boundary Ishibashi states $|\mathscr{B}_i\rangle\rangle$ satisfy the gluing conditions (2.34), we can show that the (2.77) satisfy the crosscap gluing conditions. To do so, we compute

$$e^{-\pi i L_0}\;L_n\;e^{+\pi i L_0} = (-1)^n\;L_n, \qquad e^{-\pi i L_0}\;W^i_n\;e^{+\pi i L_0} = (-1)^n\;W^i_n,$$

where we used that W^i is a primary field. For the generators of the conformal symmetry, we can then calculate

$$\begin{aligned}
&e^{-\pi\, i(L_0-h(\phi_i))}\left(L_n-(-1)^n\,\overline{L}_{-n}\right)|\,\mathscr{C}_i\rangle\rangle\\
&= e^{-\pi\, i(L_0-h(\phi_i))}\left(L_n-(-1)^n\,\overline{L}_{-n}\right)e^{\pi\, i(L_0-h(\phi_i))}\;|\mathscr{B}_i\rangle\rangle\\
&= (-1)^n\left(L_n-\overline{L}_{-n}\right)|\mathscr{B}_i\rangle\rangle\\
&= 0,
\end{aligned}$$

and the condition for the extended symmetry generators is obtained along the same lines. Therefore, the crosscap Ishibashi states (2.77) satisfy the gluing conditions (2.75).

The Cardy Condition

Similarly to the boundary states, we expect generalisations of the Cardy condition arising from the loop-channel—tree-channel equivalences of the Klein bottle and Möbius strip amplitudes. In order to study this point, we compute the Klein bottle amplitude in the following way

$$\begin{aligned}
\widetilde{\mathscr{Z}}^{\mathscr{K}}(l) &= \left\langle \Theta\; C\right| e^{-2\pi l\left(L_0+\overline{L}_0-\frac{c+\bar{c}}{24}\right)}\left|\;C\right\rangle\\
&= \sum_{i,j}\varGamma^i\;\varGamma^j\;\langle\langle\mathscr{B}_{i+}|\;e^{\pi\, i(L_0-h(\phi_i))}\;e^{2\pi i(2il)\left(L_0-\frac{c}{24}\right)}\;e^{\pi\, i\left(L_0-h(\phi_j)\right)}\;|\mathscr{B}_j\rangle\rangle\\
&= \sum_{i,j}\varGamma^i\;\varGamma^j\;\delta_{ij}\;e^{-2\pi i\left(h(\phi_j)-\frac{c}{24}\right)}\;\langle\langle\mathscr{B}_j|\;e^{2\pi i(2il+1)\left(L_0-\frac{c}{24}\right)}\;|\mathscr{B}_j\rangle\rangle\\
&= \sum_i\left(\varGamma^i\right)^2\;e^{-2\pi i\left(h(\phi_i)-\frac{c}{24}\right)}\;\chi_i(2il+1)\\
&= \sum_i\left(\varGamma^i\right)^2\;e^{-2\pi i\left(h(\phi_i)-\frac{c}{24}\right)}\;\sum_j T_{ij}\;\chi_j(2il) = \sum_i\left(\varGamma^i\right)^2\;\chi_i(2il),
\end{aligned}$$

where Θ is again the CPT operator shown for instance in (2.40), and where we employed Eq. (2.37) as well as the modular T-matrix given in (2.72). In the next step, we perform a modular S-transformation to obtain the result in the loop-channel

$$\tilde{\mathscr{Z}}^{\mathscr{K}}(l) = \sum_i \left(\Gamma^i\right)^2 \chi_i(2il) = \sum_{i,j} \left(\Gamma^i\right)^2 S_{ij}\, \chi_j(2it).$$

Now, the Cardy condition is again the requirement that the expression above can be interpreted as a partition function. Since this partition function includes the action of the orientifold projection Ω, the coefficient in front of the character has to be integer but does not need to be non-negative

$$\sum_i \left(\Gamma^i\right)^2 S_{ij} = \kappa_j \in \mathbb{Z}\,.$$

For the Möbius strip amplitude, we compute along similar lines

$$\begin{aligned}
\tilde{\mathscr{Z}}^{\mathscr{M}}(l) &= \langle \Theta\, C | e^{-2\pi l \left(L_0+\overline{L}_0-\frac{c+\overline{c}}{24}\right)} | B_\alpha \rangle \\
&= \sum_{i,j} \Gamma^i\, B^j_\alpha\ \langle\langle \mathscr{B}_{i^+}|\, e^{\pi\, i(L_0-h(\phi_i))}\, e^{2\pi i(2il)\left(L_0-\frac{c}{24}\right)}\, |\mathscr{B}_j\rangle\rangle \\
&= \sum_{i,j} \Gamma^i\, B^j_\alpha\, \delta_{ij}\, e^{-\pi i\left(h(\phi_i)-\frac{c}{24}\right)}\, \langle\langle \mathscr{B}_j|\, e^{2\pi i\left(2il+\frac{1}{2}\right)\left(L_0-\frac{c}{24}\right)}\, |\mathscr{B}_j\rangle\rangle \\
&= \sum_i \Gamma^i\, B^i_\alpha\, e^{-\pi i\left(h(\phi_i)-\frac{c}{24}\right)}\, \chi_i\!\left(2il+\tfrac{1}{2}\right) \\
&= \sum_i \Gamma^i\, B^i_\alpha\, \widehat{\chi}_i(2il) = \sum_{i,j} \Gamma^i\, B^i_\alpha\, P_{ij}\, \widehat{\chi}_j\!\left(it\right),
\end{aligned}$$

where we employed the hatted characters (2.71) together with their modular transformation. Interpreting this expression as a loop-channel partition function, we see that the coefficients have to be integer

$$\sum_i \Gamma^i\, B^i_\alpha\, P_{ij} = m_{\alpha j} \quad \in \quad \mathbb{Z}\,.$$

Similar to the Cardy boundary states, for the charge conjugate modular invariant partition function explained on page 73, one can show that these integer conditions are satisfied for the reflection coefficients of the form

$$\Gamma^i = \frac{P_{0i}}{\sqrt{S_{0i}}}, \quad B^i_\alpha = \frac{S_{\alpha\, i}}{\sqrt{S_{0i}}}\,.$$

The Klein bottle and Möbius strip coefficients can then be written as two Verlinde type formulas

$$\kappa_j = \sum_i \frac{P_{0i}P_{0i}S_{ij}}{S_{0i}} = Y^0_{j0}, \qquad m_{\alpha j} = \sum_i \frac{S_{\alpha i}P_{0i}P_{ij}}{S_{0i}} = Y^0_{\alpha j}.$$

From the relations (2.74), we can deduce $P^*_{ij} = P_{ij^+}$ and in particular $P^*_{0i} = P_{0i}$, which allows us to establish the connection to the general coefficients

$$Y_{ij}^{k} = \sum_{l} \frac{S_{il}\, P_{jl}\, P_{kl}^{*}}{S_{0l}} \, .$$

As it turns out, the coefficients Y_{ij}^{k} are integer, guaranteeing that the loop-channel Klein bottle and Möbius strip amplitudes contain only integer coefficients.

Remark

With the techniques presented in this section, it is possible to construct many orientifolds of Conformal Field Theories. However, one set of essential consistency conditions for the co-existence of crosscap and boundary states is still missing. These are the so-called tadpole cancellation conditions which we are going to discuss in a simple example in the final section of these lecture notes.

2.7 The Orientifold of the Bosonic String

We finally apply the techniques developed in this lecture to orientifold theories with boundaries and crosscaps. In particular, we are going to consider a string theory motivated but still sufficiently simple orientifold model which is the Ω projection of the bosonic string. More interestingly, this theory is actually analogous to the orientifold construction of the Type IIB superstring leading to the so-called Type I superstring. However, this needs a more detailed treatment of free fermions which we have not presented here and which is not necessary to understand the mathematical structure of such theories.

Details on the String Theory Construction

The bosonic string is only consistent in 26 flat space-time dimensions and is thus described by 26 free bosons $X^{\mu}(\sigma, \tau)$ with $\mu = 0, \ldots, 25$. The quantisation of string theory in this description, the covariant quantisation, is slightly involved. However, by defining

$$X^{+} = \frac{1}{\sqrt{2}}\Big(X^{0}(\sigma, \tau) + X^{1}(\sigma, \tau)\Big), \quad X^{-} = \frac{1}{\sqrt{2}}\Big(X^{0}(\sigma, \tau) - X^{1}(\sigma, \tau)\Big), \tag{2.78}$$

imposing so-called light-cone gauge and using constraint equations, we are left only left with the momentum p^{+} as a degree of freedom. For the computation of the characters, we can therefore simply *ignore* the contribution from $X^{0}(\sigma, \tau)$ and $X^{1}(\sigma, \tau)$ so that we are left with the Conformal Field Theory of 24 free bosons $X^{I}(\tau, \sigma)$ where $I = 2, \ldots, 25$. Since the bosonic string is made out of 24 copies of the free boson CFT, for the computation of the partition functions we can use our previous results. These have been summarized in Table 2.2 for later reference.

Table 2.2 Summary of all loop- and tree-channel amplitudes for the example of the free boson with orientifold projection (2.42)

Loop-channel	Tree-channel
$\mathcal{Z}^{\mathcal{T}}_{\text{bos.}}(\tau,\overline{\tau}) = \frac{1}{\sqrt{\tau_2}}\,\frac{1}{\lvert\eta(\tau)\rvert^2}$	
$\mathcal{Z}^{\mathcal{K}}_{\text{bos.}}(t) = \frac{1}{\sqrt{2t}}\,\frac{1}{\eta\left(2it\right)}$	$\widetilde{\mathcal{Z}}^{\mathcal{K}\,(\text{O1,O1})}_{\text{bos.}}\left(l\right) = \frac{1}{\eta(2il)}$
$\mathcal{Z}^{\mathcal{C}\,(\text{N,N})}_{\text{bos.}}(t) = \frac{1}{2\sqrt{t}}\,\frac{1}{\eta\,(it)}$	$\widetilde{\mathcal{Z}}^{\mathcal{C}\,(\text{N,N})}_{\text{bos.}}(l) = \frac{1}{2\,\eta\left(2il\right)}$
$\mathcal{Z}^{\mathcal{C}\,(\text{D,D})}_{\text{bos.}}(t) = \frac{1}{\eta\,(it)}\,e^{-\frac{t}{4\pi}\left(x_0^b-x_0^a\right)^2}$	$\widetilde{\mathcal{Z}}^{\mathcal{C}\,(\text{D,D})}_{\text{bos.}}(l) = \frac{1}{\sqrt{2l}}\,\frac{1}{\eta\left(2il\right)}\,e^{-\frac{1}{8\pi l}\left(x_0^b-x_0^a\right)^2}$
$\mathcal{Z}^{\mathcal{C}\,(\text{mixed})}_{\text{bos.}}(t) = \sqrt{\frac{\eta\,(it)}{\vartheta_4(it)}}$	$\widetilde{\mathcal{Z}}^{\mathcal{C}\,(\text{mixed})}_{\text{bos.}}(l) = \sqrt{\frac{\eta\,(2il)}{\vartheta_2(2il)}}$
$\mathcal{Z}^{\mathcal{M}\,(\text{N,N})}_{\text{bos.}}(t) = \frac{1}{2\sqrt{t}}\,\frac{1}{\eta\left(\frac{1}{2}+it\right)}\,e^{\frac{\pi i}{24}}$	$\widetilde{\mathcal{Z}}^{\mathcal{M}\,(\text{O1,N})}_{\text{bos.}}\left(l\right) = \frac{1}{\sqrt{2}}\,\frac{1}{\eta\left(\frac{1}{2}+2il\right)}\,e^{\frac{\pi i}{24}}$
$\mathcal{Z}^{\mathcal{M}\,(\text{D,D})}_{\text{bos.}}(t) = \sqrt{2}\sqrt{\frac{\eta\left(\frac{1}{2}+it\right)}{\vartheta_2\left(\frac{1}{2}+it\right)}}\,e^{\frac{\pi i}{24}}$	$\widetilde{\mathcal{Z}}^{\mathcal{M}\,(\text{O1,D})}_{\text{bos.}}\left(l\right) = \sqrt{2}\sqrt{\frac{\eta\left(\frac{1}{2}+2il\right)}{\vartheta_2\left(\frac{1}{2}+2il\right)}}\,e^{\frac{\pi i}{24}}$

In our previous definition of the open and closed sector partition functions, we employed the notion common to Conformal Field Theory. However, for the relevant quantities in string theory, we have to integrate over the modular parameter of the torus, Klein bottle, cylinder and Möbius strip. After performing the integration over the light-cone momentum p^+, the expressions relevant for the following are

$$\begin{aligned} Z^{\mathcal{T}} &= \int_{\text{Teich}} \frac{d^2\tau}{\tau_2^2}\,\mathcal{Z}^{\mathcal{T}}(\tau,\overline{\tau}), \qquad & Z^{\mathcal{C}} &= \int_0^\infty \frac{dt}{4\,t^2}\,\mathcal{Z}^{\mathcal{C}}(t), \\ Z^{\mathcal{K}} &= \int_0^\infty \frac{dt}{2\,t^2}\,\mathcal{Z}^{\mathcal{K}}(t), & Z^{\mathcal{M}} &= \int_0^\infty \frac{dt}{4\,t^2}\,\mathcal{Z}^{\mathcal{M}}(t). \end{aligned} \tag{2.79}$$

The domain of integration for the torus amplitude $Z^{\mathcal{T}}$ is the so-called Teichmüller space. It is the space of all complex structures τ of a torus $\mathbb{T}^2$ which are not related via the $SL(2,\mathbb{Z})/\mathbb{Z}_2$ symmetry. An illustration can be found in Fig. 2.9 and the precise definition reads

$$\text{Teich} = \left\{\,\tau \in \mathbb{C}\ :\ -\frac{1}{2} < \tau_1 \leq +\frac{1}{2},\ \lvert\tau\rvert \geq 1\right\}. \tag{2.80}$$

Torus Partition Function for the Bosonic String

Let us now become more concrete and determine the torus partition function for the bosonic string in light-cone gauge. Since this theory is a copy of 24 free bosons, we recall from Table 2.2 the form of $\mathcal{Z}^{\mathcal{T}}_{\text{bos.}}$ and combine it into

$$Z^{\mathcal{T}} = \int_{\text{Teich}} \frac{d^2\tau}{\tau_2^2}\left(\mathcal{Z}^{\mathcal{T}}_{\text{bos.}}(\tau,\overline{\tau})\right)^{24} = \int_{\text{Teich}} \frac{d^2\tau}{\tau_2^2}\,\frac{1}{\tau_2^{12}}\,\frac{1}{\lvert\eta^{24}(\tau)\rvert^2}\,. \tag{2.81}$$

In order to become more explicit, let us expand the Dedekind η-function in the following way

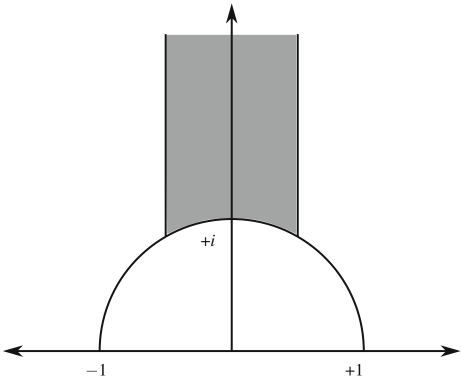


Fig. 2.9 The shaded region in this figure corresponds to the Teichmüller space of the two-torus $\mathbb{T}^2$

$$\frac{1}{\eta^{24}(\tau)} = q^{-1}\Big(\, 1 + 24\, q + 324\, q^2 + \dots \Big)\,. \tag{2.82}$$

Using this expansion in (2.81) together with the string theoretical *level-matching condition* which leaves only equal powers of q and $\overline{q}$, we arrive at

$$\begin{aligned} Z^{\mathscr{T}} &= \int_{\text{Teich}} \frac{d^2\tau}{\tau_2^{14}}\, e^{+4\pi\tau_2} \,\Big|\, 1 + 24\, e^{2\pi i \tau} + \dots \,\Big|^2 \\ &\longrightarrow \int_{\text{Teich}} \frac{d^2\tau}{\tau_2^{14}}\, e^{+4\pi\tau_2} \,\Big(\, 1 + (24)^2\, e^{-4\pi\tau_2} + \dots \Big)\,. \end{aligned} \tag{2.83}$$

Let us now study the divergent behaviour of this integral.

- Although the integrand in (2.83) diverges for $\tau_2 \to 0$ due to the factor of τ_2^{-14}, the whole integral is finite because the domain of integration (2.80) does not include $\tau_2 = 0$. Therefore, this expression is not divergent in the *infrared*, i.e. there is no singularity for small τ_2. Let us emphasize that the finiteness in this parameter region is due to the modular invariance of the torus partition function which restricts the domain of integration to the Teichmüller space.
- Next, we turn to the behaviour of (2.83) for large τ_2. We see that the first term gives rise to a divergence in the region $\tau_2 \to \infty$ which corresponds to a state with negative mass squared, i.e. a tachyon. Thus, the theory of the bosonic string is unstable. In more realistic theories, for instance the superstring, such a tachyon should be absent and we do not expect problems due to divergences in the *ultraviolet*.
- In summary, the torus partition function of the bosonic string is finite in the infrared due to modular invariance. In the ultraviolet, the partition function is divergent due to a tachyon which renders the theory unstable.

Klein Bottle Partition Function for the Bosonic String

As the title of this section suggests, we want to study the orientifold of the bosonic string and so we have to determine the Klein bottle amplitude. Following the same steps as for the torus, we arrive at

$$Z^{\mathscr{K}}(t) = \frac{1}{2}\int_0^\infty \frac{dt}{t^2}\left(\mathscr{Z}^{\mathscr{K}}_{\text{bos.}}(t)\right)^{24} = \frac{1}{2^{13}}\int_0^\infty \frac{dt}{t^{14}}\,\frac{1}{\eta^{24}(2it)}\,.$$

In order to simplify the integrand, we perform a transformation to the tree-channel with modular parameter $t = \frac{1}{4l}$ by employing the modular properties of the Dedekind η-function (2.32)

$$\begin{aligned} Z^{\mathscr{K}}(t) \quad \xrightarrow{t=\frac{1}{4l}} \quad \widetilde{Z}^{\mathscr{K}(\text{O25,O25})}(l) &= \frac{1}{2^{13}}\int_0^\infty \frac{dl}{4\,l^2}\,(4l)^{14}\,\frac{1}{\eta^{24}\left(-\frac{1}{2il}\right)} \\ &= 2\int_0^\infty dl\,\frac{1}{\eta^{24}(2il)}. \end{aligned}$$

The notation O25 deserves some explanation. Since we are studying the bosonic string in a 26-dimensional space-time, the orientifold projection naturally acts also on the light-cone coordinates (2.78). By choosing the orientifold projection (2.42), we have an orientifold plane extending over all 26 dimensions. However, the convention in string theory is such that only the space dimensions are counted which explains the term O25.

Similarly as for the torus partition function, let us now expand the tree-channel Klein bottle amplitude. Using Eq. (2.82) we obtain

$$\widetilde{Z}^{\mathscr{K}(\text{O25,O25})}(l) = 2\int_0^\infty dl\,\left(e^{4\pi\, l} + 24 + 324\, e^{-4\pi\, l} + \ldots\right). \tag{2.84}$$

The first term in (2.84) corresponds again to the tachyon and should be absent in more realistic theories. We therefore ignore this problematic behaviour. However, the second terms corresponds to massless states and gives rise to a divergence since in the present case, the domain of integration includes $t = \frac{1}{4l} = 0$. This term will not be absent in more refined theories and so at this point, the orientifold of the bosonic string is not consistent at a more severe level.

A Stack of D-Branes

As it turns out, the divergence of the Klein bottle diagram can be cancelled by introducing a to be determined number N of D25 branes. The notation D25 means that these D-branes fill out 25 spatial dimensions and it is understood that they always fill the time direction.

If we put a certain number of D-branes on top of each other, we call it a stack of D-branes. However, since there are now multiple branes, we can have new kinds of open strings. In particular, there are strings starting at D-brane i of our stack and ending on D-brane j. We thus include new labels, so-called Chan–Paton labels, to our open string states

$$\big|\mathbf{m},\ i,j\big\rangle = \big|\mathbf{m}\big\rangle \otimes \big|i,j\big\rangle,$$

where $|\mathbf{m}\rangle$ denotes the states for a single string and $i,j = 1,\ldots,N$ label the starting respectively ending points. We furthermore construct the hermitian conjugate $\langle i,j|$ in the usual way such that

$$\big\langle\, i,j \;\big|\; i',j' \,\big\rangle = \delta_{ii'}\delta_{jj'}. \tag{2.85}$$

Next, we define the action of the orientifold projection acting on the Chan–Paton labels. Since Ω changes the orientation of the world-sheet, it clearly interchanges starting and ending points of open strings. But we can also allow for rotations among the D-branes and so a general orientifold action reads

$$\Omega \;\big|\; i,j \,\big\rangle = \sum_{i',j'=1}^{N} \gamma_{jj'} \big|\, j', i' \,\big\rangle \big(\gamma^{-1}\big)_{i'i} \,, \tag{2.86}$$

where γ is a $N\times N$ matrix. Without presenting the detailed argument, we now require that the action of Ω on the Chan–Paton labels squares to the identity. For this we calculate

$$\begin{aligned}
\Omega^2 \;\big|\; i,j \,\big\rangle &= \sum_{i'',j''=1}^{N} \gamma_{ii''} \Big[\Omega \;\big|\; i,j \,\big\rangle^T\Big]_{i'',j''} \big(\gamma^{-1}\big)_{j''j} \\
&= \sum_{i',j',i'',j''=1}^{N} \gamma_{ii''}\, \big(\gamma^{-1}\big)^T_{i''i'} \;\big|\; i',j' \,\big\rangle\, \gamma^T_{j'j''}\, \big(\gamma^{-1}\big)_{j''j} \\
&= \sum_{i',j'=1}^{N} \Big[\gamma \big(\gamma^{-1}\big)^T\Big]_{ii'} \;\big|\; i',j' \,\big\rangle \Big[\gamma^T\gamma^{-1}\Big]_{j'j}\,,
\end{aligned}$$

from which we infer the constraint on the matrices γ to be symmetric or anti-symmetric

$$\gamma^T = \pm\gamma. \tag{2.87}$$

In string theory, the two different signs correspond to gauge groups $SO(N)$ and $SP(N)$ living on the stack of D-branes.

Let us now come to the final part of this paragraph which is to determine the contribution of the Chan–Paton labels to the partition function. For the Cylinder partition function, we calculate with the help of (2.85)

$$\mathscr{Z}^{\mathscr{C}}(t) = \text{Tr}_{\mathscr{H}_{\mathscr{B}}}\left(q^{L_0-\frac{c}{24}}\right) = \sum_n \langle n \mid q^{L_0-\frac{c}{24}} \mid n\rangle \times \sum_{i,j=1}^{N} \langle\, i,j \mid i,j\,\rangle$$
$$= \sum_n \langle n \mid q^{L_0-\frac{c}{24}} \mid n\rangle \times N^2.$$

Therefore, the effect of N D-branes is taken care of by including the factor N^2 for the cylinder partition function. Let us next turn to the Möbius strip partition function. Concentrating only on the Chan–Paton part, we find using (2.85) and (2.86) that

$$\sum_{i,j=1}^{N} \langle\, i,j \mid \Omega \mid i,j\,\rangle = \sum_{i,j,i',j'=1}^{N} \langle\, i,j \mid \gamma_{jj'} \mid j',i'\,\rangle (\gamma^{-1})_{i'i}$$
$$= \sum_{i,j,i',j'=1}^{N} \delta_{ij'}\, \delta_{ji'}\, \gamma_{jj'}\, \left(\gamma^{-1}\right)_{i'i}$$
$$= \text{Tr}\Big(\, \gamma^T\, \gamma^{-1}\,\Big) = \pm N,$$

where in the final step we also employed (2.87). In summary, by including a factor of $\pm N$ in the Möbius strip partition function, we can account for a stack of N D-branes.

Cylinder and Möbius-Strip Partition Function for the Bosonic String

After this discussion about stacks of D-branes, let us now compute the cylinder and Möbius strip partition functions for a stack of N D25-branes. Since the D-branes fill out the 26-dimensional space-time, the open strings always have Neumann–Neumann boundary conditions.

For the cylinder, we recall from Table 2.2 the form of a single cylinder partition function and combine it with the relevant expression from (2.79) to obtain

$$Z^{\mathscr{C}(\text{N,N})}(t) = \frac{N^2}{4} \int_0^\infty \frac{dt}{t^2} \left(\mathscr{Z}^{\mathscr{C}(\text{N,N})}_{\text{bos.}}(t)\right)^{24} = \frac{N^2}{2^{26}} \int_0^\infty \frac{dt}{t^{14}}\, \frac{1}{\eta^{24}(it)}\,,$$

where we included the factor N^2 as explained above. In order to extract the divergences, we perform a transformation from the loop- to the tree-channel via $t = \frac{1}{2l}$ to find

$$Z^{\mathscr{C}(\text{N,N})}(t) \quad \xrightarrow{t=\frac{1}{2l}} \quad \widetilde{Z}^{\mathscr{C}(\text{N,N})}(l) = \frac{N^2}{2^{26}} \int_0^\infty \frac{dl}{2\, l^2}\, (2l)^{14}\, \frac{1}{\eta^{24}\left(-\frac{1}{2il}\right)}$$
$$= \frac{N^2}{2^{25}} \int_0^\infty dl\, \frac{1}{\eta^{24}(2il)}.$$

With the help of (2.82), we can again expand this expression. The first terms read as follows

$$\tilde{Z}^{\mathscr{C}\,(\mathrm{N,N})}(l) = \frac{N^2}{2^{25}} \int_0^\infty dl \left(e^{4\pi\, l} + 24 + 324\, e^{-4\pi\, l} + \ldots \right).$$

Next, we turn to the Möbius strip contribution. Along similar lines as above, we recall from Table 2.2 the expression for the partition function of a single free boson and combine 24 copies of it into the Möbius partition function

$$Z^{\mathscr{M}\,(\mathrm{N,N})}(t) = \pm\,\frac{N}{4} \int_0^\infty \frac{dt}{t^2} \left(\mathscr{Z}^{\mathscr{M}\,(\mathrm{N,N})}_{\mathrm{bos.}}(t) \right)^{24} = \pm\,\frac{N}{2^{26}} \int_0^\infty \frac{dt}{t^{14}}\, \frac{e^{\pi i}}{\eta^{24}(\frac{1}{2} + it)} .$$

In order to extract the divergences more easily, we transform this expression into the tree-channel via the relation $t = \frac{1}{8l}$ and the modular $\mathscr{P}$ transformation (2.70)

$$\begin{aligned} Z^{\mathscr{M}\,(\mathrm{N,N})}(t) \quad \xrightarrow{t=\frac{1}{8l}} \quad \tilde{Z}^{\mathscr{M}\,(\mathrm{N,N})}(l) &= \pm\,\frac{N}{2^{26}} \int_0^\infty \frac{dl}{8\, l^2}\, (8l)^{14}\, \frac{e^{\pi i}}{\eta^{24}\left(\frac{1}{2} + \frac{i}{8l}\right)} \\ &= \pm\,\frac{N}{2^{11}} \int_0^\infty dl\, \frac{e^{\pi i}}{\eta^{24}(\frac{1}{2} + 2il)} . \end{aligned}$$

Expanding this expression with the help of (2.82), we find

$$\tilde{Z}^{\mathscr{M}\,(\mathrm{N,N})}(l) = \pm\,\frac{N}{2^{11}} \int_0^\infty dl \left(e^{4\pi\, l} - 24 + 324\, e^{-4\pi\, l} - \ldots \right).$$

Tadpole Cancellation Condition

After having determined the divergent contributions of the one-loop amplitudes, we can now combine them into the full expression. Leaving out the torus amplitude, we find

$$\begin{aligned} &\frac{1}{2} \left(\tilde{Z}^{\mathscr{K}\,(\mathrm{O25,O25})}(l) + \tilde{Z}^{\mathscr{C}\,(\mathrm{N,N})}(l) + \tilde{Z}^{\mathscr{M}\,(\mathrm{N,N})}(l) \right) \\ &\quad = 2^{-26} \int_0^\infty dl \Big(e^{4\pi l} \left(2^{26} \pm 2 \cdot 2^{13} N + N^2 \right) \\ &\qquad\qquad +24 \left(2^{26} \mp 2 \cdot 2^{13} N + N^2 \right) \\ &\qquad\qquad +324\, e^{-4\pi l} \left(2^{26} \pm 2 \cdot 2^{13} N + N^2 \right) \quad + \quad \ldots \Big). \end{aligned} \tag{2.88}$$

The first terms with prefactor $e^{4\pi l}$ stem again from the tachyon which in a more realistic theory, e.g. Superstring Theory, should be absent. We will therefore ignore this divergence. The next line with prefactor 24 corresponds to massless states which will not be absent in more refined theories. However, we can simplify this expression by noting that

$$\left(2^{26} \mp 2 \cdot 2^{13} N + N^2\right) = \left(2^{13} \mp N\right)^2.$$

We thus see that by taking $N = 2^{13} = 8192$ D25-branes and choosing the minus sign corresponding to $SO(N)$ gauge groups, the divergence is cancelled. In summary, we have found that

> For the orientifold of the bosonic string with $N = 8192$D25-branes and gauge group $SO\ (8192)$, the divergence due to massless states is cancelled. This is the famous tadpole cancellation condition for the bosonic string.

Finally, it is easy to see that the proceeding terms in (2.88) with prefactors $e^{-4\pi l}$ and powers thereof do not give rise to divergences in the integral.

Remarks

- Here we have discussed a very simple example for a CFT with boundaries. The next step is to generalise these methods for the superstring, in which case we have to define boundary and crosscap states for the CFT of the free fermion. The orientifold of the Type IIB superstring defines the so-called Type I string living in ten-dimensions and carrying gauge group $SO(32)$ instead of $SO(8192)$.
- Many examples of such orientifold models have been discussed for compactified dimensions. These include orientifolds on toroidal orbifolds and also orientifolds of Gepner models. For this purpose, one first has to find classes of boundary and crosscap states for the $\mathcal{N} = 2$ unitary models and then for Gepner models, in which the simple current construction is utilised in an essential way. Finally, one has to derive and solve the tadpole cancellation conditions. All this is a feasible exercise but beyond the scope of these lecture notes.

Acknowledgements I want to thank Erik Plauschinn for collaboration on the book [1], where these lectures notes are essentially taken from. Moreover, I am grateful to Marco Baumgartl for helping with the manuscript.

References

1. Blumenhagen, R., Plauschinn, E.: Introduction to conformal field theory, Springer Verlag. Lect. Notes Phys. **779**, 1–256 (2009)
2. Angelantonj, C., Sagnotti, A.: Open strings. Phys. Rept. **371**, 1–150 (2002) [hep-th/0204089]

Chapter 3
Introduction to Gauge/Gravity Duality

Johanna Erdmenger

3.1 Introduction

The AdS/CFT correspondence and its generalizations to gauge/gravity duality are undoubtedly one of the major developments in theoretical physics of the last decade. Following the original paper by Maldacena [1] in 1997, with the additions of Witten [2] and Gubser, Klebanov and Polyakov [3] in early 1998, a wealth of achievements has been obtained, as far as both our understanding of string theory and applications to strongly coupled systems are concerned. In its original form, the correspondence maps string theory on Anti-de Sitter spaces (AdS) to conformal quantum field theories (CFT), which have a high degree of symmetry. This is referred to as the AdS/CFT correspondence. In the meantime, many generalizations to maps between gravity theories and quantum field theories with less symmetry have been proposed. These are summarized under the notion of gauge/gravity duality.

Let us start by giving the (rough) statement of gauge/gravity duality:

- Some quantum field theories are equivalent to (quantum) gravity theories;
- In particular limits, the gravity theory becomes classical and the corresponding quantum field theory (QFT) strongly coupled.

The second point makes the duality particularly useful since by other methods, dynamical processes are inaccessible in the strongly coupled regime of QFTs: Normally, QFT calculations are done by means of perturbation theory, but this only works at weak coupling. Lattice gauge theory is a powerful way out of this dilemma in some cases, but it is hard to use for capturing dynamics, in particular at finite temperature and density.

J. Erdmenger (✉)
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut),
Föhringer Ring 6, 80805 München, Germany
e-mail: jke@mppmu.mpg.de

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_3,

Since the AdS/CFT correspondence links different areas of theoretical physics, in particular quantum field theory, general relativity, string theory, supersymmetry and conformal symmetry, we begin these lectures with an overview over some pre-requisites which are essential for understanding the correspondence. These include conformal symmetry and $\mathcal{N} = 4$ supersymmetry, large N gauge theory, the geometry of Anti-de Sitter space, and D-branes. Then we move on to explaining the correspondence in its simplest example, which is the map between $\mathcal{N} = 4\, SU(N)$ Super Yang–Mills theory and type IIB superstring theory on $AdS_5 \times S^5$. In a subtle limit, which we explain, this becomes a duality between $\mathcal{N} = 4\, SU(N)$ Super Yang–Mills theory in the limit $N \to \infty$ at strong coupling and classical type IIB supergravity on $AdS_5 \times S^5$. There is no mathematical proof of the AdS/CFT correspondence as yet, but overwhelming evidence for its correctness. The conjecture states that the two theories are equivalent including observables, states, correlation functions and dynamics.

The ten-dimensional spacetime of the string theory side contains a five-dimensional Anti-de Sitter spacetime with a four-dimensional boundary. The four-dimensional QFT can be regarded as living on this four dimensional boundary. In analogy to conventional holograms (which encode three dimensional information on a lower dimensional surface), the AdS/CFT correspondence is said to realize the *holographic principle*.

As in any field theory, symmetries are of central importance for gauge/gravity duality. The two equivalent theories have the same symmetries. Moreover, the correspondence provides a one-to-one map between classical gravity fields and quantum operators of the field theory, i.e.a holographic dictionary. This map then identifies representations of the common symmetry group.

We conclude these lecture notes with a brief outlook on more general examples of gauge/gravity duality by explaining how to include finite temperature and density on the field theory side. This is the starting point for applications for instance to the quark-gluon plasma within heavy ion physics or to strongly coupled systems of relevance for condensed matter theory.

As to the literature to this subject, let us refer to the original papers [1–4] which marked the birth of the AdS/CFT correspondence. Several review articles for the AdS/CFT correspondence followed [5–8]. Finally, [9–12] are helpful references for generalizations and applications in the context of gauge/gravity duality.

Note: The AdS/CFT correspondence can be formulated equivalently either in Minkowski or in Euclidean signature. Depending on the aspects considered, either Minkowski or Euclidean signature may be more appropriate. For instance, the calculation of n-point correlation functions is most easily performed in Euclidean signature. On the other hand, for describing time-dependent processes such as transport phenomena, using Minkowski signature is essential. In these lecture notes, we will use both signatures alternatively, while clearly indicating in each case which one is used.

3.2 Preparations

Understanding the AdS/CFT correspondence involves background knowledge in a number of areas. Most prominently, quantum field theory, general relativity, supersymmetry and some aspects of string theory are essential. In this section we recapitulate a few necessary prerequisites for AdS/CFT: These include conformal symmetry in more than two dimensions, $\mathcal{N} = 4$ supersymmetry, the geometry of Anti-de Sitter spaces and some aspects of D-branes.

3.2.1 Conformal Field Theory in d Dimensions

Conformal symmetry is of central importance for the AdS/CFT correspondence. In fact, the best known example for AdS/CFT, relating $\mathcal{N} = 4$ Super Yang–Mills theory to an appropriate supergravity theory on $AdS_5 \times S^5$, involves a conformal field theory in four dimensions. The symmetry group of conformal transformations in d dimensions coincides with the group of isometries in $d+1$-dimensional Anti-de Sitter space. This is one of the key ingredients of the correspondence.

3.2.1.1 Conformal Coordinate Transformations

Conformal coordinate transformations are defined as those local transformations $x^\mu \mapsto x'^\mu(x)$ that leave angles invariant. In a Euclidean d-dimensional space $\mathbb{R}^d$ we therefore can write

$$dx_\mu \, dx^\mu = \Omega^{-2}(x) \, dx'_\mu \, dx'^\mu . \tag{3.1}$$

The corresponding infinitesimal coordinate transformation from old coordinates x to new coordinates x' is given by

$$x'^\mu = x^\mu + v^\mu(x) \, , \tag{3.2}$$

where

$$\Omega(x) = 1 - \sigma(x), \quad \sigma(x) = \frac{1}{d} \, \partial \cdot v(x). \tag{3.3}$$

Equivalently to (3.1) we can formulate an equation for the vector v, the *conformal Killing equation*,

$$\partial_\mu v_\nu + \partial_\nu v_\mu = 2 \, \sigma(x) \, \eta_{\mu\nu}. \tag{3.4}$$

Taking its trace yields the expression (3.3) for $\sigma(x)$. We will work in d dimensional Euclidean space where $\eta_{\mu\nu} = \delta_{\mu\nu}$. Solutions v to (3.4) are referred to as *conformal Killing vectors*, the most general one reads

$$v_\mu = a_\mu + \omega_{\mu\nu} x^\nu + \lambda x_\mu + b_\mu x^2 - 2\,(b \cdot x)\, x_\mu, \quad \omega_{\mu\nu} = -\omega_{\nu\mu}. \tag{3.5}$$

This Killing vector leads to the scale factor $\sigma(x) = \lambda - 2(b \cdot x)$. Equation (3.5) is valid for any d. Note that in the special case of $d = 2$ the conformal Killing equation (3.5) is nothing but the *Cauchy-Riemann equations*

$$\partial_1 v_1 = \partial_2 v_2, \quad \partial_1 v_2 = -\partial_2 v_1. \tag{3.6}$$

Thus, in $d = 2$ all holomorphic functions $v(x)$ are solutions and generate conformal coordinate transformations. In this case we have an infinite number of functions solving (3.5), accompanied by an infinite number of associated conserved quantities. However, we will mostly consider theories in $d = 4$ dimensions, for example on the boundary of AdS_5. Here we have a finite amount of conserved quantities. Counting the independent components of the factors in the solutions (3.5) amounts to a total number of 15, namely

a_μ		4
$\omega_{\mu\nu}$	+	6
λ	+	1
b_μ	+	4
total		15.

On curved space, conformal transformations amount to a *Weyl rescaling* of the metric,

$$g_{\mu\nu}(x) \to e^{\sigma(x)} g_{\mu\nu}(x)\,. \tag{3.7}$$

The general conformal Killing vector (3.5) may be viewed as the combination of elementary transformations. The group of "large" conformal transformation is generated by infinitesimal elements of the conformal algebra. Following [13, 14], we define locally orthogonal tranformations $\mathscr{R}$ corresponding to a group element g of the conformal group as

$$\mathscr{R}^g_{\mu\alpha}(x) := \Omega^g(x)\, \frac{\partial x'_\mu}{\partial x^\alpha}. \tag{3.8}$$

One can easily show that $\mathscr{R} \in O(d)$, i.e. that $\mathscr{R}^g_{\mu\alpha}(x)\mathscr{R}^g_{\nu\alpha}(x) = \delta_{\mu\nu}$. The group multiplication and the inverse are given as follows

$$\mathscr{R}^{g'}(gx)\, \mathscr{R}^g(x) = \mathscr{R}^{g'g}(x), \quad \left(\mathscr{R}^g(x)\right)^{-1} = \mathscr{R}^{g^{-1}}(gx)\,. \tag{3.9}$$

With these we can construct translations and rotations as

$$x'_\mu = \mathscr{R}_{\mu\nu}\, x_\nu + a_\mu, \quad \Omega(x) = 1. \tag{3.10}$$

Scale transformations ($\leftrightarrow \lambda$) and *special conformal transformations* ($\leftrightarrow b_\mu$) involve a non-trivial Ω factor

$$x'_\mu = \lambda\, x_\mu, \quad \Omega(x) = \lambda, \tag{3.11}$$

$$x'_\mu = \frac{x_\mu + b_\mu x^2}{\Omega^g(x)}, \quad \Omega^g(x) = 1 + 2b \cdot x + bx^2. \tag{3.12}$$

Together, these transformations form a group isomorphic to $SO(d+1, 1)$ (or $SO(d, 2)$ in Minkowski spacetime). All transformations belonging to this group can be constructed by performing translations, rotations, and *inversions*; the latter are given by

$$x'_\mu =: (\mathrm{i}x)_\mu = \frac{x_\mu}{x^2}, \quad \Omega^{\mathrm{i}}(x) = x^2, \tag{3.13}$$

$$\mathscr{R}^{\mathrm{i}}_{\mu\nu}(x) =: I_{\mu\nu}(x) = \delta_{\mu\nu} - 2\,\frac{x_\mu x_\nu}{x^2}. \tag{3.14}$$

Special conformal transformations can be composed by concatenating inversion + translation + inversion.

3.2.1.2 Conformal Fields and Correlation Functions

So far we examined coordinate transformations. Now we will investigate the behaviour of fields. For instance, the $\mathcal{N} = 4$ *Super Yang-Mills theory* (SYM) mentioned in the introduction only contains fields transforming covariantly under the conformal group. In general QFTs (such as QED or QCD), conformal symmetry is generically broken by quantum effects (anomalies).

A necessary condition for a field theory to be conformally symmetric is a vanishing *β-function*. The latter describes the change of a coupling g with energy scales μ, i.e.

$$\beta(g) = \mu\,\frac{\partial g}{\partial \mu}, \tag{3.15}$$

so $\beta(g) = 0$ is a necessary condition scale invariance. Generically, $\beta(g) = 0$ occurs for specific values of g only, which correspond to the *renormalization group fixed points*. However, for $\mathcal{N} = 4$ Super Yang–Mills theory, it has been shown that the beta function vanishes for all values of the coupling g.

A conformally covariant operator $\mathscr{O}$ of a conformal field theory (CFT) transforms as follows under infinitesimal conformal transformations (with Killing vector v and $\sigma = \partial \cdot v/d$)

$$\delta_v \mathscr{O} = -\,(L_v\,\mathscr{O}), \quad L_v = v(x)\cdot\partial + \Delta\,\sigma(x) - \frac{1}{2}\,\partial_{[\mu} v_{\nu]}(x)\,S_{\mu\nu}\,. \tag{3.16}$$

Here, Δ denotes the *scaling dimension* of the operator $\mathscr{O}$ and $S_{\mu\nu}$ a generator of $O(d)$ in an appropriate representation. It only affects spinor, vector and tensor fields but no scalars φ

$$\delta_v \varphi = -\big(v(x)\cdot\partial + \Delta\,\sigma(x)\big)\,\varphi\,. \tag{3.17}$$

In this section, we work in Euclidean signature and follow again [13, 14]. In general QFTs, correlation functions are defined as time ordered vacuum expectation values, e.g. a two-point function of some field φ is given by

$$\langle\varphi(x)\,\varphi(y)\rangle := \langle 0|\,T\,\varphi(x)\,\varphi(y)\,|0\rangle, \tag{3.18}$$

three-, four- and higher point functions by analogous expressions. Generically, their computation is quite involved and possible only in the framework of perturbation theory.

Let us also give the path integral analogue of the definition (3.18) in the operator approach. In a scalar field theory governed by action $S[\varphi]$, the *partition function* Z and a general correlation function $\langle\mathscr{O}\rangle$ is defined by the path integrals

$$Z := \int \mathscr{D}\varphi\, \mathrm{e}^{-S[\varphi]}, \quad \langle\mathscr{O}\rangle := \frac{1}{Z}\int \mathscr{D}\varphi\,\mathscr{O}\,\mathrm{e}^{-S[\varphi]}\,. \tag{3.19}$$

In CFTs, conformal symmetry is so strong that it determines the form of the two- and three-point correlation functions up to a manageable number of parameters. In the notation $(x-y)^2 = (x-y)_\mu(x-y)^\mu$, the two- and three-point functions of scalars φ_i with scale dimensions Δ_i are given by

$$\langle\varphi_1(x)\,\varphi_2(y)\rangle := \frac{c\,\delta_{\Delta_1,\Delta_2}}{(x-y)^{2\Delta_1}}, \tag{3.20}$$

$$\langle\varphi_1(x)\,\varphi_2(y)\,\varphi_3(z)\rangle := \frac{k}{(x-y)^{\Delta_1+\Delta_2-\Delta_3}\,(y-z)^{-\Delta_1+\Delta_2+\Delta_3}\,(x-z)^{\Delta_1-\Delta_2+\Delta_3}} \tag{3.21}$$

with constants c,k determined by the field content.

Four-point correlators $\langle\varphi_1(x)\varphi_2(y)\varphi_3(z)\varphi_4(w)\rangle$ are less constrained by the symmetry since they involve dimensionless *cross ratios* $\frac{(x-y)^2}{(z-w)^2}$ and $\frac{(x-z)^2}{(y-w)^2}$.

3.2.1.3 The Energy Momentum Tensor in a CFT

The symmetric *energy-momentum tensor* $T_{\mu\nu}$ generates the Noether currents associated with conformal symmetry if the conservation law $\partial_\mu T^{\mu\nu} = 0$ (or rather $\nabla_\mu T^{\mu\nu} = 0$ in curved spacetime) is imposed. The infinitesimal transformations with conformal Killing vector v^μ gives rise to the conserved current

$$j^\mu = T^{\mu\nu}\,v_\nu\,. \tag{3.22}$$

In this subsection, we will show an important property of the energy momentum tensor in a conformal field theory, namely its tracelessness $T^\mu_\mu = 0$.

It is a common method in QFT to introduce sources for operators in a QFT's action, and then express the operator (in correlation functions) as the functional derivative of the generating functional. To do so, the action S_0 of our theory is modified by an additive term which couples the operator to its source. For instance consider some scalar operator φ and its source J,

$$S[\varphi, J] = S_0[\varphi] + \int d^d x\, \varphi(x) J(x). \tag{3.23}$$

Correlation function of that operator φ may now be calculated as the functional derivative of the generating functional $W[J] := -\ln Z[J]$ of the theory with respect to the source J, e.g.

$$\langle \varphi(x) \rangle \propto \frac{\delta W[J]}{\delta J(x)}. \tag{3.24}$$

One can also apply this procedure to vector and tensor operators,

$$S = S_0 + \int d^d x \left(\varphi J + V_\mu A^\mu + T_{\mu\nu} g^{\mu\nu} \right). \tag{3.25}$$

It can be shown that the source of the energy momentum tensor is exactly the object that has the properties of a metric. So the energy momentum tensor is obtained by calculating

$$T_{\mu\nu}(x) = -\frac{2}{\sqrt{|\det g|}} \frac{\delta W[g]}{\delta g^{\mu\nu}(x)}. \tag{3.26}$$

The metric transforms under conformal coordinate transformations induced by a vector field v as $\delta_v g^{\mu\nu} = 2\sigma g^{\mu\nu}$, so requiring invariance of W implies

$$\begin{aligned} 0 = \delta_v W[g] &= \int d^d x \frac{\delta W[g]}{\delta g^{\mu\nu}(x)}\, \delta_v g^{\mu\nu}(x) \\ &= \int d^d x \left(-\frac{\sqrt{|\det g|}\, T_{\mu\nu}}{2} \right) \cdot \left(2\, \sigma\, g^{\mu\nu} \right) \end{aligned} \tag{3.27}$$

$$= -\int d^d x \sqrt{|\det g|}\, T_\mu{}^\mu \cdot \sigma\, . \tag{3.28}$$

Since $T_\mu{}^\mu$ vanishes upon integration against an arbitrary function σ, one can conclude the announced tracelessness of the energy momentum tensor

$$T_\mu{}^\mu = 0\, . \tag{3.29}$$

3.2.2 $\mathcal{N} = 4$ Super Yang–Mills Theory

In this section we want to develop the field theory side of the AdS/CFT correspondence—the maximally supersymmetric $SU(N)$ gauge theory. This $\mathcal{N} = 4$ Super Yang–Mills theory is an example for a $d = 4$ dimensional CFT. In the following, the ingredients will be introduced step by step.

3.2.2.1 Non-Abelian Gauge Theories

Super Yang–Mills theory is a non-abelian gauge theory, i.e. its fields take values in the algebra of a non-abelian gauge group. QED, on the other hand, is associated with the abelian gauge group $U(1)$. Let us take it as an introductory example for the necessity of a gauge field.

Consider a complex scalar field $\varphi(x)$ transforming under local $U(1)$ transformations as

$$\varphi(x) \to e^{i\vartheta(x)}\,\varphi(x), \quad \partial_\mu\varphi(x) \to \partial_\mu\left(e^{i\vartheta(x)}\,\varphi(x)\right) \neq e^{i\vartheta(x)} \cdot \partial_\mu\varphi(x). \tag{3.30}$$

The derivative $\partial_\mu\varphi$ obviously does not transform like the field φ itself, so a connection A_μ is required in order to define a gauge covariant derivative:

$$D_\mu\varphi(x) := \left(\partial_\mu + iA_\mu\right)\varphi(x) \to e^{i\vartheta(x)} \cdot D_\mu\varphi(x) \Leftrightarrow A_\mu \to A_\mu - \partial_\mu\vartheta \tag{3.31}$$

With A_μ transforming like that, we can use the covariant derivative D_μ to construct gauge invariant objects (e.g. kinetic terms in the action). Furthermore, the field strength tensor

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu \tag{3.32}$$

is unaffected by gauge transformations of A_μ since $\partial_{[\mu}\partial_{\nu]}\vartheta = 0$.

The most important examples of non-abelian gauge groups in these lectures are $SU(N)$ with $N \geq 2$. One has to distinguish two transformation properties of fields under the non-abelian $SU(N)$:

- Fields transforming in the *fundamental representation* of the gauge group are elements of an N dimensional vector space

$$q_i(x) \to \left(e^{i\vartheta^a(x)T^a}\right)_i^{\ j} q_j(x), \quad i,j = 1,2,\ldots,N, \quad a = 1,\ldots,N^2-1\,. \tag{3.33}$$

  The $SU(N)$ generators T^a are traceless hermitian $N \times N$ matrices and ensure that $e^{i\vartheta^a T^a}$ is unitary. If the parameters $\vartheta^a(x)$ are infinitesimal, the field q_i is shifted by an algebra element

$$q_i(x) \to q_i(x) + i\vartheta^a(x)\,(T^a)_i^{\ j}\,q_j(x)\,. \tag{3.34}$$

- Fields transforming in the *adjoint representation* of the gauge group are aligned into the N^2-1 dimensional algebra $su(N)$,

$$\phi_i^{\ j} \equiv \phi^a\,(T^a)_i^{\ j} \to \left(e^{i\vartheta^b T^b}\right)_i^{\ k}\,\phi^a\,(T^a)_k^{\ l}\,\left(e^{-i\vartheta^c T^c}\right)_l^{\ j}. \tag{3.35}$$

  Infinitesimally, conjugation by a group element $e^{i\vartheta^a T^a}$ involves the commutator $[T^a, T^b] = if^{abc}T^c$ of the $su(N)$ generators

Fig. 3.1 A cartoon of the 3-point and 4-point vertices. The 3-point vertex is proportional to the coupling constant g, and the 4-point vertex to g^2

$$\begin{aligned}\phi^a T^a \to \phi^a T^a + i\left(\vartheta^b T^b \phi^a T^a - \phi^a T^a \vartheta^b T^b\right) \\ = \phi^a T^a - i\,\vartheta^b \phi^a \left[T^a\,,\,T^b\right] \\ = \phi^a T^a + f^{abc} \phi^a \vartheta^b T^c\,.\end{aligned} \tag{3.36}$$

Non-abelian gauge fields $A_\mu = A_\mu^a T^a$ give rise to a *non-abelian field strength tensor* in the adjoint representation

$$\begin{aligned}F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu + ig\left[A_\mu\,,\,A_\nu\right] \\ = \left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g f^{abc} A_\mu^b A_\nu^c\right) T^a.\end{aligned} \tag{3.37}$$

The transformation properties of $F_{\mu\nu}$ can be deduced from its alternative definition as a commutator of (non-abelian) gauge covariant derivatives (with g denoting the gauge coupling)

$$(D_\mu)_i{}^j := \delta_i{}^j\,\partial_\mu + ig A_\mu^a\,(T^a)_i{}^j, \quad F_{\mu\nu} = -\frac{i}{g}\left[D_\mu\,,\,D_\nu\right]. \tag{3.38}$$

One can thus form a gauge invariant action for the field strength by taking a trace over the i, j indices of the generators

$$S[A] \sim \int d^4x \mathrm{Tr}\left\{F^{\mu\nu} F_{\mu\nu}\right\}. \tag{3.39}$$

The non-linear contribution to $F_{\mu\nu}$ gives rise to interactions with the vertices (Fig. 3.1).

It turns out that in $\mathcal{N} = 4$ Super Yang–Mills theory, due to the large amount of supersymmetry all fields are in the *adjoint* representation of the gauge group. On the other hand, in QCD, the quark fields are in the *fundamental* representation of the gauge group. Given the success of the AdS/CFT correspondence, extensions of AdS/CFT have been considered where quark degrees of freedom are added to the original setup, in view of describing theories more similar to QCD in at least some respects. For a review of these extensions, see for instance [10].

3.2.2.2 The 1/*N* Expansion

As first pointed out by Gerard t'Hooft in 1974, $SU(N)$ gauge theories simplify considerably in the limit $N \to \infty$. This limit also enters gauge/gravity duality in a crucial way.

The large N limit is motivated by an expansion used in statistical mechanics, where the number of field components is taken to be large and an expansion in the inverse of this number is performed. The expansion of $SU(N)$ theory in $1/N$ rearranges the Feynman diagrams in such a way that they correspond to a string theory expansion with string coupling $1/N$. This suggests that $SU(N)$ gauge theories are equivalent to string theories, at least at large N. A particular virtue of the AdS/CFT correspondence is to make this mapping between field theory and string theory precise for a well-defined class of examples.

For illustrating the relation between a field theory expansion in the large N limit and a string theory expansion, let us consider a toy model originally advocated by Sidney Coleman [15]: This model involves a scalar field ϕ in the adjoint representation of the gauge group $SU(N)$, with $\phi = \phi^a T^a$. More explicitly, writing out the indices of the matrices T^a, we have

$$\phi_i^{\ j} \equiv \phi^a (T^a)_i^{\ j} \,. \tag{3.40}$$

Moreover it is assumed that the interaction vertices mimic Yang-Mills theory—a three point vertex $\sim g$ and a four point vertex $\sim g^2$. The toy model's Lagrangian then reads

$$\mathscr{L} \sim \mathrm{Tr}\big\{d\phi\, d\phi\big\} + g\,\mathrm{Tr}\big\{\phi^3\big\} + g^2\,\mathrm{Tr}\big\{\phi^4\big\} \,. \tag{3.41}$$

A rescaling $g\phi \mapsto \phi$ turns this into

$$\mathscr{L} \sim \frac{1}{g^2}\left(\mathrm{Tr}\big\{d\phi\, d\phi\big\} + \mathrm{Tr}\big\{\phi^3\big\} + \mathrm{Tr}\big\{\phi^4\big\}\right) . \tag{3.42}$$

To have a well-defined $N \to \infty$ limit, it is convenient to introduce the *t'Hooft coupling*

$$\lambda := {g_{\mathrm{YM}}}^2\, N. \tag{3.43}$$

If we send $N \to \infty$ at constant λ, the coefficient of (3.42) diverges but the number $N^2 - 1$ of components in the fields diverges as well. In fact, a subtle cancellation mechanism between the two infinities will take place. To see this at work, we analyze particular Feynman graphs in the t'Hooft limit. In the notation (3.40), the propagator of the field ϕ has the structure

$$\langle \phi_i^{\ j}(x)\, \phi_k^{\ l}(y) \rangle = \left(\delta_i^{\ l}\, \delta_k^{\ j} - \frac{1}{N}\, \delta_i^{\ j}\, \delta_k^{\ l} \right) \frac{1}{4\pi^2 (x-y)^2} \,, \tag{3.44}$$

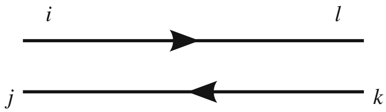


Fig. 3.2 The double line notation: fundamental and anti-fundamental fields are represented by directed lines with the color indices at both ends. An adjoint field may be seen as a direct product of a fundamental and an anti-fundamental field

which is found using the $SU(N)$ completeness relation

$$\sum_{a=1}^{N^2-1} (T^a)_i{}^j (T^a)_k{}^l = \frac{1}{2}\left(\delta_i{}^l\,\delta_k{}^j - \frac{1}{N}\,\delta_i{}^j\,\delta_k{}^l\right). \tag{3.45}$$

The space-time dependence of the propagator is the appropriate one for a scalar field in four dimensions. In the $N \to \infty$ limit, the term removing the trace, i.e.the second term in the bracket in (3.44), is subleading and may be ignored. Then the expression (3.44) for the propagator suggests a *double line notation* according to Fig. 3.2. Feynman diagrams then become networks of double lines. Vertices scale as $\frac{N}{\lambda}$, propagators as $\frac{\lambda}{N}$, and the sum over indices in a trace contributes a factor of N for each closed loop. If we introduce the shorthand notation (V,E,F) for the numbers of vertices, propagators (edges) and loops (faces) respectively, diagrams are proportional to

$$\text{diagram}(V,E,F) \sim N^{V-E+F}\,\lambda^{E-V} = N^{\chi}\,\lambda^{E-V}. \tag{3.46}$$

The power of the expansion parameter N is precisely the *Euler characteristic*

$$\chi := V - E + F = 2 - 2\,g, \tag{3.47}$$

related to the surface's number of handels (the *genus*) g. Any physical quantity in this theory is given by a perturbative expansion of type

$$\sum_{g=0}^{\infty} N^{2-2g} \sum_{i=0}^{\infty} c_{g,i}\,\lambda^i = \sum_{g=0}^{\infty} N^{2-2g} f_g(\lambda) \tag{3.48}$$

with $f_g(\lambda)$ a polynomial in the t'Hooft coupling. For large N, the series is clearly dominated by surfaces of minimal genus, the so-called *planar diagrams*. As an example let us compare the vacuum amplitudes shown in Fig. 3.3. In this simple toy model, it is not known which string theory fits to the perturbative series. For $\mathcal{N} = 4$ SYM, however, the AdS/CFT correspondence tells us which string theory leads to the correct expansion, namely to the ten-dimensional type IIB superstring theory on $\mathrm{AdS}_5 \times S^5$.

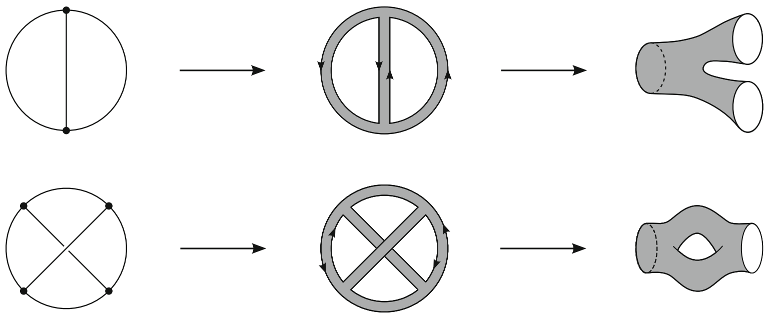


Fig. 3.3 The Feynman diagrams (*left*) can be translated to double line diagrams (*middle*), which in turn can be interpreted as Riemann surfaces of well defined topology (*shaded*). These surfaces (deformed to the shape on the *right*) can be interpreted as stringy Feynman diagrams. While the upper diagrams are planar ($\propto N^2$), the lower diagrams are non-planar of genus $g = 1$ ($\propto N^0$). The propagator and the interaction vertex of a closed string are depicted on the right pictures

3.2.2.3 Supersymmetry

We know to have Poincaré symmetry in the flat Minkowski spacetime, which is equipped with a "mostly positive" metric of signature $\eta = \mathrm{diag}(-,+,+,+)$. Generators of translations and Lorentz transformations will be denoted as P_μ and $L_{\mu\nu}$, respectively. Supersymmetry now enlarges the Poincaré algebra

$$\left[L_{\mu\nu}\,,\, P_\lambda\right] = -i\,(\eta_{\mu\lambda}\,P_\nu - \eta_{\nu\lambda}\,P_\mu)\,, \tag{3.49}$$

$$\left[L_{\mu\nu}\,,\, L_{\lambda\rho}\right] = -i\,(\eta_{\mu\lambda}\,L_{\nu\rho} - \eta_{\mu\rho}\,L_{\nu\lambda} + \eta_{\nu\rho}\,L_{\mu\lambda} - \eta_{\nu\lambda}\,L_{\mu\rho}) \tag{3.50}$$

by including *spinor supercharges* Q. In so-called Weyl notation we have a left-handed spinor Q^a_α and its right-handed counterpart $\bar{Q}_{a\dot{\alpha}} = (Q^a_\alpha)^\dagger$ where the $SL(2,\mathbb{C})$ indices $\alpha, \dot{\alpha}$ take values 1, 2 and a counts the number of independent supersymmetries $a = 1, \ldots, \mathcal{N}$. The Q's transform as Weyl spinors of $SO(1,3) \cong SL(2,\mathbb{C})/\mathbb{Z}_2$. The two-component Weyl spinor notation is related to the Dirac four-spinor notation by

$$Q^a_{\mathrm{D}} = \begin{pmatrix} Q^a_\alpha \\ \bar{Q}^{a\dot{\alpha}} \end{pmatrix}, \quad \gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu_{\alpha\dot{\beta}} \\ \bar{\sigma}^{\mu\dot{\alpha}\beta} & 0 \end{pmatrix}, \tag{3.51}$$

where $\sigma^\mu = (-\mathbb{1}, \sigma^i)$ and $\bar{\sigma}^\mu = (-\mathbb{1}, -\sigma^i)$ are four vectors of 2×2 matrices with the standard Pauli matrices σ^i as their spatial entries.

The supercharges commute with the generators of translations but otherwise obey the algebra

$$\left\{Q^a_\alpha\,,\,\bar{Q}_{b\dot{\beta}}\right\} = -\,2\,\sigma^\mu_{\alpha\dot{\beta}}\,P_\mu\,\delta^a_b, \quad \left\{Q^a_\alpha\,,\,Q^b_\beta\right\} = 2\,\varepsilon_{\alpha\beta}\,Z^{ab}. \tag{3.52}$$

Here the operators Z^{ab} are referred to as *central charges*. They commute with all the Poincaré- and supersymmetry generators Q^a and need to be antisymmetric $Z^{ab} = -Z^{ba}$ in order to respect the anticommutator's symmetry. Therefore, for $\mathcal{N} = 1$ supersymmetry, we have $Z = 0$.

The supersymmetry algebra (3.52) is invariant under global phase rotations of the supercharges $Q^a_{1,2}$ into each other. This forms an *R symmetry group* denoted as $U(1)_\mathrm{R}$. In addition, when $\mathcal{N} > 1$, the different supercharges may be rotated into one another under the unitary group $SU(N)_\mathrm{R}$ which extends the R symmetry.

The field theory in the AdS/CFT dictionary has $\mathcal{N} = 4$ supersymmetries. Let us briefly explain why this is the maximal supersymmetry for a pure gauge theory without gravity: Each supercharge $Q^a_\alpha, \bar{Q}_{a\dot{\alpha}}$ changes the spin of the state it acts on by $1/2$. In absence of gravity, helicities between -1 and $+1$ occur, hence no spin modification greater than $2 = \mathcal{N}_{\max} \cdot 1/2$ is allowed.

In the $\mathcal{N} = 4$ theory we have R symmetry $SU(4) \cong SO(6)$. This is exactly the isometry group of the sphere in the $\mathrm{AdS}_5 \times S^5$ background on the string theory side of the correspondence. The AdS_5 factor has the symmetries encoded by $SO(4, 2)$ in Minkowski space or $SO(5, 1)$ in an Euclidean formulation. These groups are isomorphic to the conformal group in $d = 4$ dimensions according to our analysis in Sect. 3.2.1.1.

3.2.2.4 Field Content of $\mathcal{N} = 4$ Supersymmetric Field Theory

Representations of the supersymmetry algebra make up the SUSY multiplets. Their components are spin 1 vector fields, spin $\frac{1}{2}$ fermion fields and spin 0 scalar fields. In $\mathcal{N} = 4$ supersymmetry we encounter *maximal supersymmetry* if $s = 1$ is the highest spin in a SUSY-multiplet. This implies that we cannot describe gravity with this theory, because the graviton is supposed to have spin 2.

For any $\mathcal{N}$ with $1 \le \mathcal{N} \le 4$ we encounter one gauge multiplet, which is a multiplet transforming in the adjoint representation of the gauge group (while we are used to have matter fields in the fundamental representation in non-supersymmetric theories). For $\mathcal{N} = 4$ this is the only possible multiplet.

Lower symmetry $\mathcal{N} = 1$ and $\mathcal{N} = 2$ also admits matter multiplets which we will not discuss here, though. (But to make you familiar with the names, the multiplet in the fundamental representation in $\mathcal{N} = 1$ SUSY is called *chiral multiplet*, and the multiplet in the fundamental representation in $\mathcal{N} = 2$ SUSY is called the *hypermultiplet*). The content of the $\mathcal{N} = 4$ multiplet is given in Table 3.1. Note that this theory is non-chiral. The Lagrangian may be written as

Table 3.1 The field content of the $\mathcal{N} = 4$ supersymmetry multiplet and the representation in which these fields transform with respect to the R symmetry group $SU(4)_{\mathrm{R}} \cong SO(6)_{\mathrm{R}}$

	Field	Range	Representation of $SU(4)_{\mathrm{R}}$
Vector	A_μ		$(\underline{1})$ singlet
Weyl fermions	λ^a_α, $\bar{\lambda}^a_{\dot\alpha}$,	$a = 1, 2, 3, 4$	$(\underline{4})$ fundamental
Real scalars	X^i,	$i = 1, 2, \ldots, 6$	$(\underline{6})$ adjoint

$$\begin{aligned}\mathcal{L} = \operatorname{Tr}\Big\{ & -\frac{1}{2g^2} F_{\mu\nu} F^{\mu\nu} + \frac{\theta^I}{8\pi^2} F_{\mu\nu} \tilde{F}^{\mu\nu} - i \sum_a \bar{\lambda}^a \bar{\sigma}^\mu D_\mu \lambda_a \\ & - \sum_i D_\mu X^i D^\mu X^i + g \sum_{a,b,i} C_i^{ab} \lambda_a \left[X^i, \lambda_b\right] \\ & + g \sum_{a,b,i} \bar{C}_{iab} \bar{\lambda}^a \left[X^i, \bar{\lambda}^b\right] + \frac{g^2}{2} \sum_{i,j} [X^i, X^j]^2 \Big\}. \end{aligned} \tag{3.53}$$

Here the trace is summing over gauge indices $\tilde{\alpha}, \tilde{\beta}$ which are suppressed in the expression above. They appear if we rewrite the adjoint fields correctly as linear combinations of the generators T^A of the gauge group, e.g. $\left(X^i\right)_{\tilde{\alpha}}^{\tilde{\beta}} = X^{iA} T^A_{\tilde{\alpha}} \tilde{\beta}$. The symbol θ^I denotes the instanton number and $\tilde{F}_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\lambda\rho} F^{\lambda\rho}$.

The C_i^{ab} are the structure constants of $SU(4)_{\mathrm{R}}$. Note that there is only one coupling constant g. On the classical level this theory is conformal with engineering dimensions of the fields as $[A_\mu] = 1$, $[\lambda] = 3/2$, $[X] = 1$ and therefore $[g] = 0$. The dimensionless coupling and absence of any mass term are necessary for conformal invariance.

The Lagrangian (3.53) is invariant under SUSY-transformations given by

$$\begin{aligned} (\delta X^i)^a{}_\alpha &= \left[Q^a{}_\alpha, X^i\right] = C^{iab} \lambda_{\alpha b}, \\ (\delta \lambda_{\beta b})^a{}_\alpha &= \left\{Q^a{}_\alpha, \lambda_{\beta b}\right\} = F^+_{\mu\nu} (\sigma^{\mu\nu} \varepsilon)_{\alpha\beta} \delta^a{}_b + \left[X^i, X^j\right] \varepsilon_{\alpha\beta} (C_{ij})^a{}_b, \\ (\delta \bar{\lambda}^b_{\dot\beta})^a{}_\alpha &= \left\{Q^a{}_\alpha, \bar{\lambda}^b_{\dot\beta}\right\} = C_i^{ab} \sigma^\mu_{\alpha\dot\beta} D_\mu X^i, \\ (\delta A^\mu)^a{}_\alpha &= \left[Q^a{}_\alpha, A^\mu\right] = \sigma^\mu_{\alpha\dot\beta} \bar{\lambda}^{\dot\beta a}. \end{aligned} \tag{3.54}$$

Note that $F^+_{\mu\nu}$ is the self-dual part $\frac{1}{2}(F_{\mu\nu} + \tilde{F}_{\mu\nu})$ of the field strength, and the constants $(C_{ij})^a_b$ are related to bilinears in Clifford Dirac matrices of $SO(6)_{\mathrm{R}}$. Upon quantization of this theory, one finds that the β-function vanishes to all orders of pertubation theory (and even non-perturbatively), therefore we are left with a CFT even at quantum level.

3.2.2.5 The Superconformal Algebra and Its Representations

The concept of supersymmetry together with the conformal group form the *superconformal group* $SU(2,2|4)$. The $SU(2,2)$ part represents the symmetry of the Weyl spinors while the $SU(4)$ refers to the R symmetry group $SU(4)_{\mathrm{R}}$ of the $\mathcal{N}=4$ supersymmetry.

The AdS/CFT map will provide a direct one to one mapping between operators on both sides of the correspondence. This relies heavily on the fact that on both sides the operators fall into representations of the same symmetry groups. The generators of the superconformal group are given by

- Conformal symmetry with generators $P_\mu, L_{\mu\nu}, D, K_\mu$: In addition to the Poincaré algebra (3.49) and (3.50), the conformal algebra involves commutators

$$\begin{aligned} \left[D\,,\,P_\mu\right] &= -i\,P_\mu\,,\\ \left[D\,,\,K_\mu\right] &= i\,K_\mu\,,\\ \left[L_{\mu\nu}\,,\,D\right] &= 0\,,\\ \left[L_{\mu\nu}\,,\,K_\rho\right] &= -i\,(\eta_{\mu\rho}\,K_\nu - \eta_{\nu\rho}\,K_\mu)\,,\\ \left[P_\mu\,,\,K_\nu\right] &= 2i\,L_{\mu\nu} - 2i\,\eta_{\mu\nu}\,D\,. \end{aligned} \tag{3.55}$$

- R symmetry $SO(6)_{\mathrm{R}} \cong SU(4)_{\mathrm{R}}$ with generators $T^A, A = 1,2,\ldots,15$. The $SO(4,2)$ and $SU(4)_{\mathrm{R}}$ subgroups commute.
- Poincaré supersymmetry with generators Q^a_α, $\bar{Q}_{\dot{\alpha}a}$, $a=1,2,3,4$ subject to (3.52).
- Conformal supersymmetry generators $S_{\alpha a}$ and $\bar{S}^{\dot{\alpha}a}$ which introduce the following anticommutation relations

$$\begin{aligned} \left\{Q^a_\alpha\,,\,Q^b_\beta\right\} = \left\{S_{\alpha a}\,,\,S_{\beta b}\right\} &= \left\{Q^a_\alpha\,,\,\bar{S}^b_{\dot\beta}\right\} = 0\,,\\ \left\{Q^a_\alpha\,,\,\bar{Q}_{\dot\beta b}\right\} &= 2\,(\sigma^\mu)_{\alpha\dot\beta}\,P_\mu\,\delta^a_b\,,\\ \left\{S^a_\alpha\,,\,\bar{S}_{\dot\beta b}\right\} &= 2\,(\sigma^\mu)_{\alpha\dot\beta}\,K_\mu\,\delta^a_b\,,\\ \left\{Q^a_\alpha\,,\,S_{\beta b}\right\} &= \varepsilon_{\alpha\beta}\,\delta^a_b\,D + \frac{1}{2}\,\delta^a_b\,L_{\mu\nu}\,(\sigma^{\mu\nu}\,\varepsilon)_{\alpha\beta}\,. \end{aligned} \tag{3.56}$$

Central charges are assumed to vanish throughout these lectures.

The fields $A_\mu(x)$, $\lambda^a_\alpha(x)$, $\bar{\lambda}^a_{\dot\alpha}(x)$ and $X^i(x)$ of the SUSY multiplet ($a=1,2,3,4$ and $i=1,2,\ldots,6$) can be used to construct composite operators of $\mathcal{N}=4$ SYM. A regularization prescription is needed when multiplying fields at the same spacetime point.

We define a *superconformal primary operator* $\mathcal{O}$ by

$$\left[S\,,\,\mathcal{O}\right] = 0, \tag{3.57}$$

i.e. the $\mathcal{O}$'s are the lowest dimensional operators in a representation of $SU(2,2|4)$. This is the generalization of the primary operator condition $[K_\mu, \mathcal{O}] = 0$ in bosonic conformal field theory (which is in fact implied by (3.57) since two S generators anticommute to K's).

An operator $\mathcal{O}'$ is a *superconformal descendant* of $\mathcal{O}$ if

$$\mathcal{O}' = \left[Q\,,\,\mathcal{O}\right], \tag{3.58}$$

$\mathcal{O}$ and $\mathcal{O}'$ then belong to the same superconformal multiplet, i.e. the same representation of $SU(2,2|4)$. The scale dimension is shifted as $\Delta_{\mathcal{O}'} = \Delta_{\mathcal{O}} + \frac{1}{2}$.

Of central importance are single trace operators (taking a trace is necessary to ensure gauge invariance)

$$\mathcal{O} = \mathrm{Tr}\left\{X^{(i_1} X^{i_2} \ldots X^{i_n)}\right\} = \mathrm{sTr}\left\{X^{i_1} X^{i_2} \ldots X^{i_n}\right\}. \tag{3.59}$$

They are also referred to as *half BPS operators* since they are annihilated by half the spinorial generators S (but not by the other half Q).

3.2.3 Anti-de Sitter Space

In this section we will examine Anti-de Sitter spacetime and compare it to flat Minkowski spacetime. As mentioned earlier, one side of the AdS/CFT correspondence is so-called type IIB superstring theory formulated on the spacetime $\mathrm{AdS}_5 \times S^5$. We will not discuss string theory now. Instead we want to get familiar with the spacetime and see how it may be connected to the more familiar Minkowski spacetime $\mathbb{R}^{1,3}$.

The most important facts about $\mathrm{AdS}_5 \times S^5$ spacetime for us are of geometrical nature. We already stated that the isometry group of this spacetime is the same as the symmetry group of the quantum field theory on the other side of the correspondence.

The key result of this section will be that the boundary of the Euclidean compactification of AdS_5 spacetime is equal to compactified $\mathbb{R}^4$, which is the Euclidean compactification of the Minkowski spacetime we live in. To see this equivalence we will make use of so called *conformal diagrams* which enable us to draw an image of the entire spacetime on a single sheet of paper making the causal structure of the spacetime visible. A short introduction to conformal diagrams is for example given in appendix H of [1].

The $(p+2)$-dimensional version AdS_{p+2} of this spacetime can be defined as the embedding of a hyperboloid (with *AdS radius* L)

$$X_0^2 + X_{p+2}^2 - \sum_{i=1}^{p+1} X_i^2 = L^2 \tag{3.60}$$

into a flat $(p+3)$-dimensional space $\mathbb{R}^{p+3}$ with metric

$$ds^2 = -\,dX_0^2 - dX_{p+2}^2 + \sum_{i=1}^{p+1} dX_i^2. \tag{3.61}$$

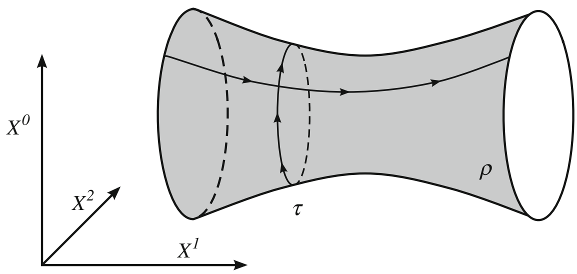


Fig. 3.4 AdS_3 spacetime as a hyperboloid with closed timelike curves

The AdS radius is a measure for the constant curvature. Riemann tensor and cosmological constant are given by

$$R_{\mu\nu\lambda\rho} = -\frac{1}{L^2}\left(g_{\mu\lambda}\, g_{\nu\rho} - g_{\mu\rho}\, g_{\nu\lambda}\right), \quad \Lambda = -\frac{d\,(d-1)}{L^2} < 0\,, \tag{3.62}$$

where d is the dimension of the boundary.

One possible parametrization of this spacetime is given by

$$\begin{aligned} X_0 &= L\,\cosh\rho\,\cos\tau\,, \\ X_{p+2} &= L\,\cosh\rho\,\sin\tau\,, \\ X_i &= L\,\Omega_i\,\sinh\rho\,, \end{aligned} \tag{3.63}$$

where Ω_i are the angular coordinates with $i = 1, \ldots, p+1$ such that $\sum_i \Omega_i^2 = 1$ and the remaining coordinates take the ranges $0 \le \rho, 0 \le \tau < 2\pi$.

Inserting (3.63) into (3.61) yields the metric

$$ds^2 = L^2\left(-\cosh^2\rho\, d\tau^2 + d\rho^2 + \sinh^2\rho d\Omega_p^2\right). \tag{3.64}$$

It features a timelike killing vector ∂_τ on the whole manifold, so τ may be called the global time coordinate. The isometry group $SO(2, p+1)$ of AdS_{p+2} has a maximal compact subgroup $SO(2) \times SO(p+1)$, the former generating translations in τ, the latter rotating the X_i's (Fig. 3.4).

Near $\rho = 0$ we have $\cosh\rho \approx 1$ and $\sinh\rho \approx \rho$, so in this environment the metric of AdS_5 looks like

$$ds^2 \approx L^2\left(-d\tau^2 + d\rho^2 + \rho^2\, d\Omega_3^2\right) \tag{3.65}$$

and thus is seen to be topologically $S^1 \times \mathbb{R}^4$. The S^1 parametrized by the time coordinate τ represents closed timelike curves. To prevent inconsistencies concerning causality, AdS_5 is therefore regarded as the causal spacetime obtained by unwrapping these circles, taking $-\infty < \tau < \infty$ without any identification.

Introducing a new coordinate θ, the metric (3.64) becomes that of the *Einstein static universe* $\mathbb{R} \times S^p$

$$\tan\theta = \sinh\rho \Rightarrow ds^2 = \frac{L^2}{\cos^2\theta}\left(-d\tau^2 + d\theta^2 + \sin^2\theta\, d\Omega_3^2\right). \tag{3.66}$$

However, since $0 \leq \theta < \frac{\pi}{2}$, this metric covers only half of $\mathbb{R} \times S^p$. The causal structure remains unchanged when scaling this metric to get rid of the overall factor. Further, adding the point $\theta = \frac{\pi}{2}$ corresponding to spatial infinity results in the compactified spacetime

$$ds^2 = -d\tau^2 + d\theta^2 + \sin^2\theta\, d\Omega_3^2, \quad 0 \leq \theta \leq \frac{\pi}{2}, \quad -\infty < \tau < \infty. \tag{3.67}$$

If we specify boundary conditions on $\mathbb{R} \times S^p$ at $\theta = \frac{\pi}{2}$, then the Cauchy problem is well-posed. As one can easily read off from (3.67), the $\theta = \frac{\pi}{2}$ boundary of conformally compactified AdS_{p+2} is identical to the conformal compactification of $(p+1)$ dimensional Minkowski spacetime.

Let us take a quick look at the special case of conformally compactified $(1+1)$ dimensional Minkowski spacetime. It is convenient to introduce light cone coordinates,

$$u_\pm := t \pm x \Rightarrow ds^2 = -dt^2 + dx^2 = -du_+\, du_-. \tag{3.68}$$

If we furthermore restrict the coordinates to a finite range, a useful choice is

$$u_\pm =: \tan\tilde{u}_\pm, \quad \tilde{u}_\pm =: \frac{\tau \pm \vartheta}{2} \Rightarrow ds^2 = \frac{-d\tau^2 + d\vartheta^2}{4\cos^2\tilde{u}_+ \cos^2\tilde{u}_-}. \tag{3.69}$$

Another neat parametrization of AdS_{p+2} is realized by the *Poincaré coordinates* which cover half of the hyperboloid. Introduce $(y, t, \mathbf{x})$ such that $y > 0$ and $\mathbf{x} \in \mathbb{R}^p$, then

$$\begin{aligned} X_0 &= \frac{1}{2y}\left(1 + y^2\left(L^2 + \mathbf{x}^2 - t^2\right)\right), \\ X_{p+1} &= \frac{1}{2y}\left(1 - y^2\left(L^2 - \mathbf{x}^2 + t^2\right)\right), \\ X_{p+2} &= L\, y\, t, \\ X_i &= L\, y\, x_i. \end{aligned} \tag{3.70}$$

The boundary at $y \to \infty$ can be better analyzed in terms of a new variable u

$$u := \frac{1}{y} \Rightarrow ds^2 = L^2\left(\frac{du^2}{u^2} + \frac{1}{u^2}\,\eta_{ij}\, dx^i\, dx^j\right). \tag{3.71}$$

After a conformal rescaling by u^2, we obtain the Minkowski metric by freezing $u = 0$.

3.2.4 D-branes

D-branes (where the 'D' stands for Dirichlet boundary conditions) are a central feature of string theory in view of the second string revolution 1995. They play a central role for the AdS/CFT correspondence: In fact, D-branes have two different interpretations within string theory and the origin of the AdS/CFT correspondence is to identify the two D-brane interpretations, together with applying a subtle limit.

D-branes have an open string and a closed string interpretation. Let us sketch these briefly in the subsequent.

3.2.4.1 Open String Interpretation

Open strings require boundary conditions for their endpoints. These may be provided at particular hyperplanes, which define the D-branes. Thus D-branes are hyperplanes on which open strings may end. Just as fundamental strings, D-branes can couple to background fields, in particular to gravity. A world-volume action describing their dynamics may be obtained as a generalization of the worldsheet action for strings. The background fields act as generalized couplings.

Let us briefly recapitulate the boundary conditions for an open string. A string is defined on the worldsheet given by the two coordinates (σ, τ). In superstring theory, the fermionic contribution to the worldsheet action reads

$$S_f = \frac{i}{2\pi\alpha'} \int d^2\sigma \left(\psi_-^\mu\, \partial_+ \psi_{-\mu} + \psi_+^\mu\, \partial_- \psi_{+\mu}\right), \tag{3.72}$$

with $\psi_\pm$ the right and left movers, respectively. For the boundary conditions, we impose $\psi_+^\mu(\tau, 0) = \psi_-^\mu(\tau, 0)$ at $\sigma = 0$. Then, the boundary condition at $\sigma = \pi$ leaves two options corresponding to the *Neveu–Schwarz*- (NS-) and the *Ramond sector* (R) of the theory:

$$\begin{aligned} \mathrm{R}: \psi_+^\mu(\tau, \pi) &= +\,\psi_-^\mu(\tau, \pi) \\ \mathrm{NS}: \psi_+^\mu(\tau, \pi) &= -\,\psi_-^\mu(\tau, \pi)\,. \end{aligned} \tag{3.73}$$

Let us now turn to the D-branes. Let ξ^a denote the coordinates for the world-volume of a Dp brane (which reduces to $\xi^0 = \tau$ and $\xi^1 = \sigma$ in case of the fundamental string). Here, 'Dp' denotes a brane in one temporal and p spatial directions.

In direct analogy to the string worldsheet area action, the bosonic part of the D-brane action is given by

$$S_{\mathrm{DBI}}^{(p)} = -\,\mu_p \int d^{p+1}\xi\, e^{-\varphi} \sqrt{\det\left(g_{ab}^* + B_{ab}^* + 2\pi\,\alpha'\, F_{ab}\right)}. \tag{3.74}$$

The action (3.74) is known as *Dirac–Born–Infeld action*, or in short, *DBI action*. Its prefactor $\mu_p = (2\pi)^{-p}\alpha'^{-(p+1)/2}$, with $\alpha' \equiv {l_s}^2$ the inverse string tension and l_s the string length, relates to the (genuinely non-perturbative) brane tension $T_p = \mu_p/g_s$.

Moreover, g^* is the induced metric on the brane obtained via pullback of the spacetime metric to the brane worldvolume,

$$g^*_{ab} = \frac{\partial X^\mu}{\partial \xi^a} \frac{\partial X^\nu}{\partial \xi^b} g_{\mu\nu}. \tag{3.75}$$

The same applies to the antisymmetric two-form B in (3.74).

Expanding the DBI action in flat spacetime (with $g^*_{ab} = \eta_{ab}$) by means of $\det(1 + M) = 1 - \frac{1}{4}\mathrm{Tr}\{M^2\}$ for antisymmetric matrices M, we see that the DBI action for D3 branes is a generalization of Yang–Mills theory,

$$S^{(p=3)}_{\mathrm{DBI}} \sim \alpha'^{-2} \int d^4\xi \,\mathrm{Tr}\big\{\mathscr{F}_{ab}\mathscr{F}^{ab}\big\}, \quad \mathscr{F}_{ab} = B^*_{ab} + 2\pi\,\alpha'\,F_{ab}. \tag{3.76}$$

D branes also carry charge under Ramond–Ramond p-form fields C_p. The full action describing a charged *BPS brane* (named after Bogomolnyi, Prasad and Sommerfield) involves a *Chern–Simons term*, $S = S_{\mathrm{DBI}} \pm S_{\mathrm{CS}}$,

$$S_{\mathrm{CS}} = \mu_p \int d^{p+1}\xi \sum_q C^*_{q+1} \wedge \mathrm{Tr}\big\{e^{\mathscr{F}}\big\}. \tag{3.77}$$

The exponential of the two form $\mathscr{F}$ has to be understood in terms of the wedge product.

BPS branes are stable due to charge conservation. In type IIA/B superstring theory, Dp branes with p even/odd are BPS stable since Ramond–Ramond gauge potentials C_{p+1} are present to which Dp branes can couple. Unlike fundamental strings, D branes are non-perturbative objects since the tension and therefore their energy scales as $1/g_s$, i.e. with the inverse string coupling.

For N coincident D-branes, the gauge group is enhanced to $U(N)$, which may be written as the semidirect product $SU(N) \times U(1)$. For N coincident D3-branes, the action reduces to $\mathscr{N} = 4 U(N)$ Super-Yang–Mills theory in the limit $\alpha' \to 0$: For D3-branes, there is a $SO(6)$ rotational symmetry in the six spatial directions perpendicular to the branes. $SO(6)$ is isomorphic to $SU(4)$, which coincides with the R symmetry group of $\mathscr{N} = 4$ Super Yang–Mills theory.

3.2.4.2 Closed String Interpretation

On the other hand, D-branes also arise as solitonic solutions to the supergravity equations of motion. Let us discuss some aspects of this in detail.

Closed string excitations contain gravity. In particular, the graviton corresponds to the quadrupole fluctuation of the closed string. The closed sector of superstring theory can be constructed in four different ways. Each of left- and right movers may be taken from open string NS or R sectors. From the spacetime point of view, we find the following statistics for the states: The NS-NS and R-R sectors correspond to spacetime bosons, whereas the NS-R and R-NS sectors correspond to spacetime fermions. The NS-NS sector contains the fields $g_{\mu\nu}, B_{\mu\nu}, \varphi$ which we had already discussed in bosonic string theory whereas the 'mixed' NS-R, R-NS sectors contain SUSY superpartners such as gravitino and dilatino.

It may be shown that to leading order in α' (i.e. at low energies when only massless excitations contribute), Weyl invariance of the string worldsheet action in curved background is equivalent to certain field equations which can be derived from a gravity action. In superstring theory, this effective target space action is precisely that of *supergravity*. For this reason, the supergravity theories are referred to as type IIA/B although they can be motivated independent of string theory.

In type IIB supergravity, the bosonic field consists of the massless closed string states, $g_{\mu\nu}, B_{\mu\nu}$ and φ from the NS-NS sector and the R-R form fields C_0, C_2 and C_4. In addition, there are fermions with an equal number of degrees of freedom as in the bosonic part.

Moreover, we define the *axio-dilaton* τ and a complex 3 form G_3 by

$$\tau := C_0 + i\, e^{-\varphi}, \quad G_3 := F_3 - \tau\, H_3 \tag{3.78}$$

where F_3, H_3 are the field strengths of C_2 and B_2 (in differential form notation $F_3 = dC_2$ and $H_3 = dB_2$). The C_4 potential is more conveniently represented by the field strength

$$\tilde{F}_5 = dC_4 + \frac{1}{2}\, B_2 \wedge F_3 - \frac{1}{2}\, C_2 \wedge H_3. \tag{3.79}$$

Let us finally introduce the rescalings $\tilde{g}_{\mu\nu} = e^{(\varphi_0 - \varphi)/6}$ and $\kappa = \kappa_0 e^{\varphi_0} = \sqrt{8\pi G_{\mathrm{N}}}$ into the *Einstein frame*, then the type IIB supergravity action is given by

$$S_{\mathrm{IIB}} = \frac{1}{2\,\kappa^2} \int d^{10}x \sqrt{-\tilde{g}} \left(\mathscr{R}_{\tilde{g}} - \frac{|\partial_\mu \tau|^2}{2\,(\mathrm{Im}\tau)^2} - \frac{|G_3|^2}{12\,\mathrm{Im}\tau} - \frac{|\tilde{F}_5|^2}{4 \cdot 5!} \right) + \frac{1}{8i\,\kappa^2} \int \frac{C_4 \wedge G_3 \wedge \bar{G}_3}{\mathrm{Im}\tau}. \tag{3.80}$$

The field strength $\tilde{F}_5$ has to be self-dual in the sense that

$$(\star F)_{\mu_1 \dots \mu_5} = F_{\mu_1 \dots \mu_5} \tag{3.81}$$

where the Hodge dual $(\star\omega)_k$ of a k form ω in D dimensions is defined by

$$(\star\omega)_{\mu_1 \dots \mu_{D-k}} = \frac{|\det g|}{k!}\, \varepsilon_{\nu_1 \dots \nu_k \mu_1 \dots \mu_{D-k}}\, \omega^{\nu_1 \dots \nu_k}, \tag{3.82}$$

e.g. $\star F_{\mu\nu} = \frac{|\det g|}{2} \varepsilon_{\mu\nu\lambda\rho} F^{\lambda\rho}$ in $D = 4$ dimensions.

Now let us look for solitonic solutions of the equations of motion due to (3.80). A Dp brane is a BPS solution of 10 dimensional supergravity, i.e. it is annihilated by half the Poincaré supercharges Q_α. It has a $p + 1$ dimensional flat hypersurface with Poincaré invariance group $\mathbb{R}^{p+1} \times SO(1,p)$. The transverse space is then of dimension $D-p-1$.

A p brane in 10 dimensions has symmetries $\mathbb{R}^{p+1} \times SO(1,p) \times SO(9-p)$. An ansatz which solves the equations of motion of type IIB supergravity is

$$ds^2 = \frac{1}{\sqrt{H(\mathbf{y})}}\, dx^\mu\, dx_\mu + \sqrt{H(\mathbf{y})}\, d\mathbf{y} \cdot d\mathbf{y} \tag{3.83}$$

where x^μ are the coordinates on the brane world volume and $\mathbf{y}$ denote the coordinates perpendicular to the brane. It turns out by means of the supergravity equations of motion that

$$e^{\varphi(\mathbf{y})} = \left[H(\mathbf{y})\right]^{\frac{3-p}{4}}, \quad H \equiv \text{harmonic function of} y = \sqrt{\mathbf{y}\cdot\mathbf{y}}. \tag{3.84}$$

Far away from the brane, i.e. at $y \to \infty$, flat space has to be recovered. This boundary condition uniquely fixes H to be

$$H(\mathbf{y}) = 1 + \left(\frac{L}{y}\right)^{D-p-3}. \tag{3.85}$$

L is a length scale related to the only dimensionful parameter α'. For a stack of N coincident Dp branes, we have

$$L^{D-p-3} = N\, g_s\, (4\pi)^{(5-p)/2}\, \Gamma\left(\frac{7-p}{2}\right)\, \alpha'^{(D-p-3)/2}. \tag{3.86}$$

For D3-branes this reduces to

$$L^4 = 4\pi\, g_\mathrm{s}\, N\, \alpha'^2 = 4\pi\, \lambda\, \alpha'^2\,, \tag{3.87}$$

where $g^2_{\mathrm{YM}} = g_s$ and $\lambda = g^2_{\mathrm{YM}} N = g_\mathrm{s} N$ is the t' Hooft coupling.

3.3 The AdS/CFT Correspondence

3.3.1 General Idea

At the level of string theory, the AdS/CFT correspondence is a duality between the open and closed string interpretations of D-branes. Thus, in the case of D3-branes, it is conjectured that the theory of open strings ending on a stack of N D3-branes, which for small inverse string tension α' corresponds to $\mathcal{N} = 4\, SU(N)$ Super Yang-Mills theory, is mapped to superstring theory on the space $AdS_5 \times S^5$, which is a full theory of quantum gravity. This is an exciting proposal, however so far it has not been possible to formulate string theory non-perturbatively on a curved space background, so it is not possible to test this proposal. Nevertheless, in his original

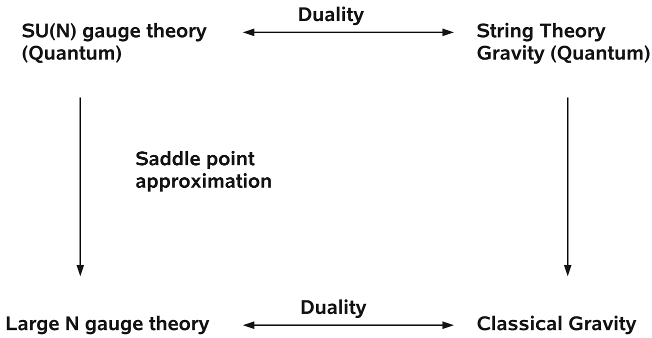


Fig. 3.5 General structure of the AdS/CFT correspondence and the limit involved

paper [1], Maldacena has proposed a subtle low-energy limit, in which the quantum gravity theory of closed strings reduces to classical supergravity. On the open string side, this limit corresponds essentially to a saddle-point approximation in which the planar limit $N \to \infty$ of the $SU(N)$ gauge theory is taken. In addition, this limit implies that the gauge theory is strongly coupled while the classical supergravity theory is weakly coupled. The general idea of AdS/CFT and the limit involved is summarized in Fig. 3.5.

3.3.2 Maldacena's Original Argument

Let us now consider the scenario outlined above in more detail, where we follow Maldacena's original argument as presented in [5]. We consider type IIB string theory in 9+1 dimensional spacetime with a stack of N D3-branes. There are two kinds of excitations, open and closed strings. The closed string fluctuations are excitations of the vacuum with the graviton as massless mode. Secondly, there are open string modes which describe excitations of the D3-branes. At energies below the string mass scale $(\alpha')^{-1/2}$, only massless string states are excited: The massless closed string states give rise to the gravity multiplet of type IIB supergravity, while the massless open string states give rise to the $\mathcal{N} = 4$ gauge multiplet and $\mathcal{N} = 4\, SU(N)$ Super Yang–Mills theory.

Let us recapitulate the open and closed string interpretations of D3 branes as introduced in the preceding section.

3.3.2.1 D3-branes From the Open String Point of View

The low energy effective action for the massless excitations of N D3-branes in flat ten dimensional space has the schematic form

$$
\begin{aligned}
S &= S_{\text{bulk}} + S_{\text{brane}} + S_{\text{int}}\,, \text{ where}\\
S_{\text{bulk}} &\equiv D = 10 \text{ supergravity incl. higher derivative terms, i.e. } \alpha' \text{ corrections}\\
S_{\text{brane}} &\equiv \text{DBI and CS action defined on 3+1 dimensional brane world volume:}\\
&\quad \text{for small } \alpha'\text{, we get SYM} \sim \operatorname{Tr}\{F_{\mu\nu}F^{\mu\nu}\} \text{ plus interactions}\\
&\quad \sim \alpha' \operatorname{Tr}\{F^4\} + \ldots\\
S_{\text{int}} &\equiv \text{bulk}-\text{brane interaction : the leading term comes from the background}\\
&\quad \text{metric } g \text{ in the brane action}\,.
\end{aligned}
$$

For $\alpha' \to 0$, the bulk action becomes the Einstein–Hilbert action with coupling $\kappa \sim g_s\alpha'^2$. In the expansion $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$ about flat space (with Minkowski metric η), the leading terms are

$$
S_{\text{bulk}} = \frac{1}{2\kappa^2}\int d^{10}x\sqrt{|g|}\,\mathscr{R}_g \sim \int d^{10}x\Big((\partial h)^2 + \kappa\,(\partial h)^2\,h + \cdots\Big). \tag{3.88}
$$

In the low-energy limit $\kappa \sim g_s\alpha'^2 \to 0$, the interaction terms $\mathscr{O}(\kappa)$ drop out, so gravity becomes free at long distances. Similar behaviour can be observed in the S_{int} sector. The term 'low energy limit' means that the relevant energies E are kept fixed while we send the dimensionful parameter $\alpha' \to 0$, therefore various dimensionless quantities such as $\alpha' E^2$ are suppressed.

In the low-energy limit, we are thus left with two decoupled theories: Supergravity of massless particles in the bulk and $\mathscr{N} = 4 SU(N)$ Super Yang–Mills theory on the brane.

3.3.2.2 D3-Branes from the Closed String Point of View and the Maldacena Limit

In their solitonic interpretation, D branes are viewed as massive charged objects which act as sources for the various supergravity fields. Specializing (3.85) to D3 branes in ten dimensions ($D = 10$ and $p = 3$) yields the metric

$$
\begin{aligned}
ds^2 &= \frac{1}{\sqrt{H(\mathbf{y})}}\,\eta_{\mu\nu}\,dx^\mu\,dx^\nu + \sqrt{H(\mathbf{y})}\,\left(dy^2 + y^2\,d\Omega_5^2\right)\\
H(\mathbf{y}) &= 1 + \left(\frac{L}{y}\right)^4 .
\end{aligned} \tag{3.89}
$$

Let us now discuss the limits of this metric: When $y^4 \gg L^4 = 4\pi g_s N\alpha'^2$, one recovers flat 10-$D$ space. When $y < L$, on the other hand, the metric appears to be singular as $y \to 0$. To examine this limit more carefully, let us define a new coordinate $u := L^2/y$. In the limit of large u (where $H = 1 + u^4/L^4 \to u^4/L^4$), the metric takes the asymptotic form

$$
ds^2\Big|_{u\to\infty} = L^2\left(\frac{1}{u^2}\,\eta_{\mu\nu}\,dx^\mu\,dx^\nu + \frac{du^2}{u^2} + d\Omega_5^2\right). \tag{3.90}
$$

In this *near horizon limit* $y \to 0 \Leftrightarrow u \gg L$, the geometry close to the brane is regular and highly symmetrical (with isometry group $SO(4, 2) \times SO(6)$). Apart from the S^5 sphere represented by $d\Omega_5^2$, we recover the AdS_5 metric (3.71).

An important property of the metric (3.89) is its non-constant *redshift factor* $\big(H(\mathbf{y})\big)^{-1/4} = g_{tt}$ with an interesting near horizon limit:

$$\big(H(\mathbf{y})\big)^{-1/4} = \Big(1 + L^4/y^4\Big)^{-1/4} = \begin{cases} \sim 1 & : \text{large } y \\ \sim y/L & : \text{small } y \end{cases} \tag{3.91}$$

The energy E_p of an object measured by an observer at constant position y differs from the energy E_i of the same object, this time measured by an observer at infinity,

$$\big(H(\mathbf{y})\big)^{-1/4} E_p = E_i. \tag{3.92}$$

When the object approaches $y \to 0$, it appears to have lower and lower energy to the observer at infinity. This gives another, geometric notion of low energy regime. We have to distinguish two kinds of low energy excitations

- particles approaching $y \to 0$,
- and massless particles propagating in the bulk (away from $y = 0$).

Their excitations decouple from each other in the low energy limit: Bulk massless particles decouple from the near horizon region around $y \to 0$. Excitations close to $y = 0$ are trapped by the gravitational potential to the $\text{AdS}_5 \times S^5$ region. Thus we have two decoupled actions in the low-energy limit: Supergravity of massless particles in flat space and supergravity in the $\text{AdS}_5 \times S^5$ region.

Comparison

As just discussed, both from the point of view of the field theory limit of open strings and from the supergravity point of view, there are two decoupled theories in the low-energy regime. In each case, one of them is supergravity of massless particles. We are thus led to identify the other theory present in each interpretation in the low-energy limit:

$$\boxed{\mathcal{N} = 4 \,\text{SYM with gauge group}\, SU(N) \overset{(*)}{\Longleftrightarrow} \text{type IIB supergravity}}\,.$$

The $(*)$ above the arrow indicates that the correspondence claimed in this *AdS/CFT conjecture* holds in the $N \to \infty$ limit at large and fixed 't Hooft coupling $\lambda = g_s N$. This field theory limit is implied by the gravity side on which $L^4 = 4\pi\lambda\alpha'^2$: Since L is finite, $\alpha' \to 0$ implies that the t'Hooft coupling $\lambda = g_{\text{YM}}^2 N \equiv g_s N$ must be large. At the same time, $g_s \to 0$, i.e. the classical limit in which the string coupling goes to zero, implies that $N \to \infty$.

3.3.2.3 Different Forms of the AdS/CFT Conjecture

We may distinguish between different strength of the AdS/CFT conjecture depending on whether the string coupling g_s, the inverse string tension α' or both are taken to zero.

The *strongest form of the AdS/CFT correspondence* conjectures that the duality between the supersymmetric $SU(N)$ gauge theory and type IIB supergravity holds for *any* value of N and g_s. This implies that $\mathscr{N} = 4$ SYM is exactly equivalent to the full type IIB superstring theory on $\mathrm{AdS}_5 \times S^5$. However, it is at present not possible to test the strongest form since there is no consistent non-perturbative quantization of string theory yet, in particular not in curved spacetime.

In the *strong form of the AdS/CFT conjecture*, we keep $\lambda = g_\mathrm{s} N$ fixed while sending $N \to \infty$. In this case the ground state is classical type IIB string theory on $\mathrm{AdS}_5 \times S^5$. The perturbative expansion parameter is $g_\mathrm{s} = \lambda/N \ll 1$ on the string theory side, this corresponds to a perturbative $1/N$ expansion on the field theory side.

Finally, there is the *weak form of the AdS/CFT conjecture* described above, in which the *Maldacena limit* $N \to \infty$ and λ very large is considered. This relates $\mathscr{N} = 4$ SYM at strong coupling and with $N \to \infty$ to classical supergravity. In contrast to the stronger versions, α' is assumed to be small now, and the α' expansion of supergravity is dual to a field theory expansion in $\lambda^{-1/2}$ powers around the strong coupling limit. The weak form can be seen as a duality in the following limits,

$$\begin{pmatrix} N \to \infty \\ \lambda \ \to \infty \end{pmatrix} \leftrightarrow \begin{pmatrix} g_\mathrm{s} \to 0 \\ \alpha' \to 0 \end{pmatrix} .$$

We see that the AdS/CFT map provides a strong/weak coupling duality: Strongly coupled quantum field theory is mapped to weakly coupled gravity. This means that if we go on to test the correspondence in the following section, we may compare only those observables in the two theories which are independent of the coupling. On the other hand, this strong/weak duality provides an interesting new tool to make non-trivial predictions about strongly coupled quantum field theories by performing calculations in weakly coupled gravity.

3.3.3 Field-Operator Map

The aim of this section is to work out the precise dictionary between objects of the two equivalent theories,

$$\begin{pmatrix} \mathscr{N} = 4\,\mathrm{SYM} \\ N, \lambda \to \infty \end{pmatrix} \leftrightarrow \begin{pmatrix} \text{type IIB supergravity} \\ \text{on } \mathrm{AdS}_5 \times S^5 \end{pmatrix} ,$$

in particular between representations of the common symmetry groups. We will relate field theory operators to supergravity fields which transform in the same representation of the superconformal group $SU(2,2|4)$ or its bosonic subgroup

$SO(6)\times SO(4,2)$. This provides a one-to-one map between gauge invariant operators in $\mathcal{N}=4$ SYM and classical fields in IIB supergravity on $\mathrm{AdS}_5\times S^5$. Moreover, we explain how to calculate correlation functions for the field theory operators by considering propagation through Anti-de Sitter space.

3.3.3.1 Correlation Functions

A crucial role in testing the AdS/CFT correspondence is played by the computation and comparison of correlation functions. Correlators which obey *non-renormalization theorems* (i.e. which are λ independent) will be of particular interest. Let us give a brief review of correlation functions in QFT.

Consider an n-point function of composite regularized gauge invariant operators $\mathcal{O}_k(x)$,

$$\langle \mathcal{O}_1(x_1)\,\mathcal{O}_2(x_2)\,\ldots\,\mathcal{O}_n(x_n)\rangle\,.$$

An important tool to compute this correlator is the generating functional $Z[J]$ (and its analogue $W[J]$ for connected diagrams) defined by

$$Z[J]:=\left\langle \exp\left(-\int d^Dx\mathcal{L}_J\right)\right\rangle = e^{-W[J]}\,, \tag{3.93}$$

where $\mathcal{L}_J$ is the Lagrangian of a given QFT with added source term coupled to a basis $\{\mathcal{O}_i\}$ of gauge invariant local operators:

$$\mathcal{L}_J=\mathcal{L}+\sum_i J_i\,\mathcal{O}_i\,. \tag{3.94}$$

The n-point correlation function is then given by

$$\langle \mathcal{O}_1(x_1)\,\mathcal{O}_2(x_2)\,\ldots\,\mathcal{O}_n(x_n)\rangle = \frac{\delta^n \ln Z[J]}{\delta J_1(x_1)\,\delta J_2(x_2)\,\ldots\,\delta J_n(x_n)}\bigg|_{J_i=0}\,. \tag{3.95}$$

To calculate correlation functions in $\mathrm{AdS}_5\times S^5$, it is convenient to work in Euclidean AdS_5 with Poincaré coordinates

$$H:=\left\{(z_0,\mathbf{z}),z_0>0,\mathbf{z}\in\mathbb{R}^4\right\},\quad \partial H=\mathbb{R}^4\,. \tag{3.96}$$

The metric

$$ds^2=\frac{1}{z_0^2}\left(dz_0^2+d\mathbf{z}^2\right) \tag{3.97}$$

diverges at the boundary $z_0\to 0$, but it is merely a coordinate singularity, not a curvature singularity. The divergence may be removed by a Weyl rescaling. As we will see later, however, sometimes it is necessary (and useful) to consider a cutoff at fixed $z_0 = \varepsilon$. The UV cutoff $\Lambda = \frac{1}{\varepsilon}$ is mapped to an IR cutoff ε in AdS.

Given the conformal symmetry, it is natural to assume that $\mathcal{N} = 4$ SYM lives on the boundary of AdS_5. This boundary may be mapped to flat space by a conformal transformation.

Typical gauge invariant operators in $SU(N)$ SYM with $\mathcal{N} = 4$ in $D = 4$ are

$$\mathcal{O}_\Delta(x) := \mathrm{sTr}\left\{X^{i_1} X^{i_2} \ldots X^{i_\Delta}\right\} = N^{(1-\Delta)/2} C_{i_1 \ldots i_\Delta} \mathrm{Tr}\left\{X^{i_1} X^{i_2} \ldots X^{i_\Delta}\right\}. \quad (3.98)$$

Here, Δ denotes the conformal dimension of the operators, X^i are the elementary scalar fields of $\mathcal{N} = 4$ SYM transforming in the representation $\underline{6}$ of $SO(6) \cong SU(4)$ and $C_{i_1 \ldots i_\Delta}$ fall into the totally symmetric rank Δ tensor representation of $SO(6)$. The trace is taken over colour indices (recall that all the fields transform in the adjoint representation of $SU(N)$). The normalization is chosen such that all planar graphs scale with N^2.

3.3.3.2 The Dual Fields of Supergravity

On the AdS side, we decompose all fields into Kaluza–Klein towers on S^5, i.e. we expand the fields in spherical harmonics $Y_\Delta(\mathbf{y})$ of S^5

$$\varphi(z, \mathbf{y}) = \sum_{\Delta=0}^{\infty} \varphi_\Delta(z)\, Y_\Delta(\mathbf{y})\,. \quad (3.99)$$

The ten-dimensional Klein–Gordon equation implies a massive wave equation in the five dimensional AdS sector,

$$\left(\Box_5 + m_\Delta^2\right) \varphi_\Delta(z) = 0, \quad m_\Delta^2 = \Delta\,(\Delta - 4). \quad (3.100)$$

It has two independent solutions which can be characterized by their asymptotics as $z_0 \to 0$,

$$\varphi_\Delta(z_0, \mathbf{z}) \sim \begin{cases} z_0^\Delta & : \text{normalizable}, \\ z_0^{4-\Delta} & : \text{non-normalizable}. \end{cases} \quad (3.101)$$

The non-normalizable fields define associated boundary fields [3] by virtue of

$$\bar{\varphi}_\Delta(\mathbf{z}) := \lim_{z_0 \to 0} \varphi_\Delta(z_0, \mathbf{z})\, z_0^{\Delta-4}. \quad (3.102)$$

We may identify the normalizable AdS modes φ_Δ as vacuum expectation values of the field theory operators $\mathcal{O}_\Delta$ and the non-normalizable modes $\bar{\varphi}_\Delta$ as sources for these operators,

$$\varphi_\Delta(z_0, \mathbf{z}) \sim \langle \mathcal{O}_\Delta \rangle\, z_0^\Delta + \bar{\varphi}_\Delta\, z_0^{4-\Delta}\,. \quad (3.103)$$

The mapping between correlation functions in Super Yang–Mills theory and the supergravity dynamics is given as follows: The generating functional $W[\bar{\varphi}_\Delta]$ for all correlators of single trace operators $\mathcal{O}_\Delta$ in Super Yang–Mills is given in terms of the source fields $\bar{\varphi}_\Delta$. The boundary values of these supergravity fields at the four-dimensional boundary of the five-dimensional AdS space become the sources for the QFT operators. In other words, on the field theory side we have

$$e^{-W[\bar{\varphi}_\Delta]} = \left\langle \exp\left(-\int_{\partial H} d^4z \bar{\varphi}_\Delta \mathcal{O}_\Delta\right)\right\rangle, \tag{3.104}$$

where the boundary fields $\bar{\varphi}_\Delta$ correspond to the sources given by (3.94). The AdS side is governed by an action in terms of the bulk fields $S[\varphi_\Delta]$ in the framework of type IIB supergravity on $\mathrm{AdS}_5 \times S^5$. The AdS/CFT conjecture for correlation functions says that precisely this classical gravity action enters the generating functional for the subclass $\{\mathcal{O}_\Delta\}$ of operators in the $\mathcal{N} = 4$ QFT. The AdS/CFT correspondence for correlators may now be expressed in the formula

$$W[\bar{\varphi}_\Delta] = S[\varphi_\Delta]\Big|_{\lim_{z_0\to 0}(\varphi_\Delta(z_0,\mathbf{z})\, z_0^{\Delta-4}) = \bar{\varphi}_\Delta(\mathbf{z})}. \tag{3.105}$$

Here, the field theory generating functional as given by (3.104) is identified with the classical action on five-dimensional Anti-de Sitter space, subject to the boundary condition that the five-dimensional fields φ_Δ assume the boundary values $\bar{\varphi}_\Delta$ in agreement with (3.103). The action S is the generating functional for tree diagrams on AdS space, i.e. for the classical expansion of correlators. It should be noted that (3.105) is a very-non-trivial statement!

The tree level graphs in AdS are referred to as *Witten diagrams* [2]. Let us give the corresponding Feynman rules:

- Each external source $\bar{\varphi}_\Delta(\mathbf{z})$ is located at the boundary;
- Propagators depart from the external sources either to another boundary point or to an interior interaction point (in which case they are called *bulk-to-boundary propagators*);
- The structure of the interior interaction points is governed by the interaction vertices of the supergravity action. These are obtained from the Kaluza–Klein reduction on S^5;
- Two interior interaction points may be connected by *bulk-to-bulk propagators*.

3.3.3.3 AdS Propagators

In this section we will derive the scalar propagator in Euclidean AdS spacetime H as defined in (3.96). For simplicity, the AdS radius is set to $L = 1$. The four vector $\mathbf{z}$ in the metric $ds^2 = \frac{1}{z_0^2}(dz_0^2 + d\mathbf{z}^2)$ parametrizes the boundary ∂H. The geodesic distance is obtained by solving the geodesic equation (where the parameter ξ is called *chordal distance*)

$$d(z,w) = \int_z^w ds = \ln\left(\frac{1+\sqrt{1-\xi^2}}{\xi}\right), \quad \xi = \frac{2\,z_0\,w_0}{z_0^2 + w_0^2 + (\mathbf{z}-\mathbf{w})^2}. \tag{3.106}$$

Let us start from the scalar part of the action which we obtain by Kaluza–Klein reduction of the ten dimensional IIB supergravity on S^5. Schematically we get

$$S[\varphi_\Delta] = \int d^5z\sqrt{|g|}\left(\frac{1}{2}\,g^{\mu\nu}\,\partial_\mu\varphi_\Delta\,\partial_\nu\varphi_\Delta + \frac{m_\Delta^2}{2}\,\varphi_\Delta + \mathscr{L}_{\text{int}}\right), \tag{3.107}$$

where $\mathscr{L}_{\text{int}}$ denotes higher order interaction terms from Kaluza–Klein reduction. Now the propagators are represented by integral kernels K_Δ, G_Δ, namely

$$\varphi_\Delta(z) = \int_{\partial H} d^4\mathbf{x} K_\Delta(z,\mathbf{x})\,\bar{\varphi}_\Delta(\mathbf{x}) \equiv \text{bulk-to-boundary propagator}\,, \tag{3.108}$$

$$\varphi_\Delta(z) = \int_H d^5x G_\Delta(z,x)\,J(x) \equiv \text{bulk-to-bulk propagator}\,, \tag{3.109}$$

where $\mathbf{x}$ denotes the boundary coordinates and $x \equiv (x_0, (x))$ denotes the bulk coordinates. The scalar Green function satisfies

$$\left(\Box_g + m_\Delta^2\right) G_\Delta(z,x) = \delta^5(z,x) \equiv \frac{\prod_{i=1}^5 \delta(z_i - x_i)}{\sqrt{|g|}}, \quad m_\Delta^2 = \Delta\,(\Delta - 4)\,, \tag{3.110}$$

where the action of the Laplacian $\Box_g$on scalar fields is in general given by

$$\Box_g\,\varphi = -\frac{1}{\sqrt{|g|}}\,\partial_\mu\sqrt{|g|}\,g^{\mu\nu}\,\partial_\nu\varphi \tag{3.111}$$

and reduces to the following expression

$$\Box_g\Big|_{\text{AdS}} = -\,z_0^2\,\partial_0^2 + 3\,z_0\,\partial_0 - z_0^2\sum_{i=1}^{d}\partial_i^2\,. \tag{3.112}$$

This turns (3.110) into a hypergeometric equation. The Green function solving (3.110) is thus given by a hypergeometric function in the argument ξ from (3.106), namely

$$\begin{aligned} G_\Delta(z,w) = G_\Delta(\xi) &= \frac{C_\Delta}{2^\Delta\,(2\Delta - d)}\,\xi^\Delta\,F_{2,1}\left(\tfrac{\Delta}{2}, \tfrac{\Delta+1}{2}, \Delta - 1; \xi^2\right), \\ C_\Delta &= \frac{\Gamma(\Delta)}{\pi^2\,\Gamma(\Delta-2)}\,. \end{aligned} \tag{3.113}$$

When x is located at the boundary, G_Δ reduces to the bulk-to-boundary propagator

$$K_\Delta(z,\mathbf{x}) = C_\Delta\left(\frac{z_0}{z_0^2 + (\mathbf{z}-\mathbf{x})^2}\right)^\Delta. \tag{3.114}$$

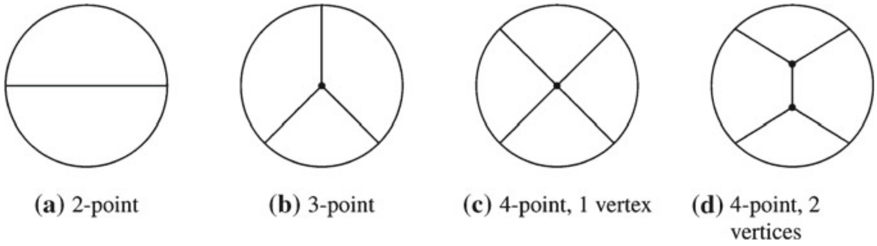


Fig. 3.6 Some examples for tree level *Witten diagrams* in AdS space-time. The circle denotes the *boundary* of AdS. The vertices denoted by the dots are in the *bulk* of AdS. Diagrams (**a**)–(**c**) involve bulk-to-boundary propagators only, while diagram (**d**) also contains a bulk-to-bulk propagator

3.3.3.4 Two-Point Function

Let us now calculate the two-point function $\langle \mathscr{O}_\Delta(\mathbf{x}) \mathscr{O}_\Delta(\mathbf{y}) \rangle$ as given by the Witten diagram (a) in Fig. 3.6 Calculation of the two-point function requires careful treatment of potential divergences at the boundary. For this purpose, we Fourier transform the boundary coordinates to momentum space. For generality, we work on $(d+1)$-dimensional AdS space with d-dimensional boundary. For the two-point function, only quadratic terms in the action are relevant. The $d+1$ dimensional bulk action

$$S[\varphi] = \int d^{d+1}z \sqrt{|g|} \left(\frac{1}{2} \partial_\mu \varphi \, \partial^\mu \varphi + \frac{m^2}{2} \varphi^2 \right) \tag{3.115}$$

gives rise to a boundary term after integration by parts,

$$S[\bar{\varphi}] = \frac{1}{2\,\varepsilon^{d-1}} \int d^d \mathbf{z} \bar{\varphi}(\mathbf{z}) \, \partial_0 \varphi(\varepsilon, \mathbf{z}) \,. \tag{3.116}$$

This action is regularized by cutting off the z_0 integral at $z_0 = \varepsilon$. In the notation $\varphi(\varepsilon, \mathbf{p}) \equiv \bar{\varphi}(\mathbf{p}) \varepsilon^{d-\Delta}$, we Fourier transform

$$\varphi(z_0, \mathbf{z}) = \int d^d \mathbf{p} e^{i\mathbf{p}\cdot\mathbf{z}} \, \varphi(z_0, \mathbf{p}) \,. \tag{3.117}$$

This yields the equation of motion in momentum space,

$$\left(z_0^2 \, \partial_0^2 - (d-1) \, z_0 \, \partial_0 - (\mathbf{p}^2 \, z_0^2 \; + \; m_\Delta^2) \right) \varphi(z_0, \mathbf{p}) = 0 \,. \tag{3.118}$$

This is a *Bessel equation* with solutions $z_0^{d/2} K_\nu(z_0 p)$ (where $\nu = \Delta - d/2$ and $p = \sqrt{\mathbf{p} \cdot \mathbf{p}}$ and K_ν is a Bessel function). The boundary asymptotics are governed by $\lim_{z_0 \to \infty} z_0^{d/2} K_\nu(z_0 p) = 0$ and $K_\nu(z_0 \to 0) \sim z_0^{d-\Delta}$.

The normalized solutions to the boundary problem read

$$\varphi(z_0, \mathbf{p}) = \frac{z_0^{d/2} K_\nu(z_0 p)}{\varepsilon^{d/2} K_\nu(\varepsilon p)} \bar{\varphi}(\mathbf{p}) \, \varepsilon^{d-\Delta}. \tag{3.119}$$

We insert the Fourier transform (3.117) into (3.118) and obtain

$$\frac{1}{2\varepsilon^{d-1}} S_p[\bar{\varphi}] = \int d^p\mathbf{p} \, d^d\mathbf{q} (2\pi)^d \, \delta^d(\mathbf{p}+\mathbf{q}) \, \varphi(\varepsilon, \mathbf{p}) \, \partial_0 \varphi(\varepsilon, \mathbf{q}) \, . \tag{3.120}$$

Using (3.119), together with the AdS/CFT conjecture (3.105), we obtain the following two-point functions for the dual CFT operators,

$$\langle \mathscr{O}_\Delta(\mathbf{p}) \, \mathscr{O}_\Delta(\mathbf{q}) \rangle_\varepsilon = -\frac{\delta^2 S_p[\bar{\varphi}]}{\delta\bar{\varphi}(\mathbf{p}) \, \delta\bar{\varphi}(\mathbf{q})} = -\frac{(2\pi)^d \, \delta^d(\mathbf{p}+\mathbf{q})}{\varepsilon^{2\Delta-d-1}} \frac{d}{d\varepsilon} \ln\Big(\varepsilon^{d/2} K_\nu(\varepsilon \mathbf{p})\Big) \, . \tag{3.121}$$

The Bessel index ν is a positive integer whenever the associated CFT operator $\mathscr{O}_\Delta$ with $\Delta = \nu + d/2$ is a chiral primary. Bessel functions have an asymptotic $u \to 0$ expansion of the schematic form

$$K_\nu(u) \to u^{-\nu} (a_0 + a_1 u^2 + a_2 u^4 + \ldots) + u^\nu \ln u \, (b_0 + b_1 u^2 + b_2 u^4 + \ldots) \, . \tag{3.122}$$

This translates as follows to the level of two point functions

$$\begin{aligned} \langle \mathscr{O}_\Delta(\mathbf{p}) \, \mathscr{O}_\Delta(\mathbf{q}) \rangle_\varepsilon = \frac{(2\pi)^d \, \delta^d(\mathbf{p}+\mathbf{q})}{\varepsilon^{2\Delta-d}} & \Bigg(-\frac{d}{2} + \nu \, (1 + c_2 + \varepsilon^2 p^2 + c_4 \, \varepsilon^4 p^4 + \ldots) \\ & - \frac{2\nu \, b_0}{a_0} \varepsilon^{2\nu} p^{2\nu} \ln(\varepsilon p) \, (1 + a_2 \, \varepsilon^2 p^2 + \ldots) \Bigg) \, . \end{aligned} \tag{3.123}$$

Explicitly, we have $\frac{2\nu b_0}{a_0} = \frac{(-1)^{\nu+1}}{2^{2(\nu-1)} \Gamma(\nu)^2}$ and $\varepsilon^{2\nu} = \varepsilon^{2\Delta-d}$, such that

$$\langle \mathscr{O}_\Delta(\mathbf{p}) \, \mathscr{O}_\Delta(-\mathbf{p}) \rangle_\varepsilon = \frac{\beta_0 + \beta_1 \, \varepsilon^2 p^2 + \ldots + \beta_\nu \, (\varepsilon p)^{2(\nu-1)}}{\varepsilon^{2\Delta-d}} - \frac{2\nu \, b_0}{a_0} p^{2\nu} \ln(\varepsilon p) + \mathscr{O}(\varepsilon^2). \tag{3.124}$$

The field theory of the first terms is governed by scheme-dependent contact terms $\sim \Box^m \delta^d(\mathbf{x}-\mathbf{y})$, and the second term gives the correct non-local result

$$\langle \mathscr{O}_\Delta(\mathbf{p}) \, \mathscr{O}_\Delta(-\mathbf{p}) \rangle = -\frac{2\nu \, b_0}{a_0} p^{2\nu} \ln(\varepsilon p) \, . \tag{3.125}$$

Transforming the non-local contribution $\propto p^{2\nu} \ln p$ back to position space yields the ε-independent result

$$\langle \mathscr{O}_\Delta(\mathbf{x}) \, \mathscr{O}_\Delta(\mathbf{y}) \rangle = \frac{\Gamma(\Delta)}{\Gamma(\Delta - d/2)} \frac{2\Delta - d}{\pi^{d/2} \, |\mathbf{x}-\mathbf{y}|^{2\Delta}} \, , \tag{3.126}$$

which agrees with the spatial dependence expected from conformal field theory. The coefficient is fixed by our AdS_5 toy model. For results in $\mathscr{N} = 4$ Super Yang–Mills theory, we also have to include the S^5 factor into our discussion. This will be done in the next chapter.

3.4 Tests of the Correspondence

3.4.1 Three-Point Function of 1/2 BPS Operators

An impressive test of the AdS/CFT correspondence is the agreement of the three-point functions of 1/2 BPS operators in $\mathcal{N} = 4$ SYM at large N with the corresponding fields in supergravity. To demonstrate this result, we will proceed as follows:

- Look at the two-point functions to fix the normalization;
- Calculate the three-point function in Super Yang–Mills theory to zeroth order in the coupling;
- Check that this is not renormalized at higher orders, i.e. prove a non-renormalization theorem to show independence of the correlator on the coupling;
- Calculate the correlation function on the gravity side (spacetime dependence from the Witten diagrams and couplings from the Kaluza–Klein reduction).

3.4.1.1 Correlation Functions of 1/2 BPS Operators

In this section, we follow the computation of Seiberg and collaborators [16] and adopt their notation: An 1/2 BPS operator of $\mathcal{N} = 4$ SYM will be denoted by

$$\mathscr{O}_k^I = C^I_{i_1 \dots i_k} \operatorname{Tr}\left\{X^{i_1} \dots X^{i_k}\right\}, \tag{3.127}$$

where $k \equiv \Delta$ and the C^I are totally symmetric traceless rank k tensors of $SO(6)$.

The SYM action is normalized such that $g^2_{\mathrm{YM}} = 4\pi g_{\mathrm{s}}$, and the normalization $\operatorname{Tr}\{T^a T^b\} = \delta^{ab}/2$ of the $SU(N)$ generators T^a allows to recast it into the form

$$\begin{aligned} S &= -\frac{1}{2\,g^2_{\mathrm{YM}}} \int d^4x \operatorname{Tr}\left\{F_{\mu\nu} F^{\mu\nu}\right\} + \text{SUSY completion}, \\ &= -\frac{1}{4\,g^2_{\mathrm{YM}}} \int d^4x F^a_{\mu\nu} F^{a\mu\nu} + \text{SUSY completion}\,. \end{aligned} \tag{3.128}$$

This gives rise to the following scalar propagators

$$\langle X^{ia}(x)\, X^{jb}(y)\rangle = \frac{g^2_{\mathrm{YM}}\,\delta^{ij}\,\delta^{ab}}{(2\pi)^2\,|x-y|^2}\,. \tag{3.129}$$

The two-point function on the field theory side to lowest order in perturbation theory is therefore given by

$$\begin{aligned} \langle \mathscr{O}_k^I(x)\, \mathscr{O}_k^J(y)\rangle &= C^I_{i_1\dots i_k}\, C^J_{j_1\dots j_k}\, \langle \operatorname{Tr}\left\{X^{i_1}(x) \dots X^{i_k}(x)\right\} \operatorname{Tr}\left\{X^{j_1}(y) \dots X^{j_k}(y)\right\}\rangle \\ &= C^I_{i_1\dots i_k}\, C^J_{j_1\dots j_k}\, \frac{N^k\, g^{2k}_{\mathrm{YM}}\left(\delta^{i_1 j_1}\,\delta^{i_2 j_2} \dots \delta^{i_k j_k} + \text{cyclic permutations}\right)}{(2\pi)^{2k}\,|x-y|^{2k}} \\ &= \frac{k\,\lambda^k\,\delta^{IJ}}{(2\pi)^{2k}\,|x-y|^{2k}}\,, \end{aligned} \tag{3.130}$$

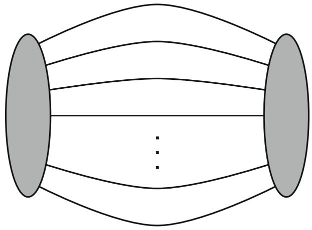


Fig. 3.7 Feynman diagram for the two-point correlation function $\langle \mathscr{O}^I_k(x)\, \mathscr{O}^J_k(y)\rangle$ to lowest order in the coupling, $\mathscr{O}({g_{\mathrm{YM}}}^0)$, in the planar limit. This is referred to as the *rainbow diagram*

where last equality only holds at leading order in N. In the large N limit, the corresponding planar Feynman diagram in shown in Fig. 3.7.

Similarly, for the three-point function to lowest order in perturbation theory and in the limit of large N we have [16]

$$\langle \mathscr{O}^I_{k_1}(x)\, \mathscr{O}^J_{k_2}(y)\, \mathscr{O}^K_{k_3}(z)\rangle = \frac{\lambda^{\Sigma/2}\, k_1\, k_2\, k_3\, \langle C^I\, C^J\, C^K\rangle}{N\,(2\pi)^{\Sigma}\, |x-y|^{2\alpha_3}\, |y-z|^{2\alpha_1}\, |x-z|^{2\alpha_2}}\,. \tag{3.131}$$

Note that the spacetime dependence is completely determined by conformal invariance. We have used the shorthand notation

$$\Sigma = k_1 + k_2 + k_3 \quad \alpha_i = \frac{\Sigma}{2} - k_i \tag{3.132}$$

(such that e.g. $\alpha_1 = \frac{k_2+k_3-k_1}{2}$) and $\langle C^I C^J C^K\rangle$ denotes a uniquely defined $SO(6)$ tensor contraction of indices determined by the Feynman graph.

It is useful to defined normalized operators $\tilde{\mathscr{O}}^I := \frac{(2\pi)^k}{\lambda^{k/2}\sqrt{k}}\, \mathscr{O}^I$. Their two-point function is normalized to one,

$$\langle \tilde{\mathscr{O}}^I_k(x)\, \tilde{\mathscr{O}}^J_k(y)\rangle = \frac{\delta^{IJ}}{|x-y|^{2k}}\ , \tag{3.133}$$

and the three point function reads

$$\langle \tilde{\mathscr{O}}^I_k(x)\, \tilde{\mathscr{O}}^J_k(y)\, \tilde{\mathscr{O}}^K_k(z)\rangle = \frac{\sqrt{k_1\, k_2\, k_3}\, \langle C^I\, C^J\, C^K\rangle}{N\, |x-y|^{2\alpha_3}\, |y-z|^{2\alpha_1}\, |x-z|^{2\alpha_2}}. \tag{3.134}$$

This holds for large values of N. Otherwise, non-planar corrections of order $\frac{1}{N^2}$ arise.

3.4.1.2 Non–Renormalization Theorem

Next, following [17], we demonstrate the absence of $\mathscr{O}(\lambda)$ terms both in $\langle \mathscr{O}\mathscr{O}\rangle$ and in $\langle \mathscr{O}\mathscr{O}\mathscr{O}\rangle$. The argument will hold for any N [17].

Define complex scalar fields $Z^i := X^i + iX^{i+3}$ making use of the embedding $SU(3) \subset SU(4)$. The Euclidean version of the $\mathscr{N} = 4, SU(N)$ SYM Lagrangian then reads

$$\mathscr{L} = \mathrm{Tr}\Bigg\{\frac{1}{4}\,F_{\mu\nu}\,F^{\mu\nu} + \frac{1}{2}\,\bar{\lambda}\,\not{D}\lambda + D_\mu Z^i\,D^\mu \bar{Z}^i + \frac{1}{2}\,\bar{\psi}^i\,\not{D}\psi^i$$

$$+ i\sqrt{2}\,g f^{abc}\left(\bar{\lambda}_a\,\bar{Z}^i_b\,L\psi^i_c - \bar{\psi}^i_a R\,Z^i_b\,\lambda_c\right). \tag{3.135}$$

$$- \frac{g}{\sqrt{2}} f^{abc}\,\varepsilon_{ijk}\left(\bar{\psi}^i_a L\,Z^j_a\,\psi^k_c + \bar{\psi}^i_a R\,\bar{Z}^i_b\,\psi^k_c\right)$$
$$- \frac{g^2}{2} f^{abc}\,\bar{Z}^i_b\,Z^i_c f_{ade}\,\bar{Z}^{jd}\,Z^{je} + \frac{g^2}{2} f^{abc} f^{ade}\,\varepsilon_{ijk}\,\varepsilon_{ilm}\,Z^i_b\,Z^k_c\,\bar{Z}^l_d\,\bar{Z}^m_e\Bigg\}, \tag{3.136}$$

where L,R denote the left and right handed chirality projectors.

Due to supersymmetry, it is sufficient to consider

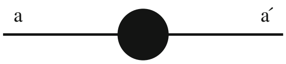
$$\langle \mathrm{Tr}\left\{(Z^1)^k(x)\right\}\mathrm{Tr}\left\{(\bar{Z}^1)^k(y)\right\}\rangle = \frac{P_{k,k,0}(N)}{\left(4\pi^2\,|x-y|\right)^k} \tag{3.137}$$

with the following polynomial in N

$$P_{k,k,0}(N) = \sum_{\sigma\in S_k} \mathrm{Tr}\left\{T^{a_1}\,T^{a_2}\,\ldots\,T^{a_k}\right\}\mathrm{Tr}\left\{T^{a_{\sigma(1)}}\,T^{a_{\sigma(2)}}\,\ldots\,T^{a_{\sigma(k)}}\right\}$$
$$= k\left(\frac{N}{2}\right)^k + \text{lower order in}\,N. \tag{3.138}$$

There are various effects to consider at leading order in the coupling, namely

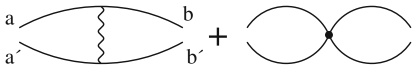
- the self energy corrections



$$= \delta^{aa'}\,N\,A(x,y)\,G(x,y)\,, \tag{3.139}$$

with $A(x,y) = a_0 + a_1\ln\left(\mu^2(x-y)^2\right)$ and the scalar propagator $G(x,y) = \frac{1}{4\pi^2|x-y|^2}$ and

- the two particle exchange interactions



$$= \left(f^{pab} f^{pa'b'} + f^{pab'} f^{pa'b}\right) B(x, y)\, G(x, y)^2 \,, \tag{3.140}$$

with $B(x, y) = b_0 + b_1 \ln\left(\mu^2 (x - y)^2\right)$.

The possible corrections to the *rainbow graph* at order g_{YM}^2 are shown in Fig. 3.8. It turns out that these three graphs cancel each other for all N and for all k. The proof goes as follows:

- Use a trace identity valid for any matrices N and M_i

$$\sum_{i=1}^{n} \mathrm{Tr}\left\{M_1 \ldots M_{i-1} \left[M_i \,,\, N\right] M_{i+1} \ldots M_n\right\} = 0, \tag{3.141}$$

- Use combinatorics for colour indices,
- Insert (3.139) between all pairs of adjacent lines using $[T^a, T^b] = i f^{abc} T^c$.
- The result for all exchange graphs (with S_k permutation σ) is then

$$\frac{1}{4}\,(-2\,B)\,\mathrm{Tr}\left\{T^{a_1} \ldots T^{a_k}\right\} \sum_{i \neq j = 1}^{k} \mathrm{Tr}\left\{T^{a_{\sigma(1)}} \ldots \left[T^{a_{\sigma(i)}} \,,\, T^p\right] \ldots \left[T^{a_{\sigma(j)}} \,,\, T^p\right] \ldots T^{a_{\sigma(k)}}\right\},.$$

- Then apply (3.14) to one of the two commutators to find

$$\begin{aligned} &\frac{B}{2}\,\mathrm{Tr}\left\{T^{a_1} \ldots T^{a_k}\right\} \sum_{i=1}^{k} \mathrm{Tr}\left\{T^{a_{\sigma(1)}} \ldots \left[\left[T^{a_{\sigma(i)}} \,,\, T^p\right] ,\, T^p\right] \ldots T^{a_{\sigma(k)}}\right\} \\ &= \frac{N\,B}{2}\,\mathrm{Tr}\left\{T^{a_1} \ldots T^{a_k}\right\} \sum_{i=1}^{k} \mathrm{Tr}\left\{T^{a_{\sigma(1)}} \ldots T^{a_{\sigma(i)}} \ldots T^{a_{\sigma(k)}}\right\}. \end{aligned} \tag{3.142}$$

The last step follows from the fact that $\left[[\cdot, T^p], T^p\right]$ is the Casimir operator of the adjoint representation of $SU(N)$ such that $\left[[T^a, T^p] T^p\right] = N T^a$ and the sum over i yields k identical terms. In the self energy corrections, we also have a factor of k by similar argument such that the overall contribution is

$$\frac{k\,N\,(B + 2A)}{2} \sum_{\sigma \in S_k} \mathrm{Tr}\left\{T^{a_1} \ldots T^{a_k}\right\} \mathrm{Tr}\left\{T^{a_{\sigma(1)}} \ldots T^{a_{\sigma(k)}}\right\} = \frac{k\,N\,(B + 2A)\,P_{k,k,0}(N)}{2}. \tag{3.143}$$

In [17] it was shown that (3.143) vanishes since $B + 2A = 0$ due to the non-renormalization theorem. The reason for that is the following: The two point function

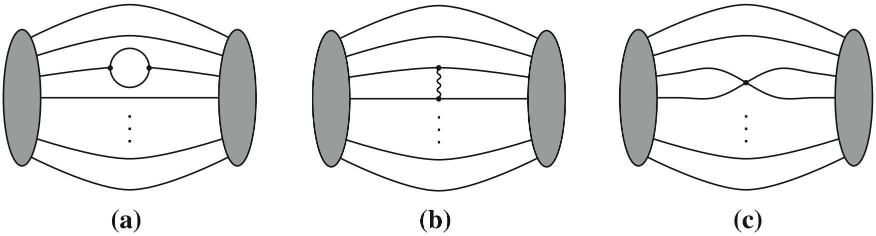


Fig. 3.8 Possible corrections to the rainbow graph at order $g_{\rm YM}^2$

$\mathrm{Tr}\{X^2\}$ falls into the same supersymmetry multiplet as the energy momentum tensor $T_{\mu\nu}$. It can be shown that the latter is not renormalized (in agreement with momentum conservation), so by supersymmetry, $\mathrm{Tr}\{X^2\}$ is protected as well. This implies that $\langle \mathscr{O}_2(\mathbf{x})\mathscr{O}_2(\mathbf{y})\rangle$ does not have any quantum corrections of order $\mathscr{O}(g_{\rm YM}^2)$, and thus $B + 2A = 0$. (3.143) then implies that $\langle \mathscr{O}_k(\mathbf{x})\mathscr{O}_k(\mathbf{y})\rangle$ is non-renormalized and independent of $g_{\rm YM}$ (and thus λ) as well for all k. Note that this non-renormalization theorem for the two-point function of 1/2 BPS operators holds for all values of N.

A similar analysis applies to the three-point functions of 1/2 BPS operators as well. Four-point functions, however, are renormalized in general, though there are special exceptional cases where they are not, see for instance [18, 19].

3.4.1.3 Three-Point Function on the Gravity Side

Having obtained an exact result for the three point function of 1/2 BPS (or chiral primary) operators on the field theory side, we are ready to compare with a gravity counterpart. Let us consider three-point functions of scalar fields in AdS spacetimes. Their Feynman diagram has the structure of a Mercedes star, as displayed in (b) of Fig. 3.6. It is specified by three edge points $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ by three bulk-to-boundary propagators and a coupling in the center determined by Kaluza-Klein reduction of S^5. The correlation function corresponding to this Witten diagram has first been calculated in [4], while the coupling has been obtained in [16].

Recall from Sect. 3.3.3.3 that the bulk-to-boundary Green functions in AdS_{d+1} is given by

$$K_\Delta(z_0, \mathbf{z}, \mathbf{x}) = \frac{\Gamma(\Delta)}{\pi^{d/2}\,\Gamma(\Delta-2)}\left(\frac{z_0}{z_0^2 + (\mathbf{z}-\mathbf{x})^2}\right)^{\Delta}. \tag{3.144}$$

Because of its defining property $\lim_{z_0\to 0}\left[z_0^{\Delta-d}K_\Delta(z_0,\mathbf{z},\mathbf{x})\right] = \delta^d(\mathbf{x}-\mathbf{z})$ we can express a bulk field ϕ in terms of its values at the boundary

$$\phi(z_0, \mathbf{z}) = \frac{\Gamma(\Delta)}{\pi^{d/2}\,\Gamma(\Delta-2)}\int d^d\mathbf{x}\left(\frac{z_0}{z_0^2 + (\mathbf{z}-\mathbf{x})^2}\right)^{\Delta}\phi_0(\mathbf{x}). \tag{3.145}$$

Now the Mercedes diagram of the gravity three point functions is evaluated as

$$A(\mathbf{x},\mathbf{y},\mathbf{z}) := \int dw_0\, d^d\mathbf{w} \frac{1}{w_0^{d+1}} \left(\frac{w_0}{(w-\mathbf{x})^2}\right)^{\Delta_1} \left(\frac{w_0}{(w-\mathbf{y})^2}\right)^{\Delta_2} \left(\frac{w_0}{(w-\mathbf{z})^2}\right)^{\Delta_3}. \tag{3.146}$$

Here, we use the notation $(w-\mathbf{x})^2 := w_0^2 + (\mathbf{w}-\mathbf{x})^2$.

The number of functions in denominator can be reduced using the trick of inversion [4]: Reexpress integration variable as $w_\mu = \frac{w'_\mu}{(w')^2}$ and similarly set $\mathbf{x} = \frac{\mathbf{x}'}{|\mathbf{x}'|^2}$, $\mathbf{y} = \frac{\mathbf{y}'}{|\mathbf{y}'|^2}$ and $\mathbf{z} = \frac{\mathbf{z}'}{|\mathbf{z}'|^2}$. Consequently, the propagators are affected as

$$K_\Delta(w,\mathbf{x}) = |\mathbf{x}'|^{2\Delta} K_\Delta(w',\mathbf{x}'). \tag{3.147}$$

The factor $|\mathbf{x}'|^{2\Delta}$ is a first parallel to field theory since $|\mathbf{x}'|^{2\Delta} = \frac{1}{|\mathbf{x}|^{2\Delta}}$. Note that inversion is an isometry of AdS, so its volume element is invariant $\frac{d^{d+1}w}{w_0^{d+1}} = \frac{d^{d+1}w'}{(w'_0)^{d+1}}$. This causes the Mercedes integral to transform as

$$A(\mathbf{x},\mathbf{y},\mathbf{z}) = |\mathbf{x}'|^{2\Delta_1} |\mathbf{y}'|^{2\Delta_2} |\mathbf{z}'|^{2\Delta_3} A(\mathbf{x}',\mathbf{y}',\mathbf{z}'). \tag{3.148}$$

To reduce the number of functions in the denominator of (3.146) from three to two, proceed as follows:

- Set one argument to zero $\mathbf{z} \to 0$ using translation invariance,

$$A(\mathbf{x},\mathbf{y},\mathbf{z}) = A(\mathbf{x}-\mathbf{z},\mathbf{y}-\mathbf{z},0) =: A(\mathbf{u},\mathbf{v},0)\,. \tag{3.149}$$

  This brings the third terms into the nice form $\left(\frac{w_0}{(w-\mathbf{z})^2}\right)^{\Delta_3} = \left(\frac{w_0}{w^2}\right)^{\Delta_3} = (w'_0)^{\Delta_3}$.
- Add an inversion to find

$$A(\mathbf{u},\mathbf{v},0) = \frac{1}{|\mathbf{u}|^{2\Delta_1}|\mathbf{v}|^{2\Delta_2}} \int \frac{d^{d+1}w'}{(w'_0)^{d+1}} \left(\frac{w'_0}{(w'-\mathbf{u}')^2}\right)^{\Delta_1} \left(\frac{w'_0}{(w'-\mathbf{v}')^2}\right)^{\Delta_2} (w'_0)^{\Delta_3}\,. \tag{3.150}$$

By translation invariance of the $\mathbf{w}$ integration variable, the integral can only depend on the difference $\mathbf{u}' - \mathbf{v}'$, and dimensional analysis fixes the power to be $|\mathbf{u}'-\mathbf{v}'|^{\Delta_3-\Delta_1-\Delta_2}$. Hence, we have already found the spacetime dependence

$$\begin{aligned} A(\mathbf{u},\mathbf{v},0) &\sim \frac{|\mathbf{u}'-\mathbf{v}'|^{\Delta_3-\Delta_1-\Delta_2}}{|\mathbf{u}|^{2\Delta_1}|\mathbf{v}|^{2\Delta_2}} \\ &= \frac{1}{|\mathbf{x}-\mathbf{y}|^{\Delta_1+\Delta_2-\Delta_3}\,|\mathbf{y}-\mathbf{z}|^{\Delta_2+\Delta_3-\Delta_1}\,|\mathbf{z}-\mathbf{x}|^{\Delta_3+\Delta_1-\Delta_2}} \\ &=: f(\mathbf{x},\mathbf{y},\mathbf{z}) \end{aligned} \tag{3.151}$$

(Note that good care has to be taken to restore the old variables before the inversion transformation. A useful formula is $(\mathbf{u}'-\mathbf{v}')^2 = \frac{(\mathbf{x}-\mathbf{y})^2}{(\mathbf{x}-\mathbf{z})^2(\mathbf{y}-\mathbf{z})^2}$.)

An exact calculation of $A(\mathbf{u}, \mathbf{v}, 0)$ can be done using Feynman parameter methods [4], the prefactor in $A(\mathbf{x}, \mathbf{y}, \mathbf{z}) = a \cdot f(\mathbf{x}, \mathbf{y}, \mathbf{z})$ is found to be

$$a = -\frac{\Gamma[\frac{1}{2}(\Delta_1 + \Delta_2 - \Delta_3)]\,\Gamma[\frac{1}{2}(\Delta_2 + \Delta_3 - \Delta_1)]\,\Gamma[\frac{1}{2}(\Delta_3 + \Delta_1 - \Delta_2)]\,\Gamma[\frac{1}{2}(\sum_i \Delta_i - d)]}{2\pi^d\,\Gamma[\Delta_1 - \frac{d}{2}]\,\Gamma[\Delta_2 - \frac{d}{2}]\,\Gamma[\Delta_3 - \frac{d}{2}]}. \tag{3.152}$$

The Gamma functions due to the Feynman parameter method have a number of poles.

Now we need to consider coupling with which the Mercedes integral (3.146) enters the three point function

$$\left\langle \mathscr{O}^I(\mathbf{x})\,\mathscr{O}^J(\mathbf{y})\,\mathscr{O}^K(\mathbf{z})\right\rangle = \lambda^{IJK}\,A(\mathbf{x}, \mathbf{y}, \mathbf{z})\,. \tag{3.153}$$

The A part was just calculated, we will next treat the cubic coupling λ^{IJK} coming from KK reduction in supergravity [16].

Recall from Sect. 3.2.4.2 that type IIB supergravity contains a self dual five form field F. It enters the equations of motion for the graviton via

$$R_{mn} = \frac{1}{3!}\,F_{mijkl}\,F_n{}^{ijkl}. \tag{3.154}$$

In the flat $\mathrm{AdS}_5 \times S^5$ background solution, the five form takes particularly simple values. Denote the AdS_5 indices by μ_i, $i = 1, 2, \ldots, 5$ and the S^5 indices by α_i, $i = 1, 2, \ldots, 5$ then the solution reads

$$\begin{aligned} ds^2 &= \frac{1}{z_0^2}\left(d\mathbf{z}^2 + dz_0^2 + d\Omega_5^2\right) =: g_{mn}\,dx^m\,dx^n\,, \\ \bar{F}_{\mu_1\ldots\mu_5} &= \varepsilon_{\mu_1\ldots\mu_5} \quad \bar{F}_{\alpha_1\ldots\alpha_5} = \varepsilon_{\alpha_1\ldots\alpha_5}\,. \end{aligned} \tag{3.155}$$

Note that the curvatures of the AdS_5 and S^5 factors cancel

$$\begin{aligned} \mathrm{AdS}_5\colon\; R_{\mu\lambda\nu\sigma} &= -(g_{\mu\nu}\,g_{\lambda\sigma} - g_{\mu\sigma}\,g_{\lambda\nu}), \quad R_{\mu\nu} = -4\,g_{\mu\nu} \quad \mathscr{R}_{\mathrm{AdS}_5} = -20\,, \\ S^5\colon\; R_{\alpha\gamma\beta\delta} &= +(g_{\alpha\beta}\,g_{\gamma\delta} - g_{\alpha\delta}\,g_{\gamma\beta}), \quad R_{\alpha\beta} = +4\,g_{\alpha\beta}, \quad \mathscr{R}_{S^5} = +20\,. \end{aligned} \tag{3.156}$$

Observe that $\mathscr{R} = \mathscr{R}_{\mathrm{AdS}_5} + \mathscr{R}_{S^5} = 0$.

Next we need to look at fluctuations ϕ_0 about this background which couple to operators $\mathscr{O}$ in the dual field theory via interaction terms $S_{\mathrm{int}} = \int d^d x\,\phi_0(x)\mathscr{O}(x)$. It was investigated in [20] how to decompose the supergravity equations of motion and how to decouple them from the fluctuations.

Starting point is the ansatz

$$G_{mn} = g_{mn} + h_{mn}, \quad F = \bar{F} + \delta F\,, \tag{3.157}$$

where the fluctiations h, δF are organized as

$$\begin{aligned} h_{\alpha\beta} &= h_{(\alpha\beta)} + \frac{h_2}{5}\,g_{\alpha\beta}, & g^{\alpha\beta}\,h_{(\alpha\beta)} &= 0\,, \\ h_{\mu\nu} &= h'_{(\mu\nu)} + \frac{h'}{5}\,g_{\mu\nu} - \frac{h_2}{3}\,g_{\mu\nu}, & g^{\mu\nu}\,h'_{(\mu\nu)} &= 0\,, \end{aligned} \tag{3.158}$$

$$\delta F_{ijklm} = 5\, \nabla_{[i}\, a_{jklm]}\,. \tag{3.159}$$

It is convenient to work in de-Donder gauge (with respect to S^5) where

$$\nabla^\alpha h_{\alpha\beta} = \nabla^\alpha h_{\mu\alpha} = \nabla^\alpha a_{\alpha\mu_1\mu_2\mu_3} = 0. \tag{3.160}$$

The KK programme requires to expand this ansatz in spherical harmonics Y^I on S^5

$$\begin{aligned} h'_{\mu\nu} &= \sum_I Y^I\, (h'_{\mu\nu})^I, \quad h_2 = \sum_I Y^I\, h_2^I\,, \\ a_{\alpha_1\ldots\alpha_4} &= \sum_I \nabla^\alpha\, Y^I\, \varepsilon_{\alpha\alpha_1\alpha_2\alpha_3\alpha_4}\, b^I\,, \\ a_{\mu_1\ldots\mu_4} &= \sum_I Y^I\, a^I_{\mu_1\ldots\mu_4}\,. \end{aligned} \tag{3.161}$$

Inserting this ansatz into the ten dimensional equations of motion leads to diagonalization and decoupling. The modes which couple to the field theory 1/2 BPS operators $\mathcal{O}^I$ are given by

$$S^I = \frac{1}{20\,(k+2)}\, \big(h_2^I - 10\,(k+4)\, b^I\big). \tag{3.162}$$

Note that $k = \Delta$ in the different notations of the original papers [4] and [16]. These S^5 modes satisfy a five dimensional equation of motion in AdS space

$$\big(\nabla_\mu \nabla^\mu - k\,(k-4)\big)\, S^I = \lambda^{IJK}\, S_J\, S_K\,. \tag{3.163}$$

where λ^{IJK} is given by

$$\lambda^{IJK} = a(k_1,k_2,k_3)\, \frac{128\, \Sigma\, \big((\Sigma/2)^2 - 1\big)\big((\Sigma/2)^2 - 4\big)\, \alpha_1\alpha_2\alpha_3\, \langle C^I C^J C^K\rangle}{(k_1+1)(k_2+1)(k_3+1)}. \tag{3.164}$$

We are using the usual shorthands $\Sigma = k_1 + k_2 + k_3$ and $\alpha_1 = \frac{k_2+k_3-k_1}{2}$ (as well as cyclic variations thereof) and the numbers $a(k_1,k_2,k_3)$ relate S^5 integrals of spherical harmonics with the $SO(6)$ tensors $\langle C^I C^J C^K\rangle$ of (3.131),

$$\begin{aligned} \int_{S^5} d\Omega\, Y^I(\Omega)\, Y^J(\Omega)\, Y^K(\Omega) &= a(k_1,k_2,k_3)\, \Big\langle C^I\, C^J\, C^K \Big\rangle\,, \\ a(k_1,k_2,k_3) &= \frac{\pi^3}{(\Sigma/2)!\, 2^{d/2(\Sigma-2)}}\, \frac{k_1!\, k_2!\, k_3!}{\alpha_1!\, \alpha_2!\, \alpha_3!}\,. \end{aligned} \tag{3.165}$$

This gives rise to the following dimensionally reduced supergravity action for the S^I modes

$$S = \frac{4\,N^2}{(2\pi)^5} \int d^5x\sqrt{g} \left[\frac{A_I}{2}\,(-\nabla S_I)^2 - k\,(k-4)\,(S_I)^2 + \frac{1}{3}\,\lambda_{IJK}\,S^I\,S^J\,S^K \right]. \tag{3.166}$$

We can identify the lower dimensional gravitation coupling and the AdS radius as

$$\frac{1}{2\,\kappa^2} = \frac{4\,N^2}{(2\pi)^5}, \quad L_{\text{AdS}} + 1, \tag{3.167}$$

and the constant A_I is determined from IIB 10d SUGRA action to be

$$A_I = 32\,\frac{k\,(k-1)\,(k-2)}{k+1}\,Z(k), \quad \int_{S^5} d\Omega\,Y^I(\Omega)\,Y^J(\Omega) =: Z(k)\,\delta^{IJ}. \tag{3.168}$$

Let us now use the action S as given above to calculate the the two point function

$$\langle S^I(x)\,S^J(y)\rangle = \frac{4\,N^2}{(2\pi)^5}\,\frac{\pi}{2^{k-7}}\,\frac{k\,(k-1)^2\,(k-2)^2}{(k+1)^2}\,\frac{\delta^{IJ}}{(x-y)^{2k}}\,, \tag{3.169}$$

then define normalized operators $\tilde{\mathcal{O}}^I(x) \sim S^I(x)$ such that $\left\langle \tilde{\mathcal{O}}^I(x)\tilde{\mathcal{O}}^J(y) \right\rangle = \frac{\delta^{IJ}}{(x-y)^{2k}}$. The three-point function is computed on the basis λ^{IJK}, the operators' normalization as given above and the result (3.151), (3.152) for $A(x, y, z)$,

$$\left\langle \tilde{\mathcal{O}}^I(x)\tilde{\mathcal{O}}^J(y)\tilde{\mathcal{O}}^K(z) \right\rangle = \frac{1}{N}\,\frac{\sqrt{k_1\,k_2\,k_3}\,\left\langle C^I\,C^J\,C^K \right\rangle}{|x-y|^{2\alpha_3}\,|y-z|^{2\alpha_1}\,|z-x|^{2\alpha_2}}\,. \tag{3.170}$$

Remarkably, this gravitational correlator coincides with the field theory result (3.134)!

Note again that for comparing quantum field theory and supergravity, it is essential to use the two-point function to normalize the operators in the same way on both sides of the correspondence. Also, it is essential to consider observables which are independent of the coupling, since the field theory calculation is performed at weak coupling while the supergravity calculation is dual to a strong coupling result in field theory. Further impressive and very non-trivial tests of the correspondence beyond non-renormalized operators, where the results *do* depend on the coupling, have been obtained in the 'integrability' approach (for a review see [21]). This requires considering the strong form of the AdS/CFT correspondence where the $\alpha' \to 0$ limit is not taken.

3.5 Introduction to Gauge/Gravity Duality

Motivated by the successes of the AdS/CFT correspondence in its original form as discussed in the previous chapters, many physicists have begun to ask the question whether similar dualities are also possible for less symmetric quantum field theories.

In particular, it would be extremely appealing to be able to map QCD to a weakly coupled gravity theory in order to solve the problem of confinement. While this goal is still a long distance away, important and interesting steps have been taken to use generalizations of AdS/CFT, i.e. gauge/gravity duality, to use weakly coupled gravity in order to make predictions for strongly coupled quantum systems which are hard to describe otherwise.

Within these lectures we have room only to give a brief outlook at this vast subject. We consider gauge/gravity duality for field theories at finite temperature and density. This is the starting point for many applications, both to the quark-gluon plasma of heavy ion physics and to strongly coupled systems of potential relevance for condensed matter physics. We recommend the reviews [9] for the quark-gluon plasma, [10] for studying mesons and quark degrees of freedom, and [11, 12] for condensed matter applications.

We starty by switching on a temperature in $\mathcal{N} = 4$ Super Yang–Mills theory. This breaks all of the supersymmetry. We still keep the $N \to \infty$ planar limit.

3.5.1 Gauge/Gravity Duality at Finite Temperature

3.5.1.1 Finite Temperature Field Theory

Finite temperature field theory is obtaine by performing a Wick rotation $t \to -i\tau$ and by compactifying Euclidean time on a circle of radius $\beta = \frac{1}{b_B T}$. This has to effect of modifying the weight factor in the correlation functions, $e^{iHt} \to e^{-\beta H}$, such that we describe a field theory in thermal equilibrium with temperature T.

The partition function of a thermal field theory is given by

$$\begin{aligned} Z_{\mathrm{E}} = \mathrm{Tr} e^{-\beta H} &= \sum_{\text{all } \beta\text{-periodic states}} \langle \varphi_\beta | e^{-\beta H} | \varphi_\beta \rangle \\ &= \int_{\text{all } \beta\text{-periodic states}} \mathscr{D}\varphi e^{-S_{\mathrm{E}}[\varphi]}, \end{aligned} \tag{3.171}$$

for a field theory with fields φ.

We now use the saddle point approximation: Let φ^* be a saddle point of the Euclidean action $S_{\mathrm{E}}[\varphi]$. Then we can approximate the generating functional semiclassically as

$$Z = \int \mathscr{D}\varphi e^{-S_{\mathrm{E}}[\varphi]} \approx e^{-S_{\mathrm{E}}[\varphi^*]}. \tag{3.172}$$

3.5.1.2 Finite Temperature on the Gravity Side

According to the weak form of the AdS/CFT correspondence, the partition function of the classical bulk theory with asymptotically AdS boundary conditions is equivalent to the partition function of the large N QFT. The metric g then takes the role of the φ field above:

$$Z_{\text{grav}} = e^{-S_{\text{E}}[g^*]} \tag{3.173}$$

The gravitational action contains a *Gibbons-Hawking boundary term* required for finiteness,

$$S_{\text{E}}[g] = -\frac{1}{2\,\kappa^2}\int d^{d+1}x\sqrt{g}\left(\mathscr{R} + \frac{d\,(d-1)}{L^2}\right) + \frac{1}{2\,\kappa^2}\int_{r\to 0} d^d x\sqrt{g}\left(-2\,\mathscr{K} + \frac{2\,(d-1)}{L^2}\right). \tag{3.174}$$

Here, $\mathscr{K}$ denotes the trace of the extrinsic curvature,

$$\mathscr{K} = \gamma^{\mu\nu}\,\nabla_\mu\, n_\nu\,, \tag{3.175}$$

where $\gamma^{\mu\nu}$ is the induced metric on the boundary at $r \to 0$ and n^μ an outward pointing unit normal vector on the boundary.

A saddle point, i.e. a solution to the equations of motion, is given by

$$ds^2 = \frac{L^2}{r^2}\left(f(r)\,d\tau^2 + \frac{dr^2}{f(r)} + dx^i\,dx^i\right), \quad f(r) = 1 - \frac{r^4}{r_{\text{H}}^4}\,. \tag{3.176}$$

This coincides with the analytic continuation of the AdS Schwarzschild metric to Euclidean space, where r_{H} denotes the Schwarzschild horizon. We thus have a black hole as solution to the equations of motion. The Euclidean metric (3.176) is defined outside the horizon only. At the horizon, we have $r = r_H$ and and the boundary, we have $r = 0$.

We now show that regularity at the horizon is obtained only if τ is periodic. The period given by $\beta = \frac{1}{T}$ is identified with the inverse temperature, with $k_B \equiv 1$ for the Boltzmann constant. Consider the behaviour of the metric near the horizon r_{H}. In this region, the metric in the (τ, r) plane becomes, defining $r' = r_{\text{H}} - r$,

$$ds^2 = L^2\left(\frac{4r'}{r_{\text{H}}}d\tau^2 + \frac{r_{\text{H}}}{4r'}\,dr'^2\right). \tag{3.177}$$

This metric may be rewritten as

$$ds^2 = d\rho^2 + \rho^2\frac{4}{{r_{\text{H}}}^2}d\tau^2 \qquad \text{with} \quad \rho^2 = r'L\,. \tag{3.178}$$

This expression corresponds to the metric of a plane in polar coordinates,

$$ds^2 = d\rho^2 + \rho^2 d\theta^2 \quad \text{with} \quad \theta \equiv \frac{2\tau}{r_{\text{H}}}\,. \tag{3.179}$$

With this choice for θ a conical singularity can be avoided. The periodicity of θ can be translated into periodicity of τ with period

$$\frac{2}{r_{\rm H}}\beta = 2\pi\ . \tag{3.180}$$

Keeping in mind that $T = 1/\beta$, we identify a relation between the horizon radius $r_{\rm H}$ and the Hawking temperature of the AdS Schwarzschild black hole,

$$T = \frac{1}{\pi\, r_{\rm H}}\ . \tag{3.181}$$

The Hawking temperature of the black hole is then identified with the temperature of the field theory on the boundary.

We obtain further thermodynamic quantities by evaluating the partition function at the saddle point $e^{-S_{\rm E}[g^*]}$. The action as given in (3.174), evaluated at the Euclidean Schwarzschild metric, is found to be

$$S_{\rm E} = -\ \frac{L^{d-1}}{2\,\kappa^2\, r_{\rm H}^d}\ \frac{V_{d-1}}{T} = -\ \frac{(4\pi)^d\, L^{d-1}\, V_{d-1}\, T^{d-1}}{2\,\kappa^2\, d^d} \tag{3.182}$$

where V_{d-1} is the spatial volume of the associated QFT.

In order to be in the classical gravity regime, we need that the spacetime is weakly curved in Planck units, i.e. that $\frac{L^{d-1}}{\kappa^2} \ll 1$. The dual field theory analogue of $\frac{L^{d-1}}{\kappa^2} \ll 1$ is $N \to \infty$, recall that $L^4 = 4\pi\, g_{\rm s} N \alpha'^2$.

From the action given by (3.182) we obtain the free energy and entropy as

$$F = -\,T\ \ln Z = T\, S_{\rm E}[g^*] = -\ \frac{(4\pi)^d\, L^{d-1}\, V_{d-1}\, T^d}{2\,\kappa^2\, d^d} \tag{3.183}$$

$$S = -\ \frac{\partial F}{\partial T} = \frac{(4\pi)^d\, L^{d-1}\, V_{d-1}\, T^{d-1}}{2\,\kappa^2\, d^{d-1}}. \tag{3.184}$$

The expression for the entropy is equal to the area of the event horizon divided by $4G_{\rm N} = \frac{\kappa^2}{2\pi}$. This area entropy relation is universally expected to be true for event horizons.

As an outlook, let us mention that here, we have considered a thermal field theory in equilibrium, dual to an Euclidean signature AdS-Schwarzschild black hole. For describing *transport phenomena* which require evolution in time and small deviations from equilibrium, it will be necessary to move on to Minkowski signature black holes.

3.5.2 Finite Density and Chemical Potential

To conclude our brief introduction to gauge/gravity duality, let us consider how to introduce a finite density and chemical potential.

3.5.2.1 Quantum Field Theory at Finite Density

Consider the Lagrangian of a quantum field theory with a $U(1)$ gauge symmetry and a scalar and a fermion charged under this symmetry

$$\mathscr{L} = (D_\mu\varphi)^* D^\mu\varphi + i\bar{\psi}\gamma^\mu D_\mu\psi + \frac{1}{g^2}F^{\mu\nu}F_{\mu\nu} \tag{3.185}$$

with the covariant derivative $D_\mu = \partial_\mu + iA_\mu$. Let's give the time-component of A_μ a non-vanishing value $\langle A_0\rangle = \mu$, such that $A_0 = \langle A_0\rangle + \delta A_0$. This generates a potential of the form

$$V = -\mu^2\varphi^*\varphi - \mu\psi^\dagger\psi\,, \tag{3.186}$$

where μ is the chemical potential, $\psi^\dagger\psi = \hat{N}$ gives the number operator and $-\mu^2$ is the mass square of the scalar field. Note that the scalar field has a negative square mass, which leads to an upside-down potential and potentially to instabilities.

Recall that the Gibbs energy (free energy in the grand canonical ensemble) is given by

$$\Omega = E - TS - \mu N\,. \tag{3.187}$$

3.5.2.2 Finite Density and Chemical Potential on the Gravity Side

In the following we will introduce finite density and chemical potential into the gauge/gravity duality. For this, we have to find a solution to the equation of motion of Einstein-Maxwell theory with $A = A_t(r)dt$. The background Maxwell $U(1)$ potential of the field theory is read off from the boundary values of the bulk Maxwell field $A_\mu(r) = A_\mu^{(0)} + \ldots$ as $r \to 0$. The Einstein equations of motion involve the energy-momentum tensor of the field strength $F_{\mu\nu}$,

$$R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} - \frac{d(d-1)}{2L^2}g_{\mu\nu} = \frac{\kappa^2}{g^2}\left(F_{\mu\lambda}F_\nu^\lambda - \frac{1}{4}g_{\mu\nu}F_{\lambda\rho}F^{\lambda\rho}\right). \tag{3.188}$$

The solution is given by the Reissner–Nordström AdS black hole, whose metric in Minkowski signature is

$$\begin{aligned} ds^2 &= \frac{L^2}{r^2}\left(-f(r)dt^2 + \frac{dr^2}{f(r)} + d\mathbf{x}^2\right), \\ f(r) &= 1 - \left(1 + \frac{r_H^2\mu}{\gamma^2}\right)\left(\frac{r}{r_H}\right)^d + \frac{r_H^2\mu^2}{\gamma}\left(\frac{r}{r_H}\right)^{2(d-1)}, \\ \text{with}\quad \gamma &= \frac{(d-1)L^2g^2}{(d-2)\kappa^2}, \end{aligned} \tag{3.189}$$

and the gauge field

$$A_0(r) = \mu\left(1 + \left(\frac{r}{r_{\mathrm{H}}}\right)^{d-2}\right). \tag{3.190}$$

This satisfies the boundary condition that $A_t(r)$ has to vanish at the horizon since ∂_t is not well-defined as a Killing vector there. Moreover, (3.190) identifies the μ parameter in the solution (3.189) with the chemical potential: In agreement with the standard AdS/CFT result for the asymptotic behaviour of gravity fields near the AdS boundary, asymptotically $A_0(r)$ gives the source and the VEV of the dual field theory operator. Here, these determine the chemical potential and the density, respectively. We readily identify the source term μ as the chemical potential. For identifying the density, we proceed as follows:

The temperature is again fixed by analytic continuation to the Euclidean regime and is given by

$$T = \frac{1}{4\pi\, r_{\mathrm{H}}}\left(d - \frac{(d-2)\, r_{\mathrm{H}}^2\, \mu^2}{\gamma^2}\right). \tag{3.191}$$

In the grand canonical ensemble, by evaluating the Euclidean action on the solution, we find the following Gibbs free energy

$$\Omega = -T\, \ln Z = -\frac{L^{d-1}}{2\,\kappa^2\, r_{\mathrm{H}}^d}\left(1 + \frac{r_{\mathrm{H}}^2\, \mu^2}{\gamma^2}\right) V_{d-1} = F\left(\frac{T}{\mu}\right) V_{d-1}\, T^d. \tag{3.192}$$

From this, we obtain the charge density

$$\rho = -\frac{1}{V_2}\,\frac{\partial\Omega}{\partial\mu} = \frac{2\, L^2\, \mu}{\kappa^2\, r_{\mathrm{H}}\, \gamma^2}\,, \tag{3.193}$$

without loss of generality in $d = 3$ dimensions.

3.6 Conclusion

In these lecture notes we have reviewed prerequisites for the AdS/CFT correspondence and introduced the correspondence itself. We went on to testing the correspondence via the calculation of correlation functions. Finally, we have considered more general models of gauge/gravity duality by considering gravity duals of field theories at finite temperature and density. With these generalizations we are now ready to consider applications of gauge/gravity duality to strongly coupled systems such as the quark-gluon plasma and particular condensed matter systems. But this is another story which will be told elsewhere.

Acknowledgements I am grateful to Martin Ammon, to Hai Ngo and to Oliver Schlotterer for help with preparing this manuscript. Moreover I would like to thank my tutors at the school, Martin Ammon, Viviane Grass, Shu Lin, Hai Ngo and Andy O'Bannon.

References

1. Maldacena, J.M.: The large N limit of superconformal field theories and supergravity. Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38** (1999) 1113][arXiv:hep-th/9711200]
2. Witten, E.: Anti-de sitter space and holography. Adv. Theor. Math. Phys. **2**, 253 (1998) [arXiv:hep-th/9802150]
3. Gubser, S.S., Klebanov, I.R., Polyakov, A.M.: Gauge theory correlators from non-critical string theory. Phys. Lett. B **428**, 105 (1998) [arXiv:hep-th/9802109]
4. Freedman, D.Z., Mathur, S.D., Matusis, A., Rastelli, L.: Correlation functions in the CFT $(d)/\mathrm{AdS}(d+1)$ correspondence. Nucl. Phys. B **546**, 96 (1999) [arXiv:hep-th/9804058]
5. Aharony, O., Gubser, S.S., Maldacena, J.M., Ooguri, H.: Oz Y., Large N field theories, string theory and gravity. Phys. Rept. **323**, 183 (2000) [arXiv:hep-th/9905111]
6. D'Hoker, E., Freedman, D.Z.: Supersymmetric gauge theories and the AdS/CFT correspondence, arXiv:hep-th/0201253
7. Petersen, J.L.: Introduction to the Maldacena conjecture on AdS/CFT. Int. J. Mod. Phys. A **14**, 3597 (1999) [arXiv:hep-th/9902131]
8. Maldacena, J.: The gauge/gravity duality, [arXiv:1106.6073 [hep-th]]
9. Son, D.T., Starinets, A.O.: Viscosity, Black Holes, and Quantum Field Theory. Ann. Rev. Nucl. Part. Sci. **57**, 95–118 (2007) [arXiv:0704.0240 [hep-th]]
10. Erdmenger, J., Evans, N., Kirsch, I., Threlfall, E.: Mesons in Gauge/Gravity Duals - A Review. Eur. Phys. J. A **35**, 81–133 (2008) [arXiv:0711.4467 [hep-th]]
11. Hartnoll, S.A.: Lectures on holographic methods for condensed matter physics. Class. Quant. Grav. **26**, 224002 (2009) [arXiv:0903.3246 [hep-th]]
12. Herzog, C.P.: Lectures on Holographic Superfluidity and Superconductivity. J. Phys. A **A42**, 343001 (2009) [arXiv:0904.1975 [hep-th]]
13. Osborn, H., Petkou, A.C.: Implications of conformal invariance in field theories for general dimensions. Annals Phys. **231**, 311–362 (1994) [hep-th/9307010]
14. Erdmenger, J., Osborn, H.: Conserved Currents and the Energy Momentum Tensor in Conformally Invariant Theories for General Dimensions. Nucl. Phys. B **483**, 431–474 (1997) [arXiv:hep-th/9605009]
15. Coleman, S.: Aspects of Symmetry, Selected Erice Lectures. Cambridge University Press, Cambridge (1988)
16. Lee, S., Minwalla, S., Rangamani, M., Seiberg, N.: Three point functions of chiral operators in D = 4, N=4 SYM at large N. Adv. Theor. Math. Phys. **2**, 697–718 (1998) [hep-th/9806074]
17. D'Hoker, E., Freedman, D.Z., Skiba, W.: Field theory tests for correlators in the AdS / CFT correspondence. Phys. Rev. D **59**, 045008 (1999) [hep-th/9807098]
18. D'Hoker, E., Freedman, D.Z., Mathur, S.D., Matusis, A., Rastelli, L.: Extremal correlators in the AdS / CFT correspondence, In: Shifman, M.A. (ed.) The many faces of the superworld, pp. 332–360 [hep-th/9908160]
19. Erdmenger, J., Pérez-Victoria, M.: Nonrenormalization of next-to-extremal correlators in N=4 SYM and the AdS / CFT correspondence. Phys. Rev. D **62**, 045008 (2000) [hep-th/9912250]
20. Kim, H.J., Romans, L.J., van Nieuwenhuizen, P.: The Mass Spectrum of Chiral $\mathcal{N}c = 2$, $D = 10$ Supergravity on S^5. Phys. Rev. D **32**, 389 (1985)
21. Beisert, N., Ahn, C., Alday, L.F., Bajnok, Z., Drummond, J.M., Freyhult, L., Gromov, N., Janik, R.A., et al.: Review of AdS/CFT integrability: An overview [arXiv:1012.3982 (hep-th)]

Chapter 4
Holography for Strongly Coupled Media

Dam Thanh Son

4.1 Motivation

Many problems of modern theoretical physics are related to strong coupling. One example is the problem of the hot and dense matter in QCD. The creation of hot QCD matter is the goal of relativistic heavy ion experiments, the most recent of which are RHIC and LHC. Although there are ample evidence that some form of matter with strong collective behavior is formed in ultra-relativistic heavy ion collisions, the theoretical problem of finding whether thermal equilibrium is achieved and at which temperature has still not been solved. (The problem can be made very sharp by imagining a world with very small electromagnetic fine structure constant so that nuclei can be very large. Can we make a quark gluon plasma by colliding very large nuclei at very high energy? What is the temperature of the system at thermal equilibration? We still do not have definite answer to these questions.) Assuming that system reaches equilibrium, one can ask questions about the properties of the thermal equilibrium state. While thermodynamics of the QGP at finite temperature and zero chemical potential can be studied by lattice methods, the latter becomes very inefficient in dealing with real time quantities, for example the viscosities. Current lattice methods are also incapable of treating QCD matter at finite chemical potential, a problem that hinders our understanding of the core of neutron stars.

Another example of a strong coupling problem is that of unitarity fermions (unitary Fermi gas). This system is that of nonrelativistic fermion interacting through a short-range potential fine tuned to resonance at threshold (see Sect. 4.6 for more discussion). The simplest version of the problem is the Bertsch problem: given a gas of spin-1/2 fermions, interacting with short-range interaction fine tuned to unitarity (defined below in the lectures), what are the properties of the ground state? This problem has

D. T. Son (✉)
Institute for Nuclear Theory, University of Washington,
Seattle, WA 98195-1550, USA
e-mail: dtson@u.washington.edu

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_4,

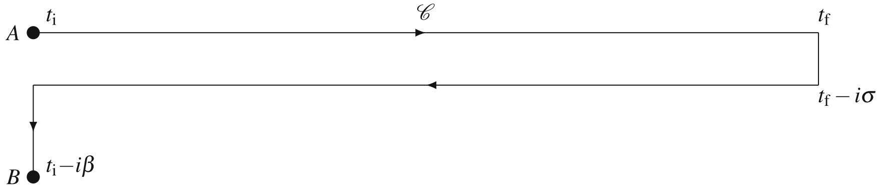


Fig. 4.1 The close time path contour

became extremely important when it became possible to realize unitarity fermions in atomic trap experiments.

Various other strong coupling problems in condensed matter physics are discussed in Subir Sachdev's lectures in this school. In these lectures, we will describe some points of contact between gauge/gravity duality and the physics of the quark gluon plasma and the unitary Fermi gas.

4.2 Thermal Field Theory

There are two main formalisms used in thermal field theory. The first formalism is the Matsubara, Euclidean formalism. It is used in lattice QCD, very convenient for thermodynamic and static quantities (like correlation length), but cannot directly address dynamic, real-time quantities. The second formalism is the real-time, close time path formalism. (For more details, see Refs. [1, 2]).

In the Matsubara formalism, the theory is formulated on a Euclidean spacetime, where the time axis is compactified to an interval $0 < \tau < \beta = 1/T$. In the close-time-path formalism, one makes a detour into real time, as in Fig. 4.1.

One can turn on source on the upper and lower parts of the contour, J_1 and J_2. The partition function of field theory now is a functional of both J_1 and J_2, $Z = Z[J_1, J_2]$, and and derivatives of $\log Z$ with respect to J gives a 2×2 matrix propagators G_{ab}, where $a, b = 1, 2$. Changing σ rescales the off-diagonal elements by a trivial factor,

$$G_{12}(\omega, q) = e^{\sigma\omega} G_{12}^{\sigma=0}(\omega, q), \tag{4.1}$$

$$G_{21}(\omega, q) = e^{-\sigma\omega} G_{21}^{\sigma=0}(\omega, q). \tag{4.2}$$

For $\sigma = 0$, the propagators G_{ab} include path-ordered, reversed path-ordered, and Wightmann Green's functions. They are related by

$$G_{11} + G_{22} = G_{12} + G_{21}, \tag{4.3}$$

but the choice $\sigma = \beta/2$ leads to symmetric 2×2 propagator matrix: $G_{12} = G_{21}$. This choice of the σ is most natural for holography, as we will see.

From the point of view of the CTP formalism, putting our system in an external source J corresponds to having, in the $\sigma = 0$ choice of the contour, $J_1 = J_2 = J$. The expectation value of the operator ϕ at a point x on an upper contour is given by an integral over the whole contour, which can be written as

$$\langle \phi_1(x) \rangle = -\int dy (G_{11}(x-y)J(y) - G_{12}(x-y)J(y)), \qquad \sigma = 0. \tag{4.4}$$

Define the retarded propagator $G_R = G_{11} - G_{12}$ ($\sigma = 0$). The retarded propagator governs the response of a system to a small external perturbation:

$$\langle \phi(x) \rangle = -\int dy G_R(x-y)J(y). \tag{4.5}$$

On the other hand, for the symmetric choice $\sigma = \beta/2,\ G_R = G_{11} - e^{-\beta\omega/2} G_{12}$.

Normally, the computations of thermal Green's function rely on summing Feynman diagrams. The set of Feynman diagrams that one has to sum in order to compute, say, the viscosity, can be quite complicated [3]. In the low-momentum limit, however, the forms of many correlation functions are simple and are dictated by an effective theory—hydrodynamics.

4.3 Hydrodynamics

Consider an interacting quantum field theory at finite temperature. One can visualize such a system as a collection of particles (or quasiparticles), moving with random velocities and colliding with each other from time to time. Such a picture is too simplistic for a strongly interacting system (with no discernible quasiparticles) but it does tell us that there is an important length scale in the problem—the mean free path, which is the length which a particle travels before colliding with other particles.

Hydrodynamics can be thought of as an effective theory describing the dynamics of a finite-temperature system at distance scales much larger than the mean free path. By definition the degrees of freedom entering hydrodynamics have to have relaxation time much larger than the mean free time. Such modes include

- Density of conserved quantities. Consider, for example, the QCD plasma, and imagine a perturbation of the system where there is a net excess of charge in a volume with size $L \gg \ell$. If one waits a long time this lump of excess charge will disappear, with the charge now distributing uniformly over the whole volume. However, since charge is conserved, causality implies that the time scale for this process cannot be smaller than L/c, where c is the speed of light (in fact in many cases the time scale is much bigger than this naive estimate. For example, if the relaxation is due to diffusion, the length scale is $\sim L^2/\ell$).
- Nambu–Goldstone modes. If there is a broken continuous symmetry, Goldstone's theorem dictates that there must be a massless particle at zero temperature. If the

symmetry remains broken at a finite temperature, the Nambu–Goldstone mode continuously deforms into a hydrodynamic mode. For example, in superfluid ^{4}He the phase of the condensate φ is a hydrodynamic degree of freedom (the superfluid velocity $\mathbf{v}_s$ is proportional to the gradient of $\varphi : \mathbf{v}_s = \nabla\varphi/m$).

- Unbroken U(1) gauge fields. At zero temperature, a U(1) gauge field which does not suffers from the Anderson-Higgs mechanism corresponds to a massless photon. At finite temperature, the electric field is screened (Debye screening) but the magnetic field is unscreened and should be included in the hydrodynamic description. An example of such a theory is magnetohydrodynamics, describing for example the interior of the Sun.

In these lectures we will consider only the simplest class of hydrodynamic theories, where the only slow degrees of freedom are the densities of conserved charges. In this case, hydrodynamics is given by the conservation equation,

$$\nabla_\mu T^{\mu\nu} = 0, \tag{4.6}$$

supplemented by the continuity equation that expresses $T^{\mu\nu}$ in terms of four variables: the local temperature T and the local fluid velocity u^μ:

$$T^{\mu\nu} = (\varepsilon + P)u^\mu u^\nu + Pg^{\mu\nu} + \tau^{\mu\nu}, \tag{4.7}$$

where $\tau^{\mu\nu}$ is the correction containing terms proportional to first derivatives. It is conventional to impose the condition $u_\mu \tau^{\mu\nu} = 0$ which eliminates any ambiguity in the definition of u^μ and T. In this case one has

$$\tau^{\mu\nu} = -\eta P^{\mu\alpha} P^{\nu\beta}(\nabla_\alpha u_\beta + \nabla_\beta u_\alpha) - \zeta P^{\mu\nu}(\nabla \cdot u), \tag{4.8}$$

where η and ζ are the shear and bulk viscosities, respectively. In a conformal plasma, the stress-energy tensor is traceless, hence $\varepsilon = 3P \sim T^4$ and $\zeta = 0$. In such a plasma, the shear viscosity has to scale with the temperature as $\eta \sim T^3$.

4.3.1 Hydrodynamics and Two Point Functions

From the hydrodynamic equations, one can easily compute the two-point functions of between two components of the stress-energy tensor. According to the general formulas of the linear response theory, the two-point function can be computed by first turning on a weak gravitational perturbation $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, $h_{\mu\nu} \ll 1$, then measuring the expectation value of the stress-energy tensor $\langle T_{\mu\nu}\rangle$. The two-point function is the coefficient of proportionality between $\langle T_{\mu\nu}\rangle$ and $h_{\mu\nu}$,

$$T_{\mu\nu}(x) \sim \int dy \langle T^{\mu\nu}(x) T^{\alpha\beta}(y)\rangle h_{\alpha\beta}(y). \tag{4.9}$$

On the other hand, when $h_{\mu\nu}$ varies with space and time very slowly, the response of the system can be determined by hydrodynamics. One first generalizes the hydrodynamic equation to curve spacetime. Assuming the system is in thermal equilibrium in the infinite past, and $h_{\mu\nu}$ is nonzero in a finite regime in spacetime, the state of the system can be completely determined.

We can re-derive the well known Kubo's formula in this way. Let us turn on a small metric perturbation whose only nonzero component is h_{xy} which is assumed to be homogeneous in space and is time dependent, $h_{xy} = h_{xy}(t)$. Then by symmetry one can right away determine that the fluid will remain in a state with constant temperature, $T = \text{const}$, and zero spatial velocity $u^{\mu} = (1, \mathbf{0})$ (a tensor perturbation cannot excite a scalar or vector mode to linear order). Nevertheless, the stress-energy tensor receives a correction

$$T^{xy} = Pg^{xy} - \eta(\nabla_x u_y + \nabla_y u_x) = -Ph_{xy} + 2\eta \Gamma^0_{xy} u_0 \tag{4.10}$$

proportional to the perturbation. Thus we find the two-point function

$$\langle T^{xy} T^{xy} \rangle = P - i\eta\omega. \tag{4.11}$$

The real part of this Green's function is a contact term, and depend on the way the two point function is defined; but one cannot get rid of the imaginary part by a redefinition of the Green function. Moreover, the imaginary part gives the value of the viscosity through the Kubo formula:

$$\eta = -\lim_{\omega \to 0} \frac{1}{\omega} \text{Im}\, G_R^{xy,xy}(\omega, \mathbf{0}). \tag{4.12}$$

4.4 AdS/CFT Prescription for Correlation Function

4.4.1 Euclidean Green's Function

Let us remind ourselves how the Euclidean Green's function is computed. For simplicity we limit ourselves to the case of an operator of dimension 4, dual to a massless scalar field ϕ. Assuming the action for the scalar field is

$$S = -\frac{K}{2} \int d^5x \sqrt{-g} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi. \tag{4.13}$$

Then the prescription tells us to solve the wave equation

$$\partial_\mu (\sqrt{-g} g^{\mu\nu} \partial_\nu \phi) = 0, \tag{4.14}$$

with boundary condition $\phi = \phi_0$ at the boundary. The solution, in momentum space, is $\phi(z,k) = \phi_0(k) f_k(z)$, where $f_k(z)$ is the solution to the field equation (at momentum k). We now rewrite S as a boundary action

$$S = \frac{K}{2} \int d^5x \frac{1}{z^3} \phi\phi'. \tag{4.15}$$

Differentiating the action with respect to the boundary value ϕ_0, we find the two point function to be

$$\langle\phi\phi\rangle_k \sim K \lim_{z\to 0} \frac{f_k'}{z^3}. \tag{4.16}$$

The boundary condition at the boundary needs to be supplemented by the boundary condition in the IR. At zero temperature, we require $\phi(z)$ to vanish as $z \to 0$. At finite temperature, spacetime is capped off at some $z = z_0$. We require the field to be regular at the horizon; in the case of the scalar, $\phi'(z_0) = 0$. The solution to the field equation is then unique, and the AdS/CFT prescription well defined.

4.4.2 *Real-Time Green's Function*

In real-time, the formulation of the AdS/CFT prescription is more subtle. The AdS/CFT rules are best formulated using the whole Penrose diagram of the black hole.

In the Poincare metric the AdS black hole looks like

$$ds^2 = -\frac{r^2}{R^2}(-f dt^2 + d\mathbf{x}^2) + \frac{R^2}{r^2 f} dr^2, \tag{4.17}$$

where $f = 1 - r_0^4/r^4$. The metric can be extended pass the horizon, one recovered four quadrants in the following Penrose diagram (Fig. 4.2).

Let us remind ourselves how it is done. Near the horizon, we expand $r = r_0 + \rho$. The (t, r) part of the metric can be rewritten as

$$ds^2 = 4\pi T\rho \left(-dt^2 + \frac{1}{(4\pi T)^2} \frac{d\rho^2}{\rho^2}\right), \tag{4.18}$$

where $T = r_0/\pi R^2$. This can be rewritten as $ds^2 = e^{4\pi T r_*}(-dt^2 + dr_*^2)$ where $r_* = (4\pi T)^{-1} \ln \rho$. Finally, we introduce Kruskal's coordinates

$$U = -e^{-2\pi T(t-r_*)}, \tag{4.19}$$

$$V = e^{2\pi T(t+r_*)}, \tag{4.20}$$

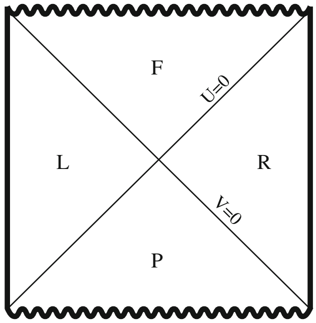


Fig. 4.2 Penrose diagram of AdS black hole

and metric is $ds^2 = -dUdV$. The Poincare coordinates cover only $U < 0,\ V > 0$ part quadrant of the diagram. There is another copy with the same metric, corresponding to the $U > 0,\ V < 0$ part. There are two boundaries.

The extension of the AdS/CFT duality was suggested by Maldacena [4], and then explicitly considered in Ref. [5]. The idea is that the two boundaries correspond to two horizontal parts of the close time path contour. The AdS/CFT prescription is then identifies the logarithm partition function of the thermal field theory, with sources J_1 and J_2 on the two parts of the contour, with the classical action of a configuration where the bulk field ϕ reaches the values J_1 and J_2 on the right and left boundaries, respectively.

In addition, one should also put boundaries conditions near the horizon. The choice of the boundary condition should be that when the bulk field ϕ is considered as function of the complex Kruskal coordinates U and V, it is analytic in the U upper half plane and V lower half plane.

The solution to the linearized field equation can be written in terms of the mode function $f_k(r)$, defined as the radial profile of a solution to the wave equation with momentum k, and is incoming wave at the horizon. One can write the solution down separately in the right and left quadrants,

$$\phi(k,r)|_R = \big((n+1)f_k^*(r_R) - nf_k(r_R)\big)\,\phi_1(k) + \sqrt{n(n+1)}\,\big(f_k(r_R) - f_k^*(r_R)\big)\,\phi_2(k), \tag{4.21}$$

$$\phi(k,r)|_L = \sqrt{n(n+1)}\,\big(f_k^*(r_L) - f_k(r_L)\big)\,\phi_1(k) + \big((n+1)f_k(r_L) - nf_k^*(r_L)\big)\,\phi_2(k). \tag{4.22}$$

Here $n = (e^{\omega/T} - 1)^{-1}$ is the Fermi–Dirac distribution function at frequency ω. Substituting the solution into the quadratic action, using the boundary form of the on-shell action,

$$\frac{K}{2}\int_R \sqrt{-g}g^{rr}\phi(-k,r)\partial_r\phi(k,r)\frac{d^4k}{(2\pi)^4} - \frac{K}{2}\int_L \sqrt{-g}g^{rr}\phi(-k,r)\partial_r\phi(k,r)\frac{d^4k}{(2\pi)^4} \tag{4.23}$$

(where K is a normalization factor) and differentiate it, one obtains the CTP propagators. Taking the appropriate linear combination of G_{11} and G_{12} we then find the retarded Green function,

$$G_R(k) = -K\sqrt{-g}g^{rr} f_k(r)\partial_r f_k^*(r)|_{r\to\infty}. \tag{4.24}$$

This formula coincides with the prescription first proposed in Ref. [6]. We now use this formula to compute the shear viscosity of the $\mathcal{N} = 4$ plasma.

4.4.3 Viscosity

Let us compute the a two-point function. We assume the momentum to be $q = (\omega, 0, 0, q)$, and we compute the two-point function T^{xy}, we consider gravitation perturbation with the only perturbation being $h_{xy}(t, z)$. One can show that the quadratic action of h_{xy} is that of a minimally coupled theory when written in terms of $\phi = g^{xx}h_{xy}$:

$$S = \frac{1}{2\kappa_{10}^2}\int d^{10}x\sqrt{-g}(R - 2\Lambda) = -\frac{V(S^5)}{4\kappa_1 0^2}\int d^5x\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi. \tag{4.25}$$

We now the write down the mode equation

$$\left(\frac{f(z)}{z^3}f_k'\right)' + \left(\frac{\omega^2}{z^3 f} - \frac{q^2}{z^3}\right)f_k(z) = 0. \tag{4.26}$$

The solution to this equation is

$$f_k(z) = \left(1 - \frac{z}{z_0}\right)^{-i\omega/4\pi T}. \tag{4.27}$$

Inserting the solution into the formula for G_R, we find the imaginary part of the retarded propagator,

$$\operatorname{Im} G_R(k) = -\frac{V(S^5)}{2\kappa_{10}^2}\frac{R^3}{z_0^3}i\omega. \tag{4.28}$$

To compute the real part of G_R one needs to be more careful with holographic renormalization. But we can already read out the viscosity from $\operatorname{Im} G_R$,

$$\eta = \frac{V(S^5)}{2\kappa_{10}^2}\frac{R^3}{z_0^3}. \tag{4.29}$$

This can be compared with the entropy density,

$$s = \frac{S}{V} = V(S^5)\frac{R^3}{z_0^3}\frac{2\pi}{\kappa_{10}^2}, \tag{4.30}$$

and we find that $\eta/s = 1/4\pi$. This is the common feature of all theories with gravitational dual [7].

4.5 Fluid-Gravity Correspondence: Diffusion

There is an alternative method to compute the kinetic coefficients. This method, sometimes called fluid-gravity correspondence, allows one to see directly the emergence of the *nonlinear* hydrodynamic equations from the field equations in the bulk [8]. The approach is thus complementary to standard AdS/CFT method based on the calculations of correlation functions.

We will illustrate the technique of fluid-gravity correspondence on a very simple example where the higher-dimensional theory is an abelian gauge theory in a black hole background,

$$S = -\frac{1}{4g_{\rm YM}^2}\int d^{d+1}x\, F_{\mu\nu}F^{\mu\nu}. \tag{4.31}$$

The background is chosen to be

$$ds^2 = r^2(-f(r)dt^2 + d\mathbf{x}^2) + \frac{dr^2}{r^2 f(r)}, \tag{4.32}$$

where $f(r)$ is a function that vanishes at the horizon, $f(r_0) = 0$ and tends to 1 at the AdS boundary, $f(\infty) = 1$. This is the usual back hole (black brane) background.

We are interested in solution to Maxwell's equations which satisfies outgoing wave boundary conditions. To enforce the incoming-wave boundary condition, it is more convenient to use the incoming Eddington–Fikelstein coordinates,

$$ds^2 = -r^2 f dv^2 + 2dvdr + r^2 d\mathbf{x}^2. \tag{4.33}$$

The usefulness of the Eddington–Fikelstein coordinates is that regularity at the horizon in these coordinates correspond to incoming wave boundary conditions in the usual coordinates. We go on to construct such a solution. The Maxwell equations are

$$\partial_r(r^d F_{vr}) + r^{d+2}\partial_i F_{ir} = 0, \tag{4.34}$$

$$r^d\partial_v F_{rv} + r^{d-2}\partial_i F_{iv} + f r^d \partial_i F_{ir} = 0, \tag{4.35}$$

$$\partial_r(r^{d-2}F_{vi}) + \partial_r(f r^d F_{ri}) + r^{d-2}\partial_v F_{ri} + r^{d-4}\partial_j F_{ji} = 0. \tag{4.36}$$

We start from gauge field of a charged black hole,

$$A_0 = \frac{q}{r^{d-1}}. \tag{4.37}$$

This is a one-parameter family of solutions to the Maxwell equations, parameterized by the charge density q. It describes a state in the field theory with a constant charge density in complete thermal equilibrium. Note that the solution is translationally invariant in all field-theory directions, t and $\mathbf{x}$.

What happens if we make q a function of the space and time? As one can easily verify, now the configuration (4.37) is not an exact solution to the Maxwell equation. However, when q varies slowly in space and time, one should be able to still find the solution by expanding it in powers of $\partial_t q$ and $\partial_x q$, which are small parameters. This is exactly the strategy that we will follow.

First we need to settle on a power counting scheme. Anticipating the end result to be a diffusion equation $\partial_t \rho = D\nabla^2 \rho$, we shall treat $\partial_i q$ as $O(\varepsilon)$ and $\partial_t q$ as $O(\varepsilon^2)$, where ε is small. We then expand the solutions, using the gauge $A_r = 0$,

$$A_v(r, x) = \frac{q(x)}{r^{d-2}} + A_v^{(1)}, \qquad A_i = A_i^{(1)}. \tag{4.38}$$

We can demand that $A_0^{(1)}$ falls off faster than $r^{-(d-2)}$ at large r. Otherwise, one can redefine $q(x)$ to absorb any $r^{-(d-2)}$ piece in $A_0^{(1)}$. Consistency requires that we treat $A_v^{(1)}$ as a quantity of order ε^2 and $A_i^{(1)} \sim \varepsilon$.

Substituting the ansatz in the the Maxwell equations, collecting terms with the same smallness in ε, we find

$$\partial_r(r^d \partial_r A_v^{(1)}) + r^{d+2}\partial_i \partial_r A_i^{(1)} = 0, \tag{4.39}$$

$$(d-1)\partial_v q + f r^d \partial_i \partial_r A_i^{(1)} - \frac{1}{r}\partial_i^2 q = 0, \tag{4.40}$$

$$-\partial_r \left(\frac{\partial_i q}{r}\right) + \partial_r(f r^d \partial_r A_i) = 0. \tag{4.41}$$

Integrating the last equation, we find

$$\partial_r A_i^{(1)} = \frac{C}{f r^d} + \frac{\partial_i q}{f r^{d+q}}, \tag{4.42}$$

where C is an x-dependent integration constant. Both terms in the right hand side have pole at the horizon $r = r_0$, and regularity at the horizon requires that the singularities cancel out between the two terms. Therefore we find

$$C = -r_0 \partial_i q. \tag{4.43}$$

Integrating once more, we then can find A_i. Actually for our purposes, we just need to know the asymptotic behavior of A_i at large r,

$$A_i = \frac{r_0}{d-1}\frac{\partial_i q}{r^{d-1}} + O(r^{-d}). \tag{4.44}$$

We can now substitute A_i into the second equation of (4.39), taking the large r limit and derive the following equation for q:

$$\partial_v q - \frac{r_0}{d-1}\nabla^2 q = 0, \tag{4.45}$$

which is nothing but the diffusion equation. Thus we have found a more general solution to the Maxwell equation which is parameterized by solutions to the diffusion equation. Maxwell equation in the background of black brane metric reduces to the diffusion equation in the long-wavelength limit.

4.6 Nonrelativistic Conformal Invariance

Fermions interacting through a unitarity interaction form a simplest nonrelativistic strongly interacting system. This system is beautiful because of its simplicity and universality. It has attracted enormous attention since being realized in cold atom experiments.

Let us first define fermions at unitarity. Consider two nonrelativistic particles, interacting through a potential,

$$H = \frac{\mathbf{p}_1^2}{2} + \frac{\mathbf{p}_2^2}{2} + V(|\mathbf{x}_1 - \mathbf{x}_2|). \tag{4.46}$$

For simplicity, we can consider V of the form of a square well potential, with size r_0 and depth $-V_0$: $V(r < r_0) = -V_0$ and $V(r > r_0) = 0$. If the potential is shallow, it does not have any bound state; but if it is deep enough it may have one, two, or more bound states. There is a critical value of $V_0 \sim r_0^{-2}$ at which the potential starts to have exactly one bound state. We tune V_0 to be exactly this value.

Then we take the range of interaction r_0 to zero, keeping V_0 always tuned to the critical value (in other words, keeping $V_0 r_0^2$ fixed). This limit is called the unitary limit, and the system of fermions interacting with this interaction the unitary Fermi gas.

The stability of such a system is not a trivial issue. It is relatively easy to see that for bosons, and for fermions of three or more different species, the finite-density system is not stable. This fact is related to the so-called Efimov effect: in the limit of zero range interaction, the Hamiltonian is unbounded from below (there is an infinite number of bound states, the lowest of which has an energy determined by the UV cutoff of the theory—the range of the potential).

4.6.1 Quantum Mechanics Formulation: Boundary Condition

The quantum mechanics of unitary fermions can be formulated in a way which gets rid of the interaction potential completely. Let us start with the case of two particles, one spin-up with coordinate **x**, and another spin-down with coordinate **y**. Neglecting the center of mass motion, the Schrödinger equation has the form

$$\frac{\partial^2}{\partial \mathbf{r}^2}\Psi(\mathbf{r}) + V(r)\Psi(\mathbf{r}) = -E\Psi(\mathbf{r}). \tag{4.47}$$

In the limit of zero range, the potential $V(r)$ is zero at any nonzero r. In the limit of $\mathbf{r} \to 0$, the right hand side can be neglected ($E \ll r^{-2}$), and we have the Laplace equation $\nabla_{\mathbf{r}}^2\Psi = 0$. Now it is known that the Laplace equation has two independent solutions, 1 and r^{-1}. The behavior of the wavefunction at small r is, in general,

$$\Psi(\mathbf{r}) = \frac{C}{r} + C_1 + O(r), \tag{4.48}$$

where C and C_1 are some numbers. In the usual problem of free particles, the wavefunction is assumed to be regular at $r = 0$, which means $C = 0$. On the other hand, from the mathematical point of view one can impose a general boundary condition

$$\Psi(\mathbf{r}) \sim \left(\frac{1}{r} - \frac{1}{a}\right), \tag{4.49}$$

with any value of a ($a = 0$ corresponding to free particles). Physically a is obtained by solving the zero-energy Schrödinger equation inside the potential $r < r_0$ and then match it to the solution to the Laplace equation outside the potential $r > r_0$; a therefore characterizes low-energy scatterings and is called the scattering length. The fine-tuning of the potential corresponds to the limit $a \to \infty$.

For the case of a general number of particles, the Hamiltonian is the sum of the kinetic terms of all particles,

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} = -\frac{1}{2m}\sum_i \frac{\partial^2}{\partial \mathbf{x}_i^2} \tag{4.50}$$

(where i numerates all particles) but the Hilbert space is nontrivial: the wave function of a system, $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N; y_1, \ldots, \mathbf{y}_N)$ satisfies the following condition when one spin-up and one spin-down particles approach each other,

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N; y_1, \ldots, \mathbf{y}_M) \to \frac{C}{|\mathbf{x}_i - y_j|} + O(1) + O(|\mathbf{x}_i - y_j|), \tag{4.51}$$

where $\mathbf{x}_i$ and $\mathbf{y}_j$ are the coordinates of the spin-up particles and spin-down particles, respectively. The Hamiltonian is trivial, but the nontriviality of the problem is in the Hilbert space.

For example, we can put a spin-up and a spin-down fermion in a harmonic potential. The Hamiltonian is now

$$H = \frac{1}{2}(\mathbf{p}_1^2 + p_2^2) + \frac{1}{2}\omega^2(\mathbf{x}_1^2 + \mathbf{x}_2^2). \tag{4.52}$$

The problem can be solved exactly even when the interaction between particles is unitary. The ground state is

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) \sim \frac{e^{-\omega(|\mathbf{x}_1|^2+|\mathbf{x}_2|^2)/2}}{|\mathbf{x}_1 - \mathbf{x}_2|}, \tag{4.53}$$

and the ground state energy is $E = 2\omega$. This is lower than the ground state energy in the case of a 3ω, in consistency with the attractiveness of the interaction.

The two-particle problem is special because it can be solved analytically. For three particles in a harmonic potentials, the energy levels are also known exactly (they are solutions to a trigonometric equation). For four particles and more (unless they have the same spin), the many-body problem cannot be solved exactly.

4.6.2 Symmetries of Unitary Fermions

A general nonrelativistic system is invariant under translation (in space and time), rotation, and Galilean boosts. In addition, the conserved mass (particle number) corresponds to a phase symmetry, $\psi \to e^{i\alpha}\psi$. These symmetries are enhanced to a new symmetry group called the Schrödinger group. The Schrödinger group contains two new symmetries

- Dilatation: $t \to \lambda^2 t,\ \mathbf{x} \to \lambda\mathbf{x}$,
- Proper conformal transformatio: $t = t/(1-\lambda t),\ \mathbf{x} \to \mathbf{x}/(1-\lambda t)$.

Saying that a theory theories has these symmetries means that if one has a solution to the time-dependent Schrödinger equation, $\Psi(t, \mathbf{x}_i)$, then one can generate new solutions. For example, the solution obtained by dilatation is

$$\Psi'(t, \mathbf{x}) = \lambda^{3N/2}\Psi(\lambda^2 t, \lambda x) \tag{4.54}$$

(the prefactor is to keep the normalization of Ψ). It is obvious that if Ψ solve the free Schrödinger equation, then Ψ' also does. More nontrivially, the boundary condition at short distances for unitary particle is preserved under dilatation. Similarly, the proper conformal transformation corresponds to the following family of new solutions,

$$\Psi'(t, \mathbf{x}) = C(t, \mathbf{x})\Psi\left(\frac{t}{1-\lambda t}, \frac{\mathbf{x}}{1-\lambda t}\right). \tag{4.55}$$

We leave the determination of $C(t, \mathbf{x})$ to the reader.

In the theory of unitarity fermions, the dilatation operator D and the proper conformal transformation C can be expressed in terms of the operators creating and annihilating a particle,

$$D = -\frac{i}{2}\int d\mathbf{x}, \mathbf{x}\cdot(\boldsymbol{\psi}^\dagger\nabla\boldsymbol{\psi} - \nabla\boldsymbol{\psi}^\dagger\boldsymbol{\psi}), \qquad C = \frac{1}{2}\int d\mathbf{x}, \mathbf{x}^2\boldsymbol{\psi}^\dagger(\mathbf{x})\boldsymbol{\psi}(\mathbf{x}). \tag{4.56}$$

One can check that the operators D, C and the Hamiltonian H form a close SO(2,1) subalgebra of the Schrödinger algebra,

$$[D, C] = -2iC, \quad [D, H] = 2iH, \quad [C, H] = iD. \tag{4.57}$$

The full Schrödinger algebra can be found in Ref. [9].

4.6.3 Local Operators

The local operators (for example $\boldsymbol{\psi}$, $\boldsymbol{\psi}^\dagger$, or $\boldsymbol{\psi}^\dagger\boldsymbol{\psi}$) depend on time T and space $\mathbf{x}$. Its commutators with time and space rotation are completely defined. The local operators can be classified by particle number by taking it commutator with the particle number operator $M = \int d^{\mathbf{x}}, \boldsymbol{\psi}^\dagger\boldsymbol{\psi}$. For example $\boldsymbol{\psi}$ has particle number -1 while for $\boldsymbol{\psi}^\dagger$ it is $+1$. Each operator can be associated with a dimension by

$$[D, O(0)] = i\Delta_O O(0). \tag{4.58}$$

For example, $\Delta_\psi = 3/2$ ($d/2$ in d spatial dimensions). One example of a nontrivial composite operator is obtained when one tries to construct the product of two annihilation operators of particles with different spins, $\boldsymbol{\psi}_\uparrow\boldsymbol{\psi}_\downarrow$. We know that the matrix element of $\boldsymbol{\psi}_\uparrow(\mathbf{x})\boldsymbol{\psi}_\downarrow(\mathbf{y})$ between a two-body state and vacuum is just the wave function,

$$\langle 0|\boldsymbol{\psi}_\uparrow(\mathbf{x})\boldsymbol{\psi}_\downarrow(\mathbf{y})|\Psi\rangle = \Psi(\mathbf{x}, \mathbf{y}). \tag{4.59}$$

The problem is that when we tries to take $\mathbf{x} \to \mathbf{y}$ to have a local operator $\psi_\uparrow\psi_\downarrow$, the matrix element diverges due to the boundary condition at $\mathbf{x} \to \mathbf{y}$. On the other hand, one can define the following operator

$$O_2(\mathbf{x}) = \lim_{\mathbf{y}\to\mathbf{x}} 4\pi|\mathbf{x} - \mathbf{y}|\boldsymbol{\psi}_\uparrow(\mathbf{x})\boldsymbol{\psi}_\downarrow(\mathbf{y}), \tag{4.60}$$

which has finite matrix elements between states in the Hamiltonian. Another way to write the equation above is

$$\boldsymbol{\psi}_\uparrow(\mathbf{x})\boldsymbol{\psi}_\downarrow(\mathbf{y}) = \frac{O_2(x)}{4\pi|\mathbf{x} - y|} + \cdots, \tag{4.61}$$

which has the form of an operator product expansion for unitarity fermions. Operator product expansions have been applied very recently for unitarity fermions; as in particle physics they are most useful at short distances. The operator $O_2^\dagger O_2$ has a special role: its expectation value is called in the literature the Tan's parameter, or the contact.

As in relativistic CFTs, one can introduce the notion of primary and descendant operators. Primary operators are called those which commute, at zero coordinates, with Galilean boosts and the proper conformal transformation: $[K_i, O(0)] = [C, O(0)] = 0$. By taking derivatives with respect to coordinates and time, descendants are obtained.

The SO(2,1) commutators are important to prove what we call the operator-state correspondence for systems with Schrödinger symmetry. Namely, a primary operator, which does not annihilate the vacuum (or its Hermitian conjugate does not annihilate the vacuum) can be put into correspondence with an eigenstate of the system of a few particles in a harmonic potential [9]. This statement can be proven by first noticing (recall the form of the operator C in Eq. (4.56) that the Hamiltonian in a harmonic potential can be written as

$$H_{\rm osc} = H + \omega^2 C. \tag{4.62}$$

Then for an primary operator O, one can construct a state $\Psi_O\rangle$ as follows,

$$|\Psi_O\rangle = e^{-H/\omega} O^\dagger(0)|0\rangle. \tag{4.63}$$

Physically, first we use $O^\dagger$ to create a state which is localized at the origin of coordinates, and then evolve that state in imaginary time using the free-space Hamiltonian during a time $1/\omega$. The resulting state, whose wavefunction is a Gaussian-type wavepacket, can be shown, by using the SO(2,1) commutators, to be an eigenstate of the Hamiltonian $H_{\rm osc}$ with energy 2ω.

This operator-state correspondence can be illustrated explicitly in a few example. The operator ψ has dimension $3/2$, which matches with the ground state energy of a single particle in an isotropic harmonic potential, $3\omega/2$. The operator O_2 has dimension 2, and corresponds to the ground state of a system of one spin up and one spin down particle in a harmonic potential, whose energy was shown above to be 2ω.

4.6.4 Schrödinger Space

If one wants to move in the direction of constructing a gravitational dual of the unitarity fermions, it seems reasonable to start by asking the question: what is the space-time that realizes the Schrödinger symmetry? (Recall that in the original Maldacena's duality, the symmetry of $AdS_5 \times S^5$ space matches with the symmetry of the quantum field theory). An example of such a space was constructed in Refs. [10, 11]. The space has *two* extra dimensions compared to one extra dimensions in standard holography. The metric is

$$ds^2 = -\frac{2(dx^+)^2}{z^4} + \frac{-2dx^+dx^- + dx^i dx^i + dz^2}{z^2}. \tag{4.64}$$

One can check that this spacetime has realizes all generators of the Schrödinger algebra as Killing vectors. In particular, the total mass and the proper conformal transformation are

$$
\begin{aligned}
&M: x^- \to x^- + a,\\
&C: z \to (1-ax^+)z, \quad x^i \to (1-ax^+)x^i, \quad x^+ \to (1-ax^+)x^+,\\
&\qquad x^- \to x^- - \frac{a}{2}(x^i x^i + z^2).
\end{aligned}
\tag{4.65}
$$

One can see that the translational symmetry along the direction x^- realizes the conservation of mass in the nonrelativistic theory. The simplest action which gives rise to the Schrödinger spacetime is that of Einstein gravity with negative cosmological constant, coupled to a massive gauge field with a suitably chosen mass.

Subsequently the five-dimensional Schrödinger spacetime (corresponding to two spatial directions in field theory) have been constructed in string theory. As by-product of the construction, one also found black hole solutions, which describe a medium with finite chemical potential and temperature. One might think that these solutions may be the first holographic model for the unitarity Fermi gas. Unfortunately, closer inspection reveals a serious undesirable feature: the equation of state of the black hole is $P(T,\mu) \sim \frac{T^4}{\mu^2}$, which has the correct scaling behavior but is more restrictive than required. The more general equation of state is

$$
P(T,\mu) = \mu^2 F(T/\mu), \tag{4.66}
$$

where the function F is not constrained. This is in contrast to the situation in relativistic holography, where fitting the equation of state of QCD is not really a problem in the bottom-up approach.

It seems that one should try to devise a more general way to realize Schrödinger symmetry. Attempts in this direction are being made. At the more general level, one should not expect the gravity dual of unitarity fermions to be a classical theory due to the lack of a large N parameter. One can generalize the unitarity fermions to a many-favor theory with Sp(2N) vector symmetry. This theory is trivial to solve; at large N the BCS theory becomes exact. The situation is very similar to the relativistic O(N) vector model. One can hope that there is a nonrelativistic high-spin theory that is dual to the Sp(2N) version of unitary fermions, similar to the case of the O(N) model [12]. As far as I know, to date no serious attempt has been made to uncover such a theory.

4.7 Summary

In these lectures we have considered some applications of gauge/gravity duality to systems with finite temperature and chemical potential. We have left out some very important applications of gauge gravity duality, most notably jet quenching and heavy quark energy loss.

Acknowledgements I thank the organizers of the Munich, Cargese and TASI schools for inviting me to give this series of lectures. This work is supported, in part, by the DOE grant No. DE-FG02-00ER41132.

References

1. Kapusta, J.I., Gale, C.: Finite-temperature field theory: principles and applications. Cambridge University Press, Cambridge (2006)
2. Le Bellac, M.: Thermal field theory. Cambridge University Press, Cambridge (2000)
3. Jeon, S., Yaffe, L.G.: From quantum field theory to hydrodynamics: transport coefficients and effective kinetic theory. Phys. Rev. D **53**, 5799 (1996) [arXiv:hep-ph/9512263]
4. Maldacena, J.M.: Eternal black holes in Anti-de-Sitter. JHEP **0304**, 021 (2003) [arXiv:hep-th/0106112]
5. Herzog, C.P., Son, D.T.: Schwinger-keldysh propagators from AdS/CFT correspondence. JHEP **0303**, 046 (2003) [arXiv:hep-th/0212072]
6. Son, D.T., Starinets, A.O.: Minkowski-space correlators in AdS/CFT correspondence: recipe and applications. JHEP **0209**, 042 (2002) [arXiv:hep-th/0205051]
7. Kovtun, P., Son, D.T., Starinets, A.O.: Viscosity in strongly interacting quantum field theories from black hole physics. Phys. Rev. Lett. **94**, 111601 (2005) [arXiv:hep-th/0405231]
8. Bhattacharyya, S., Hubeny, V.E., Minwalla, S., Rangamani, M.: Nonlinear fluid dynamics from gravity. JHEP **0802**, 045 (2008) [arXiv:0712.2456 [hep-th]]
9. Nishida, Y., Son, D.T.: Nonrelativistic conformal field theories. Phys. Rev. D **76**, 086004 (2007) [arXiv:0706.3746 [hep-th]]
10. Son, D.T.: Toward an AdS/cold atoms correspondence: a geometric realization of the Schroedinger symmetry. Phys. Rev. D **78**, 046003 (2008) [arXiv:0804.3972 [hep-th]].
11. Balasubramanian, K., McGreevy, J.: Gravity duals for non-relativistic CFTs. Phys. Rev. Lett. **101**, 061601 (2008) [arXiv:0804.4053 [hep-th]]
12. Klebanov, I.R., Polyakov, A.M.: AdS dual of the critical O(N) vector model. Phys. Lett. B **550**, 213 (2002) [arXiv:hep-th/0210114]

Chapter 5
Quantum Black Holes

Atish Dabholkar and Suresh Nampuri

5.1 Introduction

The entropy of black holes supplies us with very useful quantitative information about the fundamental degrees of freedom of quantum gravity. One of the important successes of string theory is that one can explain the thermodynamic entropy of certain supersymmetric black holes as a logarithm of the microscopic degeneracy as required by the Boltzmann relation. These results imply that at the quantum level, one should regard a black hole as an ensemble of quantum states in the Hilbert space of the theory.

In any consistent quantum theory of gravity such as string theory, the requirement that the thermodynamic entropy must equal the statistical entropy of a black hole is an extremely stringent theoretical constraint. This constraint is also *universal* in that it must hold in any 'phase' or compactification of the theory that admits a black hole. It is therefore a particularly useful guide in our explorations of string theory in the absence of direct experimental guidance, especially given the fact that we do not know which phase of the theory might describe the real world.

Much of the earlier work concerning quantum black holes has been in the limit of large charges when the area of the even horizon is also large. In recent years there has been substantial progress in understanding the entropy of supersymmetric

A. Dabholkar (✉)
Laboratoire de Physique Théorique et Hautes Energies (LPTHE),
Université Pierre et Marie Curie-Paris 6; CNRS UMR 7589,
Tour 13-14, 5^{éme} étage, Boite 126, 4 Place Jussieu,
75252 Paris Cedex 05, France
e-mail: atish@lpthe.jussieu.fr

S. Nampuri
Arnold Sommerfeld Center for Theoretical Physics,
Ludwig-Maximilians-Universität München,
Theresienstrasse 37, 80333 München, Germany
e-mail: suresh.nampuri@physik.uni-muenchen.de

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_5,

black holes within string theory going well beyond the large charge limit. It has now become possible to begin exploring finite size effects in perturbation theory in inverse size and even nonperturbatively, with highly nontrivial agreements between thermodynamics and statistical mechanics. Unlike the leading Bekenstein–Hawking entropy which follows from the two-derivative Einstein–Hilbert action, these finite size corrections depend sensitively on the 'phase' under consideration and contain a wealth of information about the details of compactification as well as the spectrum of nonperturbative states in the theory. Finite-size corrections are therefore very interesting as a valuable window into the microscopic degrees of freedom of the theory.

In these notes we describe recent progress in understanding finite size corrections to the black hole entropy. To simplify the discussion we consider the compactification of the heterotic string on $T^4 \times T^2$ which is dual to the compactifcation of Type-II string on $K_3 \times T^2$. This leads to a four-dimensional theory with $\mathcal{N} = 4$ supersymmetry and 22 vector multiplets. Our objective will be to understand the entropy of half-BPS and quarter-BPS black holes in this theory both from the thermodynamic and statistical view points. A lot is known about generalization of these results to other compactifications. There has also been more progress both in defining the quantum entropy using AdS/CFT correspondence and in computing it using localization. We will not discuss these recent topics here to keep the discussion simple and more accessible.

The organization is as follows. We review aspects of classical and semiclassical black holes in Sects. 5.2 and 5.3, and elements of string theory in Sect. 5.4. The microscopic counting is then described in Sects. 5.5 and 5.6 and the comparison with macroscopic entropy is discussed in Sect. 5.7. Relevant mathematical background is covered in Sect. 5.8.

These lecture notes are aimed at beginning graduate students but assume some basic background in general theory of relativity, quantum field theory, and string theory. A good introductory textbook on general relativity from a modern perspective see [1]. For a more detailed treatment see [2] which has become a standard reference among relativists, and [3] which remains a classic for various aspects of general relativity. For quantum field theory in curved spacetime see [4]. For relevant aspects of string theory see [5–8]. For additional details of some of the material covered here relating to $\mathcal{N} = 4$ dyons see [9].

These notes are based primarily on lectures delivered at the summer school 2010 in Munich on "Strings and Fundamental Physics." As well as at various lectures courses by AD on "Quantum Black Holes" taught at the Université Pierre et Marie Curie, Paris VI together with Ashoke Sen; at the 'School on D-Brane Instantons, Wall Crossing and Microstate Counting ' at the ICTP Trieste in 2010; at the "School on Black Objects in Supergravity" at the INFN, Frascati in 2010. Some of the material was used in earlier lecture courses by AD at Shanghai, CERN, Cargèse, and Seoul.

5.2 Classical Black Holes

A black hole is at once the most simple and the most complex object.

It is the most simple in that it is completely specified by its mass, spin, and charge. This remarkable fact is a consequence of a the so-called 'No Hair Theorem'. For an astrophysical object like the earth, the gravitational field around it depends not only on its mass but also on how the mass is distributed and on the details of the oblateness of the earth and on the shapes of the valleys and mountains. Not so for a black hole. Once a star collapses to form a black hole, the gravitational field around it forgets all details about the star that disappears behind the even horizon except for its mass, spin, and charge. In this respect, a black hole is very much like a structure-less elementary particle such as an electron.

And yet it is the most complex in that it possesses a huge entropy. In fact the entropy of a solar mass black hole is enormously bigger than the thermal entropy of the star that might have collapsed to form it. Entropy gives an account of the number of microscopic states of a system. Hence, the entropy of a black hole signifies an incredibly complex microstructure. In this respect, a black hole is very unlike an elementary particle.

Understanding the simplicity of a black hole falls in the realm of classical gravity. By the early seventies, full 50 years after Schwarzschild, a reasonably complete understanding of gravitational collapse and of the properties of an event horizon was achieved within classical general relativity. The final formulation began with the singularity theorems of Penrose, area theorems of Hawking and culminated in the laws of black hole mechanics.

Understanding the complex microstructure of a black hole implied by its entropy falls in the realm of quantum gravity and is the topic of present lectures. Recent developments have made it clear that a black hole is 'simple' not because it is like an elementary particle, but rather because it is like a statistical ensemble. An ensemble is also specified by a few conserved quantum numbers such as energy, spin, and charge. The simplicity of a black hole is no different than the simplicity that characterizes a thermal ensemble.

To understand the relevant parameters and the geometry of black holes, let us first consider the Einstein–Maxwell theory described by the action

$$\frac{1}{16\pi G}\int R\sqrt{g}d^4x - \frac{1}{16\pi}\int F^2\sqrt{g}d^4x, \tag{5.1}$$

where G is Newton's constant, $F_{\mu\nu}$ is the electro-magnetic field strength, R is the Ricci scalar of the metric $g_{\mu\nu}$. In our conventions, the indices μ, ν take values 0, 1, 2, 3 and the metric has signature $(-, +, +, +)$.

5.2.1 Schwarzschild Metric

Consider the Schwarzschild metric which is a spherically symmetric, static solution of the vacuum Einstein equations $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} = 0$ that follow from (5.1) when no electromagnetic fields are excited. This metric is expected to describe the spacetime outside a gravitationally collapsed non-spinning star with zero charge. The solution for the line element is given by

$$ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu = -\left(1 - \frac{2GM}{r}\right)dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega^2,$$

where t is the time, r is the radial coordinate, and Ω is the solid angle on a 2-sphere. This metric appears to be singular at $r = 2GM$ because some of its components vanish or diverge, $g_{00} \to \infty$ and $g_{rr} \to \infty$. As is well known, this is not a real singularity. This is because the gravitational tidal forces are finite or in other words, components of Riemann tensor are finite in orthonormal coordinates. To better understand the nature of this apparent singularity, let us examine the geometry more closely near $r = 2GM$. The surface $r = 2GM$ is called the 'event horizon' of the Schwarzschild solution. Much of the interesting physics having to do with the quantum properties of black holes comes from the region near the event horizon.

To focus on the near horizon geometry in the region $(r - 2GM) \ll 2GM$, let us define $(r - 2GM) = \xi$, so that when $r \to 2GM$ we have $\xi \to 0$. The metric then takes the form

$$ds^2 = -\frac{\xi}{2GM}dt^2 + \frac{2GM}{\xi}(d\xi)^2 + (2GM)^2 d\Omega^2, \tag{5.2}$$

up to corrections that are of order $\left(\frac{1}{2GM}\right)$. Introducing a new coordinate ρ,

$$\rho^2 = (8GM)\xi \quad \text{so that} \quad d\xi^2 \frac{2GM}{\xi} = d\rho^2,$$

the metric takes the form

$$ds^2 = -\frac{\rho^2}{16G^2M^2}dt^2 + d\rho^2 + (2GM)^2 d\Omega^2. \tag{5.3}$$

From the form of the metric it is clear that ρ measures the geodesic radial distance. Note that the geometry factorizes. One factor is a 2-sphere of radius $2GM$ and the other is the (ρ, t) space

$$ds_2^2 = -\frac{\rho^2}{16G^2M^2}dt^2 + d\rho^2. \tag{5.4}$$

We now show that this 1+1 dimensional spacetime is just a flat Minkowski space written in funny coordinates called the Rindler coordinates.

5.2.2 Rindler Coordinates

To understand Rindler coordinates and their relation to the near horizon geometry of the black hole, let us start with 1+1 Minkowski space with the usual flat Minkowski metric,

$$ds^2 = -dT^2 + dX^2. \tag{5.5}$$

In light-cone coordinates,

$$U = (T + X) \quad V = (T - X), \tag{5.6}$$

the line element takes the form

$$ds^2 = -dU\, dV. \tag{5.7}$$

Now we make a coordinate change

$$U = \frac{1}{\kappa} e^{\kappa u}, \quad V = -\frac{1}{\kappa} e^{-\kappa v}, \tag{5.8}$$

to introduce the Rindler coordinates (u, v). In these coordinates the line element takes the form

$$ds^2 = -dU\, dV = -e^{\kappa(u-v)} du\, dv. \tag{5.9}$$

Using further coordinate changes

$$u = (t + x), \quad v = (t - x), \quad \rho = \frac{1}{\kappa} e^{\kappa x}, \tag{5.10}$$

we can write the line element as

$$ds^2 = e^{2\kappa x}(-dt^2 + dx^2) = -\rho^2 \kappa^2 dt^2 + d\rho^2. \tag{5.11}$$

Comparing (5.4) with this Rindler metric, we see that the (ρ, t) factor of the Schwarzschild solution near $r \sim 2GM$ looks precisely like Rindler spacetime with metric

$$ds^2 = -\rho^2 \kappa^2\, dt^2 + d\rho^2 \tag{5.12}$$

with the identification

$$\kappa = \frac{1}{4GM}.$$

This parameter κ is called the surface gravity of the black hole. For the Schwarzschild solution, one can think of it heuristically as the Newtonian acceleration GM/r_H^2 at the horizon radius $r_H = 2GM$. Both these parameters–the surface gravity κ and the horizon radius r_H play an important role in the thermodynamics of black hole.

This analysis demonstrates that the Schwarzschild spacetime near $r = 2GM$ is not singular at all. After all it looks exactly like flat Minkowski space times a sphere of radius $2GM$. So the curvatures are inverse powers of the radius of curvature $2GM$ and hence are small for large $2GM$.

5.2.3 Exercises

5.2.3.1 Uniformly Accelerated Observer and Rindler Coordinates

Consider an astronaut in a spaceship moving with constant acceleration a in Minkowski spacetime with Minkowski coordinates $(T, \mathbf{X})$. This means she feels a constant normal reacting from the floor of the spaceship in her rest frame:

$$\frac{d^2\mathbf{X}}{dt^2} = \mathbf{a}, \quad \frac{dT}{d\tau} = 1 \tag{5.13}$$

where τ is proper time and $\mathbf{a}$ is the acceleration 3-vector.

1. *Write the equation of motion in a covariant form and show that her* 4*-velocity* $u^\mu := \frac{dX^\mu}{d\tau}$ *is timelike whereas her* 4*-acceleration* a^μ *is spacelike.*
2. *Show that if she is moving along the x direction, then her trajectory is of the form*

$$T = \frac{1}{a}\sinh(a\tau), \quad X = \frac{1}{a}\cosh(a\tau) \tag{5.14}$$

which is a hyperboloid. Find the acceleration 4-vector.

3. *Show that it is natural for her to use her proper time as the time coordinate and introduce a coordinate frame of a family of observers with*

$$T = \zeta\sinh(a\eta), \quad X = \zeta\cosh(a\eta). \tag{5.15}$$

By examining the metric, show that $v = \eta - \zeta$ and $u = \eta + \zeta$ are precisely the Rindler coordinates introduced earlier with the acceleration parameter a identified with the surface gravity κ.

5.2.4 Kruskal Extension

One important fact to note about the Rindler metric is that the coordinates u, v do not cover all of Minkowski space because even when the vary over the full range

$$-\infty \leq u \leq \infty, \quad -\infty \leq v \leq \infty$$

the Minkowski coordinate vary only over the quadrant

$$0 \leq U \leq \infty, \quad -\infty < V \leq 0. \tag{5.16}$$

If we had written the flat metric in these 'bad', 'Rindler-like' coordinates, we would find a fake singularity at $\rho = 0$ where the metric appears to become singular. But we can discover the 'good', Minkowski-like coordinates U and V and extend them to run from $-\infty$ to ∞ to see the entire spacetime.

Since the Schwarzschild solution in the usual (r, t) Schwarzschild coordinates near $r = 2GM$ looks exactly like Minkowski space in Rindler coordinates, it suggests that we must extend it in properly chosen 'good' coordinates. As we have seen, the 'good' coordinates near $r = 2GM$ are related to the Schwarzschild coordinates in exactly the same way as the Minkowski coordinates are related the Rindler coordinates.

In fact one can choose 'good' coordinates over the entire Schwarzschild spacetime. These 'good' coordinates are called the Kruskal coordinates. To obtain the Kruskal coordinates, first introduce the 'tortoise coordinate'

$$r^* = r + 2GM \log\left(\frac{r - 2GM}{2GM}\right). \tag{5.17}$$

In the (r^*, t) coordinates, the metric is conformally flat, i.e., flat up to rescaling

$$ds^2 = \left(1 - \frac{2GM}{r}\right)\left(-dt^2 + dr^{*2}\right). \tag{5.18}$$

Near the horizon the coordinate r^* is similar to the coordinate x in (5.11) and hence $u = t + r^*$ and $v = t - r^*$ are like the Rindler (u, v) coordinates. This suggests that we define U, V coordinates as in (5.8) with $\kappa = 1/4GM$. In these coordinates the metric takes the form

$$ds^2 = -e^{-(u-v)\kappa} dU\, dV = -\frac{2GM}{r} e^{-r/2GM} dU\, dV \tag{5.19}$$

We now see that the Schwarzschild coordinates cover only a part of spacetime because they cover only a part of the range of the Kruskal coordinates. To see the entire spacetime, we must extend the Kruskal coordinates to run from $-\infty$ to ∞. This extension of the Schwarzschild solution is known as the Kruskal extension.

Note that now the metric is perfectly regular at $r = 2GM$ which is the surface $UV = 0$ and there is no singularity there. There is, however, a real singularity at $r = 0$ which cannot be removed by a coordinate change because physical tidal forces become infinite. Spacetime stops at $r = 0$ and at present we do not know how to describe physics near this region.

5.2.5 Event Horizon

We have seen that $r = 2GM$ is not a real singularity but a mere coordinate singularity which can be removed by a proper choice of coordinates. Thus, locally there is nothing special about the surface $r = 2GM$. However, globally, in terms of the causal structure of spacetime, it is a special surface and is called the 'event horizon'. An event horizon is a boundary of region in spacetime from behind which no causal signals can reach the observers sitting far away at infinity.

To see the causal structure of the event horizon, note that in the metric (5.11) near the horizon, the constant radius surfaces are determined by

$$\rho^2 = \frac{1}{\kappa^2}e^{2\kappa x} = \frac{1}{\kappa^2}e^{\kappa u}e^{-\kappa v} = -UV = \text{constant} \tag{5.20}$$

These surfaces are thus hyperbolas. The Schwarzschild metric is such that at $r \gg 2GM$ and observer who wants to remain at a fixed radial distance r=constant is almost like an inertial, freely falling observers in flat space. Her trajectory is time-like and is a straight line going upwards on a spacetime diagram. Near $r = 2GM$, on the other hand, the constant r lines are hyperbolas which are the trajectories of observers in uniform acceleration.

To understand the trajectories of observers at radius $r > 2GM$, note that to stay at a fixed radial distance r from a black hole, the observer must boost the rockets to overcome gravity. Far away, the required acceleration is negligible and the observers are almost freely falling. But near $r = 2GM$ the acceleration is substantial and the observers are not freely falling. In fact at $r = 2GM$, these trajectories are light like. This means that a fiducial observer who wishes to stay at $r = 2GM$ has to move at the speed of light with respect to the freely falling observer. This can be achieved only with infinitely large acceleration. This unphysical acceleration is the origin of the coordinate singularity of the Schwarzschild coordinate system.

In summary, the surface defined by r = contant is timelike for $r > 2GM$, space-like for $r < 2GM$, and light-like or null at $r = 2GM$.

In Kruskal coordinates, at $r = 2GM$, we have $UV = 0$ which can be satisfied in two ways. Either V=0, which defines the 'future event horizon', or U=0, which defines the 'past event horizon'. The future event horizon is a one-way surface that signals can be sent into but cannot come out of. The region bounded by the event horizon is then a black hole. It is literally a hole in spacetime which is black because no light can come out of it. Heuristically, a black hole is black because even light cannot escape its strong gravitation pull. Our analysis of the metric makes this notion more precise. Once an observer falls inside the black hole she can never come out because to do so she will have to travel faster than the speed of light.

As we have noted already $r = 0$ is a real singularity that is inside the event horizon. Since it is a spacelike surface, once a observer falls insider the event horizon, she is sure to meet the singularity at $r = 0$ sometime in future no matter how much she boosts the rockets.

In our example of the Schwarzschild black hole, the event horizon is static because it is defined as a hypersurface $r = 2GM$ which does not change with time. More precisely, the time-like Killing vector $\frac{\partial}{\partial t}$ leaves it invariant. It is at the same time null because g^{rr} vanishes at $r = 2GM$. In general, as for a spinning Kerr–Newman black hole, the horizon is not static but only stationary and null. More precisely, a linear combination of the time-like Killing vector and a space-like vector leaves it invariant and moreover, the norm of this vector vanishes on the event horizon.

To summarize, an event horizon is a surface that is simultaneously *stationary* and *null*. Such a surface causally separates the inside and the outside of a black hole.

5.2.6 Black Hole Parameters

From our discussion of the Schwarzschild black hole we are ready to abstract some important general concepts that are useful in describing the physics of more general black holes.

To begin with, a *black hole* is an asymptotically flat spacetime that contains a region which is not in the backward lightcone of future timelike infinity. The boundary of such a region is a stationary null surface call the *event horizon*. The fixed t slice of the event horizon is a two sphere.

There are a number of important parameters of the black hole. We have introduced these in the context of Schwarzschild black holes. For a general black holes their actual values are different but for all black holes, these parameters govern the thermodynamics of black holes.

1. The radius of the event horizon r_H is the radius of the two sphere. For a Schwarzschild black hole, we have $r_H = 2GM$.
2. The area of the event horizon A_H is given by $4\pi r_H^2$. For a Schwarzschild black hole, we have $A_H = 16\pi G^2 M^2$.
3. The surface gravity is the parameter κ that we encountered earlier. As we have seen, for a Schwarzschild black hole, $\kappa = 1/4GM$.

5.2.7 Laws of Black Hole Mechanics

One of the remarkable properties of black holes is that one can derive a set of laws of black hole mechanics which bear a very close resemblance to the laws of thermodynamics. This is quite surprising because *a priori* there is no reason to expect that the spacetime geometry of black holes has anything to do with thermal physics.

(0) Zeroth Law: In thermal physics, the zeroth law states that the temperature T of a body at thermal equilibrium is constant throughout the body. Otherwise heat will flow from hot spots to the cold spots. Correspondingly for stationary black holes one can show that surface gravity κ is constant on the event horizon. This is obvious for spherically symmetric horizons but is true also more generally for non-spherical horizons of spinning black holes.

Table 5.1 Laws of black hole mechanics

Laws of thermodynamics	Laws of black hole mechanics
Temperature is constant throughout a body at equilibrium. T = constant.	Surface gravity is constant on the event horizon. κ = constant.
Energy is conserved. $dE = TdS + \mu dQ + \Omega dJ$.	Energy is conserved. $dM = \frac{\kappa}{8\pi} dA + \mu dQ + \Omega dJ$.
Entropy never decrease. $\Delta S \geq 0$.	Area never decreases. $\Delta A \geq 0$.

(1) First Law: Energy is conserved, $dE = TdS + \mu dQ + \Omega dJ$, where E is the energy, Q is the charge with chemical potential μ and J is the spin with chemical potential Ω. Correspondingly for black holes, one has $dM = \frac{\kappa}{8\pi G} dA + \mu dQ + \Omega dJ$. For a Schwarzschild black hole we have $\mu = \Omega = 0$ because there is no charge or spin.

(2) Second Law: In a physical process the total entropy S never decreases, $\Delta S \geq 0$. Correspondingly for black holes one can prove the area theorem that the net area in any process never decreases, $\Delta A \geq 0$. For example, two Schwarzschild black holes with masses M_1 and M_2 can coalesce to form a bigger black hole of mass M. This is consistent with the area theorem, since the area is proportional to the square of the mass, and $(M_1 + M_2)^2 \geq M_1^2 + M_2^2$. The opposite process where a bigger black hole fragments is however, disallowed by this law.

Thus the laws of black hole mechanics, crystallized by Bardeen, Carter, Hawking, and other bears a striking resemblance with the three laws of thermodynamics for a body in thermal equilibrium. We summarize these results below in Table 5.1 for a black hole of mass M, spin J, and charge Q.

Here A is the area of the horizon, and κ is the surface gravity which can be thought of roughly as the acceleration at the horizon, μ is the chemical potential conjugate to Q, and Ω is the angular speed conjugate to J.

We will see that this formal analogy between the laws of black hole mechanics and thermodynamics is actually much more than an analogy. Bekenstein and Hawking discovered that there is a deep connection between black hole geometry, thermodynamics and quantum mechanics. Quantum mechanically, a black hole is not quite black.

5.2.8 Historical Aside

Apart from its physical significance, the entropy of a black hole makes for a fascinating study in the history of science. It is one of the very rare examples where a scientific idea has gestated and evolved over several decades into an important conceptual and quantitative tool almost entirely on the strength of theoretical considerations. That we can proceed so far with any confidence at all with very little guidance from experiment is indicative of the robustness of the basic tenets of physics. It is therefore worthwhile to place black holes and their entropy in a broader context before coming to the more recent results pertaining to the quantum aspects of black holes within string theory.

A black hole is now so much a part of our vocabulary that it can be difficult to appreciate the initial intellectual opposition to the idea of 'gravitational collapse' of a star and of a 'black hole' of nothingness in spacetime by several leading physicists, including Einstein himself.

To quote the relativist Werner Israel,

> There is a curious parallel between the histories of black holes and continental drift. Evidence for both was already non-ignorable by 1916, but both ideas were stopped in their tracks for half a century by a resistance bordering on the irrational.

On January 16, 1916, barely two months after Einstein had published the final form of his field equations for gravitation [10], he presented a paper to the Prussian Academy on behalf of Karl Schwarzschild [11], who was then fighting a war on the Russian front. Schwarzschild had found a spherically symmetric, static and exact solution of the full nonlinear equations of Einstein without any matter present.

The Schwarzschild solution was immediately accepted as the correct description within general relativity of the gravitational field outside a spherical mass. It would be the correct approximate description of the field around a star such as our sun. But something much more bizzare was implied by the solution. For an object of mass M, the solution appeared to become singular at a radius $R = 2GM/c^2$. For our sun, for example, this radius, now known as the Schwarzschild radius, would be about 3 km. Now, as long the physical radius of the sun is bigger than 3 km, the 'Schwarzschild's singularity' is of no concern because inside the sun the Schwarzschild solution is not applicable as there is matter present. But what if the entire mass of the sun was concentrated in a sphere of radius smaller than 3 km? One would then have to face up to this singularity.

Einstein's reaction to the 'Schwarzschild singularity' was to seek arguments that would make such a singularity inadmissible. Clearly, he believed, a physical theory could not tolerate such singularities. This drove his to write as late as 1939, in a published paper,

> The essential result of this investigation is a clear understanding as to why the 'Schwarzschild singularities' do not exist in physical reality.

This conclusion was however, based on an incorrect argument. Einstein was not alone in this rejection of the unpalatable idea of a total gravitational collapse of a physical system. In the same year, in an astronomy conference in Paris, Eddington, one of the leading astronomers of the time, rubbished the work of Chandrasekhar who had concluded from his study of white dwarfs, a work that was to earn him the Nobel prize later, that a large enough star could collapse.

It is interesting that Einstein's paper on the inadmissibility of the Schwarzschild singularity appeared only two months before Oppenheimer and Snyder published their definitive work on stellar collapse with an abstract that read,

> When all thermonuclear sources of energy are exhausted, a sufficiently heavy star will collapse.

Once a sufficiently big star ran out of its nuclear fuel, then there was nothing to stop the inexorable inward pull of gravity. The possibility of stellar collapse meant that a star could be compressed in a region smaller than its Schwarzschild radius and the 'Schwarzschild singularity' could no longer be wished away as Einstein had desired. Indeed it was essential to understand what it means to understand the final state of the star.

It is thus useful to keep in mind what seems now like a mere change of coordinates was at one point a matter of raging intellectual debate.

5.3 Semiclassical Black Holes

In the semiclassical treatment of a black hole, we treat the spacetime geometry of the black hole classically but treat various fields such as the electromagnetic field in this fixed spacetime background quantum mechanically. This semiclassical inclusion of quantum effects already reveals a deep and unexpected connection between the spacetime geometry of a black hole and thermodynamics.

5.3.1 Hawking Temperature

Bekenstein asked a simple-minded but incisive question. If nothing can come out of a black hole, then a black hole will violate the second law of thermodynamics. If we throw a bucket of hot water into a black hole then the net entropy of the world outside would seem to decrease. Do we have to give up the second law of thermodynamics in the presence of black holes?

Note that the energy of the bucket is also lost to the outside world but that does not violate the first law of thermodynamics because the black hole carries mass or equivalently energy. So when the bucket falls in, the mass of the black hole goes up accordingly to conserve energy. This suggests that one can save the second law of thermodynamics if somehow the black hole also has entropy. Following this reasoning and noting the formal analogy between the area of the black hole and entropy discussed in the previous section, Bekenstein proposed that a black hole must have entropy proportional to its area [12].

This way of saving the second law is however, in contradiction with the classical properties of a black hole because if a black hole has energy E and entropy S, then it must also have temperature T given by

$$\frac{1}{T} = \frac{\partial S}{\partial E}.$$

For example, for a Schwarzschild black hole, the area and the entropy scales as $S \sim M^2$. Therefore, one would expect inverse temperature that scales as M

$$\frac{1}{T} = \frac{\partial S}{\partial M} \sim \frac{\partial M^2}{\partial M} \sim M. \tag{5.21}$$

Now, if the black hole has temperature then like any hot body, it must radiate. For a classical black hole, by its very nature, this is impossible.

Hawking showed that after including quantum effects, however, it is possible for a black hole to radiate [13]. In a quantum theory, particle-antiparticle are constantly being created and annihilated even in vacuum. Near the horizon, an antiparticle can fall in once in a while and the particle can escapes to infinity. In fact, Hawking's calculation showed that the spectrum emitted by the black hole is precisely thermal with temperature $T = \frac{\hbar\kappa}{2\pi} = \frac{\hbar}{8\pi GM}$. With this precise relation between the temperature and surface gravity the laws of black hole mechanics discussed in the earlier section become identical to the laws of thermodynamics. Using the formula for the Hawking temperature and the first law of thermodynamics

$$dM = TdS = \frac{\kappa\hbar}{8\pi G\hbar} dA,$$

one can then deduce the precise relation between entropy and the area of the black hole:

$$S = \frac{Ac^3}{4G\hbar}.$$

Before discussing the entropy of a black hole, let us derive the Hawking temperature in a somewhat heuristic way using a Euclidean continuation of the near horizon geometry. In quantum mechanics, for a system with Hamiltonian H, the thermal partition function is

$$Z = \mathrm{Tr} e^{-\beta \hat{H}}, \tag{5.22}$$

where β is the inverse temperature. This is related to the time evolution operator $e^{-itH/\hbar}$ by a Euclidean analytic continuation $t = -i\tau$ if we identify $\tau = \beta\hbar$. Let us consider a single scalar degree of freedom Φ, then one can write the trace as

$$\mathrm{Tr} e^{-\tau\hat{H}/\hbar} = \int d\phi \langle\phi| e^{-\tau_E \hat{H}/\hbar} |\phi\rangle$$

and use the usual path integral representation for the propagator to find

$$\mathrm{Tr} e^{-\tau\hat{H}/\hbar} = \int d\phi \int D\Phi e^{-S_E[\Phi]}.$$

Here $S_E[\Phi]$ is the Euclidean action over periodic field configurations that satisfy the boundary condition

$$\Phi(\beta\hbar) = \Phi(0) = \phi.$$

This gives the relation between the periodicity in Euclidean time and the inverse temperature,

$$\beta\hbar = \tau \quad \text{or} \quad T = \frac{\hbar}{\tau}. \tag{5.23}$$

Let us now look at the Euclidean Schwarzschild metric by substituting $t = -it_E$. Near the horizon the line element (5.11) looks like

$$ds^2 = \rho^2\kappa^2 dt_E^2 + d\rho^2.$$

If we now write $\kappa t_E = \theta$, then this metric is just the flat two-dimensional Euclidean metric written in polar coordinates provided the angular variable θ has the correct periodicity $0 < \theta < 2\pi$. If the periodicity is different, then the geometry would have a conical singularity at $\rho = 0$. This implies that Euclidean time t_E has periodicity $\tau = \frac{2\pi}{\kappa}$. Note that far away from the black hole at asymptotic infinity the Euclidean metric is flat and goes as $ds^2 = d\tau_E^2 + dr^2$. With periodically identified Euclidean time, $t_E \sim t_E + \tau$, it looks like a cylinder. Near the horizon at $\rho = 0$ it is nonsingular and looks like flat space in polar coordinates for this correct periodicity. The full Euclidean geometry thus looks like a cigar. The tip of the cigar is at $\rho = 0$ and the geometry is asymptotically cylindrical far away from the tip.

Using the relation between Euclidean periodicity and temperature, we then conclude that Hawking temperature of the black hole is

$$T = \frac{\hbar\kappa}{2\pi}. \tag{5.24}$$

5.3.2 Bekenstein–Hawking Entropy

Even though we have "derived" the temperature and the entropy in the context of Schwarzschild black hole, this beautiful relation between area and entropy is true quite generally essentially because the near horizon geometry is always Rindler-like. For *all* black holes with charge, spin and in number of dimensions, the Hawking temperature and the entropy are given in terms of the surface gravity and horizon area by the formulae

$$T_H = \frac{\hbar\kappa}{2\pi}, \quad S = \frac{A}{4G\hbar}.$$

This is a remarkable relation between the thermodynamic properties of a black hole on one hand and its geometric properties on the other.

The fundamental significance of entropy stems from the fact that even though it is a quantity defined in terms of gross thermodynamic properties, it contains nontrivial information about the *microscopic* structure of the theory through Boltzmann relation

$$S = k \log(d),$$

where d is the the degeneracy or the total number of microstates of the system of for a given energy, and k is Boltzmann constant. Entropy is not a kinematic quantity like energy or momentum but rather contains information about the total number microscopic degrees of freedom of the system. Because of the Boltzmann relation, one can learn a great deal about the microscopic properties of a system from its thermodynamics properties.

The Bekenstein–Hawking entropy behaves in every other respect like the ordinary thermodynamic entropy. It is therefore natural to ask what microstates might account for it. Since the entropy formula is given by this beautiful, general form

$$S = \frac{Ac^3}{4G\hbar},$$

that involves all three fundamental dimensionful constants of nature, it is a valuable piece of information about the degrees of freedom of a quantum theory of gravity.

5.3.3 Exercises

5.3.3.1 Reissner–Nordström (RN) Black Hole

The most general static, spherically symmetric, charged solution of the Einstein–Maxwell theory (5.1) gives the Reissner–Nordström (RN) black hole. In what follows we choose units so that $G = \hbar = 1$. The line element is given by

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 d\Omega^2, \qquad (5.25)$$

and the electromagnetic field strength by

$$F_{tr} = Q/r^2.$$

The parameter Q is the charge of the black hole and M is the mass. For $Q = 0$, this reduces to the Schwarzschild black hole.

From the metric (5.25) we see that the event horizon for this solution is located at where $g^{rr} = 0$, or

$$1 - \frac{2M}{r} + \frac{Q^2}{r^2} = 0.$$

Since this is a quadratic equation in r,

$$r^2 - 2QMr + Q^2 = 0,$$

it has two solutions.

$$r_\pm = M \pm \sqrt{M^2 - Q^2}.$$

Thus, r_+ defines the outer horizon of the black hole and r_- defines the inner horizon of the black hole. The area of the black hole is $4\pi r_+^2$.

1. *Identify the horizon for this metric and examine the near horizon geometry to show that it has two-dimensional Rindler spacetime as a factor.*
2. *Using the relation to the Rindler geometry determine the surface gravity κ as for the Schwarzschild black hole and thereby determine the temperature and entropy of the black hole.*

$$\begin{aligned} T =& \frac{\kappa\hbar}{2\pi} = \frac{\sqrt{M^2 - Q^2}}{2\pi(2M(M + \sqrt{M^2 - Q^2}) - Q^2)} \\ S =& \pi r_+^2 = \pi(M + \sqrt{M^2 - Q^2})^2. \end{aligned}$$

Recover the formulae for Schwarzschild black hole in the limit $Q = 0$.

3. *Show that in the extremal limit $M \to Q$ the temperature vanishes but the entropy has a nonzero limit. Show that for the extremal Reissner–Nordström black hole the near horizon geometry is of the form $AdS_2 \times S^2$.*

5.3.4 Bekenstein–Hawking–Wald Entropy

In our discussion of Bekenstein–Hawking entropy of a black hole, the Hawking temperature could be deduced from surface gravity or alternatively the periodicity of the Euclidean time in the black hole solution. These are geometric asymptotic properties of the black hole solution. However, to find the entropy we needed to use the first law of black hole mechanics which was derived in the context of Einstein–Hilbert action

$$\frac{1}{16\pi}\int R\sqrt{g}d^4x.$$

Generically in string theory, we expect corrections (both in α' and g_s) to the effective action that has higher derivative terms involving Riemann tensor and other fields.

$$I = \frac{1}{16\pi}\int (R + R^2 + R^4F^4 + \cdots).$$

How do the laws of black hole thermodynamics get modified?

Wald derived the first law of thermodynamics in the presence of higher derivative terms in the action [14–16]. This generalization implies an elegant formal expression for the entropy S given a general action I including higher derivatives

$$S = 2\pi \int_{\rho^2} \frac{\delta I}{\delta R_{\mu\nu\alpha\beta}} \varepsilon^{\mu\alpha} \varepsilon^{\nu\beta} \sqrt{h} d^2 \Omega,$$

where $\varepsilon^{\mu\nu}$ is the binormal to the horizon, h the induced metric on the horizon, and the variation of the action with respect to $R_{\mu\nu\alpha\beta}$ is to be carried out regarding the Riemann tensor as formally independent of the metric $g_{\mu\nu}$.

As an example, let us consider the Schwarzschild solution of the Einstein Hilbert action. In this case, the event horizon is S^2 which has two normal directions along r and t. We can construct an antisymmetric 2-tensor $\varepsilon_{\mu\nu}$ along these directions so that $\varepsilon_{rt} = \varepsilon_{tr} = -1$.

$$\mathscr{L} = \frac{1}{16\pi} R_{\mu\nu\alpha\beta} g^{\nu\alpha} g^{\mu\beta}, \quad \frac{\partial \mathscr{L}}{\partial R_{\mu\nu\alpha\beta}} = \frac{1}{16\pi} \frac{1}{2} (g^{\mu\alpha} g^{\nu\beta} - g^{\nu\alpha} g^{\mu\beta})$$

Then the Wald entropy is given by

$$\begin{aligned} S &= \frac{1}{8} \int \frac{1}{2} (g^{\mu\alpha} g^{\nu\beta} - g^{\nu\alpha} g^{\mu\beta}) (\varepsilon_{\mu\nu} \varepsilon_{\alpha\beta}) \sqrt{h} d^2 \Omega \\ &= \frac{1}{8} \int g^{tt} g^{rr} \cdot 2 = \frac{1}{4} \int_{S^2} \sqrt{h} d^2 \Omega = \frac{A_H}{4}, \end{aligned}$$

giving us the Bekenstein–Hawking formula as expected.

5.3.5 *Extremal Black Holes*

For a physically sensible definition of temperature and entropy in (5.26) the mass must satisfy the bound $M^2 \geq Q^2$. Something special happens when this bound is saturated and $M = |Q|$. In this case $r_+ = r_- = |Q|$ and the two horizons coincide. We choose Q to be positive. The solution (5.25) then takes the form,

$$ds^2 = -(1 - Q/r)^2 dt^2 + \frac{dr^2}{(1 - Q/r)^2} + r^2 d\Omega^2, \tag{5.26}$$

with a horizon at $r = Q$. In this extremal limit (5.26), we see that the temperature of the black hole goes to zero and it stops radiating but nevertheless its entropy has a finite limit given by $S \to \pi Q^2$. When the temperature goes to zero, thermodynamics does not really make sense but we can use this limiting entropy as the definition of the zero temperature entropy.

For extremal black holes it is sometimes more convenient to use isotropic coordinates in which the line element takes the form

$$ds^2 = H^{-2}(\mathbf{x})dt^2 + H^2(\mathbf{x})d\mathbf{x}^2$$

where $d\mathbf{x}^2$ is the flat Euclidean line element $\delta_{ij}dx^i dx^j$ and $H(\mathbf{x})$ is a harmonic function of the flat Laplacian

$$\delta^{ij}\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}.$$

The extremal Reissner–Nordström solution is obtained by choosing

$$H(\mathbf{x}) = \left(1 + \frac{Q}{\rho}\right),$$

and the field strength is given by $F_{0i} = \partial_i H(\mathbf{x})$.

One can in fact write a multi-centered Reissner–Nordström solution by choosing a more general harmonic function

$$H = 1 + \sum_{i=1}^{N} \frac{Q_i}{|\mathbf{x} - \mathbf{x}i|}. \tag{5.27}$$

The total mass M equals the total charge Q and is given additively

$$Q = \sum Q_i. \tag{5.28}$$

The solution is static because the electrostatic repulsion between different centers balances the gravitational attraction between them.

Note that the coordinate ρ in the isotropic coordinates should not be confused with the coordinate r in the spherical coordinates. In the isotropic coordinates the line-element is

$$ds^2 = -\left(1 + \frac{Q}{\rho}\right)^2 dt^2 + (1 + \frac{Q}{\rho})^{-2}\left(d\rho^2 + \rho^2 d\Omega^2\right),$$

and the horizon occurs at $\rho = 0$. Contrast this with the metric in the spherical coordinates (5.26) that has the horizon at $r = Q$. The near horizon geometry is quite different from that of the Schwarzschild black hole. The line element is

$$\begin{aligned} ds^2 &= -\frac{\rho^2}{Q^2}dt^2 + \frac{Q^2}{\rho^2}(d\rho^2 + \rho^2 d\Omega^2) \\ &= \left(-\frac{\rho^2}{Q^2}dt^2 + \frac{Q^2}{\rho^2}dr^2\right) + (Q^2 d\Omega^2). \end{aligned}$$

The geometry thus factorizes as for the Schwarzschild solution. One factor the 2-sphere S^2 of radius Q but the other (r, t) factor is now not Rindler any more but is a two-dimensional Anti-de Sitter or AdS_2. The geodesic radial distance in AdS_2 is

$\log r$. As a result the geometry looks like an infinite throat near $r = 0$ and the radius of the mouth of the throat has radius Q.

Extremal black holes are interesting because they are stable against Hawking radiation and nevertheless have a large entropy. We now try to see if the entropy can be explained by counting of microstates. In doing so, supersymmetry proves to be a very useful tool.

5.3.6 Wald Entropy for Extremal Black Holes

The horizon of extremal black holes has additional symmetries. For non-spinning black holes, the geometry is spherically symmetric. At extremality, the near horizon geometry becomes $AdS_2 \times S^2$ just as in the case of Reissner–Nordström black hole. The formula for the Wald entropy can be simplified considerably by exploiting these symmetries [17, 18].

The Reissner–Nordström metric is

$$ds^2 = -(1 - r_+/r)(1 - r_-/r)dt^2 + \frac{dr^2}{(1 - r_+/r)(1 - r_-/r)} + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{5.29}$$

Here (t, r, θ, ϕ) are the coordinates of space-time and r_+ and r_- are two parameters labelling the positions of the outer and inner horizon of the black hole respectively $(r_+ > r_-)$. The extremal limit corresponds to $r_- \to r_+$. We take this limit keeping the coordinates $\theta,\ \phi$, and

$$\sigma := \frac{(2r - r_+ - r_-)}{(r_+ - r_-)}, \quad \tau := \frac{(r_+ - r_-)t}{2r_+^2}, \tag{5.30}$$

fixed. In this limit the metric and the other fields take the form:

$$ds^2 = r_+^2 \left(-(\sigma^2 - 1)d\tau^2 + \frac{d\sigma^2}{\sigma^2 - 1}\right) + r_+^2 \left(d\theta^2 + \sin^2(\theta)d\phi^2\right). \tag{5.31}$$

This is the metric of $AdS_2 \times S^2$, with AdS_2 parametrized by (σ, τ) and S^2 parametrized by (θ, ϕ). Although in the original coordinate system the horizons coincide in the extremal limit, in the (σ, τ) coordinate system the two horizons are at $\sigma = \pm 1$. The AdS_2 space has $SO(2, 1) \equiv SL(2, \mathbb{R})$ symmetry—the time translation symmetry is enhanced to the larger $SO(2, 1)$ symmetry. All known extremal black holes have this property. Henceforth, we will take this as a definition of the near horizon geometry of an extremal black hole. In four dimensions, we also have the S^2 factor with $SO(3)$ isometries. Our objective will be to exploit the $SO(2, 1) \times SO(3)$ isometries of this spacetime to considerably simply the formula for Wald entropy.

Consider an arbitrary theory of gravity in four spacetime dimensions with metric $g_{\mu\nu}$ coupled to a set of $U(1)$ gauge fields $A^{(i)}_\mu$ ($i = 1, \ldots, r$ for a rank r gauge group) and neutral scalar fields ϕ_s $(s = 1, \ldots N)$. Let x^μ ($\mu = 0, \ldots, 3$ be local coordinates on spacetime and $\mathscr{L}$ be an arbitrary general coordinate invariant local lagrangian. The action is then

$$I = \int d^4x \sqrt{-det(g)}\mathscr{L}. \tag{5.32}$$

For an extremal black hole solution of this action, the most general form of the near horizon geometry and of all other fields consistent with $SO(2, 1) \times SO(3)$ isometry is given by

$$ds^2 = v_1 \left(-(\sigma^2 - 1)d\tau^2 + \frac{d\sigma^2}{\sigma^2 - 1} \right) + v_2(d\theta^2 + \sin^2(\theta)d\phi^2), \tag{5.33}$$

$$F^{(i)}_{\sigma\tau} = e_i, \quad F^{(i)}_{\theta\phi} = \frac{p_i}{4\pi} \sin(\theta), \quad \phi_s = u_s. \tag{5.34}$$

We can think of e_i and p_i $(i = 1, \ldots, r)$ as the electric and magnetic fields respectively near the black hole horizon. The constants v_a $(a = 1, 2)$ and $u_s(s = 1, \ldots, N)$ are to be determined by solving the equations of motion. Let us define

$$f(u, v, e, p) := \int d\theta d\phi \sqrt{-\det(g)}\mathscr{L}|_{horizon}. \tag{5.35}$$

Using the fact that $\sqrt{-\det(g)} = \sin(\theta)$ on the horizon, we conclude

$$f(u, v, e, p) := 4\pi v_1 v_2 \mathscr{L}|_{horizon} \tag{5.36}$$

Finally we define the entropy function

$$\mathscr{E}(q, u, v, e, p) = 2\pi(e_i q_i - f(u, v, e, p)), \tag{5.37}$$

where we have introduced the quantities

$$q_i := \frac{\partial f}{\partial e_i} \tag{5.38}$$

which by definition can be identified with the electric charges carried by the black hole. This function called the 'entropy function' is directly related to the Wald entropy as we summarize below.

1. For a black hole with fixed electric charges $\{q_i\}$ and magnetic charges $\{p_i\}$, all near horizon parameters v, u, e are determined by extremizing $\mathscr{E}$ with respect to the near horizon parameters:

$$\frac{\partial \mathscr{E}}{\partial e_i} = 0 \quad i = 1, \ldots r; \tag{5.39}$$

$$\frac{\partial \mathscr{E}}{\partial v_a} = 0, \quad a = 1, 2; \tag{5.40}$$

$$\frac{\partial \mathscr{E}}{\partial u_s} = 0, \quad s = 1, \dots N. \tag{5.41}$$

Equation (5.39) is simply the definition of electric charge whereas the other two equations (5.40) and (5.41) are the equations of motion for the near horizon fields. This follows from the fact that the dependence of $\mathscr{E}$ on all the near horizon parameters other than e_i comes only through $f(u, v, e, p)$ which from (5.36) is proportional to the action near the horizon. Thus extremization of the near horizon action is the same as the extremization of $\mathscr{E}$. This determines the variables (u, v, e) in terms of (q, p) and as a result the value of the entropy function at the extremum $\mathscr{E}^*$ is a function only of the charges

$$\mathscr{E}^*(q, p) := \mathscr{E}(q, u^*(q, p), v^*(q, p), e^*(q, p), p). \tag{5.42}$$

2. Once we have determined the near horizon geometry, we can find the entropy using Wald's formula specialized to the case of external black holes:

$$S_{wald} = -8\pi \int d\theta d\phi \frac{\partial S}{\partial R_{rtrt}} \sqrt{-g_{rr} g_{tt}}. \tag{5.43}$$

   With some algebra it is easy to see that the entropy is given by the value of the entropy function at the extremum:

$$S_{wald}(q, p) = \mathscr{E}^*(q, p). \tag{5.44}$$

This 'entropy function formalism' described above allows one to compute the entropy of various extremal black holes very efficiently by simply solving certain algebraic equations (instead of partial differential equations). It also allows one to incorporate effects of higher derivative corrections to the two-derivative action with relative ease.

5.3.6.1 Wald Entropy for a Reissner–Nordström Black Hole

To illustrate the use of the entropy function formalism for concrete computations, consider the Einstein–Maxell theory given by the action (5.1) and a solution given by

$$\begin{aligned} ds^2 &= v_1 \left(-(\sigma^2 - 1) d\tau^2 + \frac{d\sigma^2}{\sigma^2 - 1} \right) + v_2 \left(d\theta^2 + \sin^2(\theta) d\phi^2 \right) \\ F_{\sigma\tau} &= e, \quad F_{\theta\phi} = \frac{p}{4\pi} \sin(\theta) \end{aligned} \tag{5.45}$$

Substituting into the action we obtain the entropy function

$$\begin{aligned} \mathscr{E}(q, v, e, q, p) &\equiv 2\pi \, (e_i q_i - f(v, e, p)) \\ &= 2\pi \left[eq - 4\pi v_1 v_2 \left\{ \frac{1}{16\pi} \left(-\frac{2}{v_1} + \frac{2}{v_2} \right) + \frac{1}{2v_1^2} e^2 - \frac{1}{32\pi^2 v_2^2} p^2 \right\} \right]. \end{aligned} \tag{5.46}$$

The extremization equations

$$\frac{\partial \mathscr{E}}{\partial e} = 0, \quad \frac{\partial \mathscr{E}}{\partial v_1} = 0, \quad \frac{\partial \mathscr{E}}{\partial v_2} = 0 \tag{5.47}$$

can be easily solved to obtain

$$v_1 = v_2 = \frac{q^2 + p^2}{4\pi}, \quad e = \frac{q}{4\pi} \tag{5.48}$$

and

$$S_{wald}(q,p) = \mathscr{E}^*(q,p) = \frac{q^2 + p^2}{4}. \tag{5.49}$$

5.4 Elements of String Theory

5.4.1 BPS States in $\mathscr{N} = 4$ String Compactifications

Superstring theories are naturally formulated in ten-dimensional Lorentzian spacetime $\mathscr{M}_{10}$. A 'compactification' to four-dimensions is obtained by taking $\mathscr{M}_{10}$ to be a product manifold $\mathbb{R}^{1,3} \times X_6$ where X_6 is a compact Calabi-Yau threefold and $\mathbb{R}^{1,3}$ is the noncompact Minkowski spacetime. We will focus in these lectures on a compactification of Type-II superstring theory when X_6 is itself the product $X_6 = K3 \times T^2$. A highly nontrivial and surprising result from the 1990s is the statement that this compactification is quantum equivalent or 'dual' to a compactification of heterotic string theory on $T^4 \times T^2$ where T^4 is a four-dimensional torus [19, 20]. One can thus describe the theory either in the Type-II frame or the heterotic frame.

The four-dimensional theory in $\mathbb{R}^{1,3}$ resulting from this compactification has $\mathscr{N} = 4$ supersymmetry.[1] The massless fields in the theory consist of 22 vector multiplets in addition to the supergravity multiplet. The massless moduli fields consist of the S-modulus λ taking values in the coset

$$SL(2,\mathbb{Z})\backslash SL(2;\mathbb{R})/O(2;\mathbb{R}), \tag{5.50}$$

and the T-moduli μ taking values in the coset

$$O(22,6;\mathbb{Z})\backslash O(22,6;\mathbb{R})/O(22;\mathbb{R}) \times O(6;\mathbb{R}). \tag{5.51}$$

[1] This supersymmetry is a super Lie algebra containing $ISO(1,3) \times SU(4)$ as the bosonic subalgebra where $ISO(1,3)$ is the Poincaré symmetry of the $\mathbb{R}^{1,3}$ spacetime and $SU(4)$ is an internal symmetry called R-symmetry in physics literature. The odd generators of the superalgebra are called supercharges. With $\mathscr{N} = 4$ supersymmetry, there are eight complex supercharges which transform as a spinor of $ISO(1,3)$ and a fundamental of $SU(4)$.

The group of discrete identifications $SL(2,\mathbb{Z})$ is called S-duality group. In the heterotic frame, it is the electro-magnetic duality group [21, 22] whereas in the type-II frame, it is simply the group of area- preserving global diffeomorphisms of the T^2 factor. The group of discrete identifications $O(22,6;\mathbb{Z})$ is called the T-duality group. Part of the T-duality group $O(19,3;\mathbb{Z})$ can be recognized as the group of geometric identifications on the moduli space of K3; the other elements are stringy in origin and have to do with mirror symmetry.

At each point in the moduli space of the internal manifold $K3\times T^2$, one has a distinct four- dimensional theory. One would like to know the spectrum of particle states in this theory. Particle states are unitary irreducible representations, or supermultiplets, of the $\mathcal{N}=4$ superalgebra. The supermultiplets are of three types which have different dimensions in the rest frame. A long multiplet is 256-dimensional, an intermediate multiplet is 64-dimensional, and a short multiplet is 16-dimensional. A short multiplet preserves half of the eight supersymmetries (i.e. it is annihilated by four supercharges) and is called a half-BPS state; an intermediate multiplet preserves one quarter of the supersymmetry (i.e. it is annihilated by two supercharges), and is called a quarter-BPS state; and a long multiplet does not preserve any supersymmetry and is called a non-BPS state. One consequence of the BPS property is that the spectrum of these states is ‘topological’ in that it does not change as the moduli are varied, except for jumps at certain walls in the moduli space [23].

An important property of the BPS states that follows from the superalgebra is that their mass is determined by the charges and the moduli [23]. Thus, to specify a BPS state at a given point in the moduli space, it suffices to specify its charges. The charge vector in this theory transforms in the vector representation of the T-duality group $O(22,6;\mathbb{Z})$ and in the fundamental representation of the S-duality group $SL(2,\mathbb{Z})$. It is thus given by a vector $\Gamma^{i\alpha}$ with integer entries

$$\Gamma^{i\alpha}=\begin{pmatrix}Q^i\\ P^i\end{pmatrix}\quad where\quad i=1,2,\ldots 28;\quad \alpha=1,2 \tag{5.52}$$

transforming in the (2, 28) representation of $SL(2,\mathbb{Z})\times O(22,6;\mathbb{Z})$. The vectors Q and P can be regarded as the quantized electric and magnetic charge vectors of the state respectively. They both belong to an even, integral, self-dual lattice $\Pi^{22,6}$. We will assume in what follows that $\Gamma=(Q,P)$ in (5.52) is primitive in that it cannot be written as an integer multiple of (Q_0,P_0) for Q_0 and P_0 belonging to $\Pi^{22,6}$. A state is called purely electric if only Q is non-zero, purely magnetic if only P is non-zero, and dyonic if both P and Q are non-zero.

To define S-duality transformations, it is convenient to represent the S-modulus as a complex field S taking values in the upper half plane. An S-duality transformation

$$\gamma\equiv\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in SL(2;\mathbb{Z}) \tag{5.53}$$

acts simultaneously on the charges and the S-modulus by

$$\begin{pmatrix}Q\\ P\end{pmatrix}\to\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}Q\\ P\end{pmatrix};\qquad S\to\frac{aS+b}{cS+d} \tag{5.54}$$

To define T-duality transformations, it is convenient to represent the T-moduli by a 28×28 of matrix μ_I^A satisfying

$$\mu^t L \mu = L \tag{5.55}$$

with the identification that $\mu \sim k\mu$ for every $k \in O(22;\mathbb{R}) \times O(6;\mathbb{R})$. Here L is the (28×28) matrix

$$L_{IJ} = \begin{pmatrix} -\mathbf{C}_{16} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_6 \\ \mathbf{0} & \mathbf{I}_6 & \mathbf{0} \end{pmatrix}, \tag{5.56}$$

with $\mathbf{I}_s$ the $s \times s$ identity matrix and $\mathbf{C}_{16}$ is the Cartan matrix of $E_8 \times E_8$. The T-moduli are then represented by the matrix

$$\mathscr{M} = \mu^t \mu \tag{5.57}$$

which satisifies

$$\mathscr{M}^t = \mathscr{M}, \quad \mathscr{M}^t L \mathscr{M} = L \tag{5.58}$$

In this basis, a T-duality transformation can then be represented by a (28×28) matrix R with integer entries satisfying

$$R^t L R = L, \tag{5.59}$$

which acts simultaneously on the charges and the T-moduli by

$$Q \to RQ; \quad P \to RP; \quad \mu \to \mu R^{-1} \tag{5.60}$$

Given the matrix μ_I^A, one obtains an embedding $\Lambda^{22,6} \subset \mathbb{R}^{22,6}$ of $\Pi^{22,6}$ which allows us to define the moduli-dependent charge vectors Q and P by

$$Q^A = \mu_I^A Q_I \quad P^A = \mu_I^A P_I. \tag{5.61}$$

Note that while Q^I are integers Q^A are not. In what follows we will not always write the indices explicitly assuming that it will be clear from the context. In any case, the final answers will only depend on the T-duality invariants which are all integers. The matrix L has a 22-dimensional eigensubspace with eigenvalue -1 and a 6-dimensional eigensubspace with eigenvalue +1. Given Q and P, one can define the 'right-moving' charges[2] Q_R and P_R as the projections of Q and P respectively onto the subspace with eigenvalue +1. and the 'left-moving' charges as projections onto the subspace with eigenvalue -1. These definitions can be compactly written as

[2] The right-moving charges couple to the graviphoton vector fields associated with the right-moving chiral currents in the conformal field theory of the dual heterotic string.

$$Q_{R,L} = \frac{(1 \pm L)}{2} Q; \quad P_{R,L} = \frac{(1 \pm L)}{2} P \tag{5.62}$$

The right-moving charges since for the heterotic string, Q_R are related to the right-moving momenta. The central charges Z_1 and Z_2 of the $\mathcal{N} = 4$ superalgebra can then be defined in terms of the right-moving charges and moduli (For details of these definitions and the superalgebra, see Sect. 5.8.1.2).

If the vectors Q and P are nonparallel, then the state is quarter-BPS. On the other hand, if $Q = pQ_0$ and $P = qQ_0$ for some $Q_0 \in \Pi^{22,6}$ with p and q relatively prime integers, then the state is half-BPS.

An important piece of nonperturbative information about the dynamics of the theory is the exact spectrum of all possible dyonic BPS- states at all points in the moduli space. More specifically, one would like to compute the number $d(\Gamma)|_{\lambda,\mu}$ of dyons of a given charge Γ at a specific point (λ, μ) in the moduli space. Computation of these numbers is of course a very complicated dynamical problem. In fact, for a string compactification on a general Calabi-Yau threefold, the answer is not known. One main reason for focusing on this particular compactification on $K3 \times T^2$ is that in this case the dynamical problem has been essentially solved and the exact spectrum of dyons is now known. Furthermore, the results are easy to summarize and the numbers $d(\Gamma)|_{\lambda,\mu}$ are given in terms of Fourier coefficients of various modular forms.

In view of the duality symmetries, it is useful to classify the inequivalent duality orbits labeled by various duality invariants. This leads to an interesting problem in number theory of classification of inequivalent duality orbits of various duality groups such as $SL(2, \mathbb{Z}) \times O(22, 6; \mathbb{Z})$ in our case and more exotic groups like $E_{7,7}(\mathbb{Z})$ for other choices of compactification manifold X_6. It is important to remember though that a duality transformation acts simultaneously on charges and the moduli. Thus, it maps a state with charge Γ at a point in the moduli space (λ, μ) to a state with charge Γ' but at some other point in the moduli space (λ', μ'). In this respect, the half-BPS and quarter-BPS dyons behave differently.

- For half-BPS states, the spectrum does not depend on the moduli. Hence $d(\Gamma)|_{\lambda',\mu'} = d(\Gamma)|_{\lambda,\mu}$. Furthermore, by an S-duality transformation one can choose a frame where the charges are purely electric with $P = 0$ and $Q \neq 0$. Single-particle states have Q primitive and the number of states depends only on the T-duality invariant integer $n \equiv Q^2/2$. We can thus denote the degeneracy of half-BPS states $d(\Gamma)|_{S',\mu'}$ simply by $d(n)$.
- For quarter-BPS states, the spectrum does depend on the moduli, and $d(\Gamma)|_{\lambda',\mu'} \neq d(\Gamma)|_{\lambda,\mu}$. However, the partition function turns out to be independent of moduli and hence it is enough to classify the inequivalent duality orbits to label the partition functions. For the specific duality group $SL(2, \mathbb{Z}) \times O(22, 6; \mathbb{Z})$ the partition functions are essentially labeled by a single discrete invariant [24–26].

$$I = \gcd(Q \wedge P), \tag{5.63}$$

The degeneracies themselves are Fourier coefficients of the partition function. For a given value of I, they depend only on[3] the moduli and the three T-duality invariants $(m, n, \ell) \equiv (P^2/2, Q^2/2, Q \cdot P)$. Integrality of (m, n, ℓ) follows from the fact that both Q and P belong to $\Pi^{22,6}$. We can thus denote the degeneracy of these quarter-BPS states $d(\Gamma)|_{\lambda,\mu}$ simply by $d(m, n, l)|_{\lambda,\mu}$. For simplicity, we consider only $I = 1$ in these lectures. Generalization for higher I can be found in [27, 28].

5.4.2 Exercises

5.4.2.1 Elements of String Compactifications

The heterotic string theory in ten dimensions has 16 supersymmetries. The bosonic massless fields consist of the metric g_{MN}, a 2-form field $B^{(2)}$, 16 abelian 1-form gauge fields $A^{(r)}$ $r = 1, \ldots 16$, and a real scalar field ϕ called the dilaton. The Type-IIB string theory in ten dimensions has 32 supersymmetries. The bosonic massless fields consist of the metric g_{MN}; two 2-form fields $C^{(2)}$, $B^{(2)}$; a self-dual 4-form field $C^{(4)}$; and a complex scalar field λ called the dilaton-axion field.

One of the remarkable strong-weak coupling dualities is the 'string–string' duality between heterotic string compactified on $T^4 \times T^2$ and Type-IIB string compactified on $K3 \times T^2$. One piece of evidence for this duality is obtained by comparing the massless spectrum for these compactifications and certain half-BPS states in the spectrum.

Exercise 1 *how that the heterotic string compactified on $T^4 \times S^1 \times \tilde{S}^1$ leads a four dimensional theory with $\mathcal{N} = 4$ supersymmetry with 22 vector multiplets.*

Exercise 2 *Show that the Type-IIB string compactified on $K3 \times S^1 \times \tilde{S}^1$ leads a four dimensional theory with $\mathcal{N} = 4$ supersymmetry with 22 vector multiplets.*

Exercise 3 *Show that the Kaluza–Klein monopole in Type-IIB string associated with the circle $\tilde{S}^1$ has the right structure of massless fluctuations to be identified with the half-BPS perturbative heterotic string in the dual description.*

5.4.3 String–String Duality

It will be useful to recall a few details of the string–string duality between heterotic compactified on $T^4 \times S^1 \times \tilde{S}^1$ and Type-IIB compactified on $K3 \times S^1 \times \tilde{S}^1$. Two pieces of evidence for this duality will be relevant to our discussion.

[3] There is an additional dependence on arithmetic T-duality invariants but the degeneracies for states with nontrivial values of these T-duality invariants can be obtained from the degeneracies discussed here by demanding S-duality invariance [26].

Low energy effective action

Both these compactifications result in $\mathcal{N} = 4$ supergravity in four dimensions. With this supersymmetry, the two-derivative effective action for the massless fields receives no quantum corrections. Hence, if the two theories are to be dual to each other, they must have identical 2-derivative action.

This is indeed true. Even though the field content and the action are very different for the two theories in ten spacetime dimensions, upon respective compactifications, one obtains $\mathcal{N} = 4$ supergravity with 22 vector multiplets coupled to the supergravity multiplet. This has been discussed briefly in one of the tutorials. For a given number of vector multiplets, the two-derivative action is then completely fixed by supersymmetry and hence is the same for the two theories. This was one of the properties that led to the conjecture of a strong–weak coupling duality between the two theories.

For our purposes, we will be interested in the 2-derivative action for the bosonic fields. This is a generalization of the Einstein–Hilbert–Maxwell action (5.1) which couples the metric, the moduli fields and 28 abelian gauge fields:

$$\begin{aligned} I =& \frac{1}{32\pi} \int d^4x \sqrt{-\det G} S [R_G + \frac{1}{S^2} G^{\mu\nu} (\partial_\mu S \partial_\nu S - \frac{1}{2} \partial_\mu a \partial_\nu a) \\ &+ \frac{1}{8} G^{\mu\nu} Tr(\partial_\mu ML \partial_\nu ML) - G^{\mu\mu'} G^{\nu\nu'} F^{(i)}_{\mu\nu} (LML)_{ij} F^{(j)}_{\mu'\nu'} \\ &- \frac{a}{S} G^{\mu\mu'} G^{\nu\nu'} F^{(i)}_{\mu\nu} L_{ij} \tilde{F}^{(j)}_{\mu'\nu'}] \qquad i, j = 1, \ldots, 28. \end{aligned} \tag{5.64}$$

In the heterotic string picture, the expectation value of the dilaton field S is related to the four-dimensional string coupling g_4

$$S \sim \frac{1}{g_4^2}, \tag{5.65}$$

and a is the axion field. The metric $G_{\mu\nu}$ is the metric in the string frame and is related to the metric $g_{\mu\nu}$ in Einstein frame by the Weyl rescaling

$$g_{\mu\nu} = S G_{\mu\nu} \tag{5.66}$$

BPS spectrum

Another requirement of duality is that the spectrum of BPS states should match for the two dual theories. Perturbative states in one description will generically get mapped to some non-perturbative states in the dual description. As a result, this leads to highly nontrivial predictions about the nonpertubative spectrum in the dual description given the perturbative spectrum in one description.

As an example, consider the perturbative BPS-states in heterotic string theory on $K3 \times S^1 \times \tilde{S}^1$. A heterotic string wrapping w times on S^1 and carrying momentum n gets mapped in Type-IIA to the NS5-brane wrapping w times on $K3 \times S^1$ and carrying momentum n. One can go from Type-IIA to Type-IIB by a T-duality along the $\tilde{S}^1$

circle. Under this T-duality, the NS5-brane gets mapped to a KK-monopole with monopole charge w associated with the circle $\tilde{S}^1$ and carrying momentum n. This thus leads to a prediction that the spectrum of KK-monopole carrying momentum in Type-IIB should be the same as the spectrum of perturbative heterotic string discussed earlier. We will verify this highly nontrivial prediction in the next subsection for the case of $w=1$.

5.4.4 Kaluza–Klein Monopole and the Heterotic String

The metric of the Kaluza–Klein monopole is given by the so-called Taub-NUT metric

$$ds^2_{TN} = \left(1+\frac{R_0}{r}\right)\left(dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right) + R_0^2\left(1+\frac{R_0}{r}\right)^{-1}(2d\psi + \cos\theta d\phi)^2 \tag{5.67}$$

with the identifications:

$$(\theta,\phi,\psi) \equiv (2\pi-\theta, \phi+\pi, \psi+\frac{\pi}{2}) \equiv (\theta, \phi+2\pi, \psi+\pi) \equiv (\theta,\phi,\psi+2\pi). \tag{5.68}$$

Here R_0 is a constant determining the size of the Taub-NUT space $\mathscr{M}_{TN}$. This metric satisfies the Einstein equations in four-dimensional Euclidean space. The metric (5.67) admits a normalizable self-dual harmonic form ω, given by

$$\omega^{KK} = \frac{r}{r+R_0}d\sigma_3 + \frac{R_0}{(r+R_0)^2}dr\wedge\sigma_3, \qquad \sigma_3 \equiv \left(d\psi + \frac{1}{2}\cos\theta d\phi\right). \tag{5.69}$$

We are interested in the Type-IIB string theory compactified on $K_3\times\tilde{S}^1\times S^1$ in the presence of a Kaluza–Klein monopole, with $\tilde{S}^1$ identified with the asymptotic circle of the Taub-NUT space labeled by the coordinate ψ in (5.67). Thus, we want analyze the massless fluctuations of Type-IIB string on $K_3\times S^1\times\mathscr{M}_{TN}$ space. Let y and $\tilde{y}$ be the coordinates of S^1 and $\tilde{S}^1$ respectively with $y\sim y+2\pi R$ and $\tilde{y}\sim\tilde{y}+2\pi\tilde{R}$. When the radius R of the S^1 is large compared to the size of the $K3$ and the radius $\tilde{R}$ of the $\tilde{S}^1$ circle, we obtain an 'effective string' wrapping the S^1 with massless spectrum that agrees with the massless spectrum of a fundmental heterotic string wrapping S^1. These massless modes can be deduced as follows:

- The center-of-mass of the KK-monopole can be located anywhere in $\mathbb{R}^3$ and its position is specified by a vector $\mathbf{a}$. Thus, we have

$$r := |\mathbf{x}-\mathbf{a}|, \quad \cos\theta := \frac{x^3-a^3}{r}, \quad \tan\phi := \frac{x^1-a^1}{x^2-a^2}. \tag{5.70}$$

if (x^1, x^2, x^3) are the coordinates of $\mathbb{R}^3$. We can allow these coordinates to fluctuate in the t and y directions and hence we will obtain three non-chiral massless $a^i(t, y)$ scalar fields along the effective string associated with oscillations of the three coordinates of the center-of-mass of the KK monopole.

- There are two additional non-chiral scalar fields $b(t, y)$ and $c(t, y)$ obtained by reducing the two 2-form fields $B^{(2)}$ and C^2 of Type-IIB along the harmonic 2-form (5.69):

$$B^{(2)} = b(t, y) \cdot \omega^{KK} \quad C^{(2)} = c(t, y) \cdot \omega^{KK} \tag{5.71}$$

- There are 3 right-moving $a^r_R(t + y), r = 1, 2, 3$ and 19 left-moving scalars $a^s_L(t - y), s = 1, \ldots, 19$ obtained by reducing the self-dual 4-form field $C^{(4)}$ of type IIB theory. This works as follows. The field $C^{(4)}$ can be reduced taking it as a tensor product of the harmonic 2-form (5.69) and a harmonic 2-form $\omega^{K_3}_\alpha$ for $\alpha = 1, \ldots, 22$ on K_3. This gives rise to rise to a chiral scalar field on the world-volume. The chirality of the scalar field is correlated with whether the corresponding harmonic 2-form $\omega^{K_3}_\alpha$ is self-dual or anti-self-dual. Since $K3$ has three self-dual $\omega^{K_3+}_r$ and nineteen anti-selfdual harmonic 2-forms $\omega^{K_3-}_s$, we get 3 right-moving and 19 left-moving scalars:

$$C^{(4)} = \sum_{r=1}^{3} a^s_R(t + y) \cdot \omega^{K_3-}_s \wedge \omega^{KK} + \sum_{s=1}^{19} a^s_L(t - y) \cdot \omega^{K_3-}_s \wedge \omega^{KK}. \tag{5.72}$$

The KK-monopole background breaks 8 of the 16 supersymmetries of Type-II on $K3 \times S^1$. Consequently, there are eight right-moving fermionic fields

$$S^a(t + y) \quad a = 1, \ldots, 8$$

which arise as the goldstinos of these eight broken supersymmetries. This is precisely the field content of the 1+1 dimensional worldsheet theory of the heterotic string wrapping S^1 as we discussed in the tutorial (Sect. 5.5.1).

5.4.5 Supersymmetry and Extremality

Some of the special properties of external black holes can be understood better by embedding them in supergravity. We will be interested in these lectures in string compactifications with $\mathcal{N} = 4$ supersymmetry in four spacetime dimensions. The $\mathcal{N} = 4$ supersymmetry algebra contains in addition to the usual Poincaré generators, sixteen real supercharges which can be grouped into 8 complex charges Q^a_α and their complex conjugates. Here $\alpha = 1, 2$ is the usual Weyl spinor index of 4d Lorentz symmetry. and the internal index $a = 1, \ldots, 4$ in the fundamental **4** representation of an $SU(4)$, the R-symmetry of the superalgebra. The relevant anticommutators for our purpose are

$$\{Q^a_\alpha, \bar{Q}_{\dot{\beta}b}\} = -2P_\mu \sigma^\mu_{\alpha\dot{\beta}} \delta^a_b$$
$$\{Q^a_\alpha, Q^b_\beta\} = \varepsilon_{\alpha\beta} Z^{ab} \quad \{\bar{Q}_{\dot{\alpha}a}, \bar{Q}_{\dot{\beta}b}\} = \bar{Z}_{ab}\varepsilon_{\dot{\alpha}\dot{\beta}} \tag{5.73}$$

where σ^μ are (2×2) matrices with $\sigma_0 = -\mathbf{1}$ and σ^i for $i = 1, 2, 3$ are the usual Pauli matrices. Here P_μ is the momentum operator and Q are the supersymmetry generators and the complex number Z^{ab} is the central charge matrix.

Let us first look at the representations of this algebra when the central charge is zero. In this case the massive and massless representation are qualitatively different.

1. Massive Representation, $M > 0, P^\mu = (M, 0, 0, 0)$
   In this case (5.73) becomes $\{Q^a_\alpha, \bar{Q}_{\dot{\beta}b}\} = 2M\delta_{\alpha\dot{\beta}}\delta^a_b$ and all other anti-commutators vanish. Up to overall scaling, these are the commutation relations for eight complex fermionic oscillators. Each oscillator has a two-state representation, which is either filled or empty. These states together define a unitary irreducible representation, called a supermultiplet, of the superalgebra. The total dimension of the representation is $2^8 = 256$ which is CPT self-conjugate.
2. Massless Representation $M = 0, P^\mu = (E, 0, 0, E)$
   In this case (5.73) becomes $\{Q^a_1, \bar{Q}_{\dot{1}b}\} = 2E\delta^a_b$ and all other anti-commutators vanish. Up to overall scaling, these are now the anti-commutation relations of *four* fermionic oscillators and hence the total dimension of the representation is $2^4 = 16$ which is also CPT-self-conjugate.

The important point is that for a massive representation, with $M = \varepsilon > 0$, no matter how small ε, the supermultiplet is long and precisely at $M = 0$ it is short. Thus the size of the supermultiplet has to change discontinuously if the state has to acquire mass. Furthermore, the size of the supermultiplet is determined by the number of supersymmetries that are *broken* because those have non-vanishing anti-commutations and turn into fermionic oscillators.

Note that there is a bound on the mass $M \geq 0$ which simply follows from the fact the using (5.73) one can show that the mass operator on the right hand side of the equation equals a positive operator, the absolute value square of the supercharge on the left hand side. The massless representation saturates this bound and is 'small' whereas the massive representation is long.

There is an analog of this phenomenon also for nonzero Z_{ab}. As explained in the appendix, the central charge matrix Z_{ab} can be brought to the standard form by an $U(4)$ rotation

$$\tilde{Z} = UZU^T, \quad U \in U(4), \quad \tilde{Z}_{ab} = \left(\begin{array}{c|c} Z_1\varepsilon & 0 \\ \hline 0 & Z_2\varepsilon \end{array}\right), \quad \varepsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.74}$$

so we have two 'central charges' Z_1 and Z_2. Without loss of generality we can assume $|Z_1| \geq |Z_2|$. Using the supersymmetry algebra one can prove the BPS bound $M - |Z_1| \geq 0$ by showing that this operator is equal to a positive operator (see appendix for details). States that saturate this bound are the BPS states. There are three types of representations:

- If $M = |Z_1| = |Z_2|$, then eight of of the sixteen supersymmetries are preserved. Such states are called half-BPS. The broken supersymmetries result in four complex fermionic zero modes whose quantization furnishes a 2^4-dimensional short multiplet.
- If $M = |Z_1| > |Z_2|$, then and four out of the sixteen supersymmetries are preserved. Such states are called quarter-BPS. The broken supersymmetries result in six complex fermionic zero modes whose quantization furnishes a 2^6-dimensional intermediate multiplet.
- If $M > |Z_1| > |Z_2|$, then no supersymmetries are preserved. Such states are called non-BPS.The sixteen broken supersymmetries result in eight complex fermionic zero modes whose quantization furnishes a 2^8-dimensional long multiplet.

The significance of BPS states in string theory and in gauge theory stems from the classic argument of Witten and Olive which shows that under suitable conditions, the spectrum of BPS states is stable under smooth changes of moduli and coupling constants. The crux of the argument is that with sufficient supersymmetry, for example $\mathcal{N} = 4$, the coupling constant does not get renormalized. The central charges Z_1 and Z_2 of the supersymmetry algebra depend on the quantized charges and the coupling constant which therefore also does not get renormalized. This shows that for BPS states, the mass also cannot get renormalized because if the quantum corrections increase the mass, the states will have to belong a long representation . Then, the number of states will have to jump discontinuously from, say from 16 to 256 which cannot happen under smooth variations of couplings unless there is some kind of a 'Higgs Mechanism' or there is some kind of a phase transition. [4]

As a result, one can compute the spectrum at weak coupling in the region of moduli space where perturbative or semiclassical counting methods are available. One can then analytically continue this spectrum to strong coupling. This allows us to obtain invaluable non-perturbative information about the theory from essentially perturbative commutations.

5.4.6 BPS Dyons in $\mathcal{N} = 4$ Compactifications

The massless spectrum of the toroidally compactified heterotic string on T^6 contains 28 different "photons" or $U(1)$ gauge fields—one from each of the 22 vector multiplets and 6 from the supergravity multiplet.As a result, the electric charge of a state is

---

[4] Such 'phase transitions' do occur and the degeneracies can jump upon crossing certain walls in the moduli space. This phenomenon called 'wall-crossing' occurs not because of Higgs mechanism but because at the walls, single particle states have the same mass as certain multi-particle states and can thus mix with the multi-particle continuum states. The wall-crossing phenomenon complicates the analytic continuation of the degeneracy from weak coupling from strong coupling since one may encounter various walls along the way. However, in many cases, the jumps across these walls can be taken into account systematically.

specified by a 28-dimensional charge vector Q and the magnetic charge is specified by a 28-dimensional charge vector P. Thus, a dyonic state is specified by the charge vector

$$\Gamma = \begin{pmatrix} Q \\ P \end{pmatrix} \tag{5.75}$$

where Q and P are the electric and magnetic charge vectors respectively. Both Q and P are elements of a self-dual integral lattice $\Pi^{22,6}$ and can be represented as 28-dimensional column vectors in $\mathbb{R}^{22,6}$ with integer entries, which transform in the fundamental representation of $O(22, 6; \mathbb{Z})$. We will be interested in BPS states.

- For half-BPS state the charge vectors Q and P must be parallel. These states are dual to perturbative BPS states.
- For a quarter-BPS states the charge vectors Q and P are not parallel. There is no duality frame in which these states are perturbative.

There are three invariants of $O(22, 6; \mathbb{Z})$, quadratic in charges, and given by P^2, Q^2 and $Q \cdot P$. These three T-duality invariants will be useful in later discussions.

5.5 Spectrum of Half-BPS Dyons

An instructive example of BPS of states is provided by an infinite tower of BPS states that exists in perturbative string theory [29, 30].

5.5.1 Perturbative Half-BPS States

Consider a perturbative heterotic string state wrapping around S^1 with winding number w and quantized momentum n. Let the radius of the circle be R and $\alpha' = 1$, then one can define left-moving and right-moving momenta as usual,

$$p_{L,R} = \sqrt{\frac{1}{2}} \left(\frac{n}{R} \pm wR \right). \tag{5.76}$$

Recall that the heterotic strings consists of a right-moving superstring and a left-moving bosonic string. In the NSR formalism in the light-cone gauge, the worldsheet fields are:

- Right moving superstring $X^i(\sigma^-)$ $\tilde{\psi}^i(\sigma^-)$ $\quad i = 1 \cdots 8$
- Left-moving bosonic string $X^i(\sigma^+), X^I(\sigma^+)$ $\quad I = 1 \cdots 16,$

where X^i are the bosonic transverse spatial coordinates, $\tilde{\psi}^i$ are the worldsheet fermions, and X^I are the coordinates of an internal $E_8 \times E_8$ torus. A BPS state is obtained by keeping the right-movers in the ground state (that is, setting the right-moving oscillator number $\tilde{N} = \frac{1}{2}$ in the NS sector and $\tilde{N} = 0$ in the R sector).

The Virasoro constraints are then given by

$$\tilde{L_0} - \frac{M^2}{4} + \frac{p_R^2}{2} = 0 \tag{5.77}$$

$$L_0 - \frac{M^2}{4} + \frac{p_L^2}{2} = 0, \tag{5.78}$$

where N and $\tilde{N}$ are the left-moving and right-moving oscillation numbers respectively.

The left-moving oscillator number is then

$$L_0 = \sum_{n=1}^{\infty}\left(\sum_{i=1}^{8} n a^i_{-n} a^i_n + \sum_{I=1}^{16} n\beta^I_{-n}\beta^I_{-n}\right) - 1 := N - 1, \tag{5.79}$$

where a^i are the left-moving Fourier modes of the fields X^i, and β^I are the Fourier modes of the fields X^I. Note that the right-moving fermions satisfy anti-periodic boundary condition in the NS sector and have half-integral moding, and satisfy periodic boundary conditions in the R sector and have integral moding. The oscillator number operator is then given by

$$\tilde{L}_0 = \sum_{n=1}^{\infty}\sum_{i=1}^{8}(n\tilde{a}^i_{-n}\tilde{a}^i_n + r\tilde{\psi}^i_{-r}\tilde{\psi}^i_r - \frac{1}{2}) := \tilde{N} - \frac{1}{2}, \tag{5.80}$$

with $r \equiv -(n - \frac{1}{2})$ in the NS sector and by

$$\tilde{L}_0 = \sum_{n=1}^{\infty}\sum_{i=1}^{8}(n\tilde{a}^i_{-n}\tilde{a}^i_n + r\tilde{\psi}^i_{-r}\tilde{\psi}^i_r) \tag{5.81}$$

with $r \equiv (n - 1)$ in the R sector.

In the NS-sector then one then has $\tilde{N} = \frac{1}{2}$ and the states are given by

$$\tilde{\psi}^i_{-\frac{1}{2}}|0\rangle, \tag{5.82}$$

that transform as the vector representation 8_v of $SO(8)$. In the R sector the ground state is furnished by the representation of fermionic zero mode algebra $\{\psi^i_0, \psi^j_0\} = \delta^{ij}$ which after GSO projection transforms as 8_s of $SO(8)$. Altogether the right-moving ground state is thus 16-dimensional $8_v \oplus 8_s$. From the Virasoro constraint (5.77) we see that a BPS state with $\tilde{N} = 0$ saturates the BPS bound

$$M = \sqrt{2}p_R, \tag{5.83}$$

and thus $\sqrt{2}p_R$ can be identified with the central charge of the supersymmetry algebra. The right-moving ground state after the usual GSO projection is indeed 16-dimensional as expected for a BPS-state in a theory with $\mathcal{N} = 4$ supersymmetry.

We thus have a perturbative BPS state which looks pointlike in four dimensions with two integral charges n and w that couple to two gauge fields $g_{5\mu}$ and $B_{5\mu}$ respectively. It saturates a BPS bound $M = \sqrt{2}p_R$ and belongs to a 16-dimensional short representation. This point-like state is our 'would-be' black hole. Because it has a large mass, as we increase the string coupling it would begin to gravitate and eventually collapse to form a black hole.

Microscopically, there is a huge multiplicity of such states which arises from the fact that even though the right-movers are in the ground state, the string can carry arbitrary left-moving oscillations subject to the Virasoro constraint. Using $M = \sqrt{2}p_R$ in the Virasoro constraint for the left-movers gives us

$$N - 1 = \frac{1}{2}(p_R^2 - p_L^2) := Q^2/2 = nw. \tag{5.84}$$

We would like to know the degeneracy of states for a given value of charges n and w which is given by exciting arbitrary left-moving oscillations whose total worldsheet oscillator excitation number adds up to N. Let us take $w = 1$ for simplicity and denote the degeneracy by $d(n)$ which we want to compute. As usual, it is more convenient to evaluate the canonical partition function

$$Z(\beta) = \mathrm{Tr}\left(e^{-\beta L_0}\right) \tag{5.85}$$

$$\equiv \sum_{-1}^{\infty} d(n)q^n \quad q := e^{-\beta}. \tag{5.86}$$

This is the canonical partition function of 24 left-moving massless bosons in 1+1 dimensions at temperature $1/\beta$. The micro-canonical degeneracy $d(N)$ is given then given as usual by the inverse Laplace transform

$$d(N) = \frac{1}{2\pi i}\int d\beta e^{\beta N} Z(\beta). \tag{5.87}$$

Using the expression (5.79) for the oscillator number s and the fact that

$$\mathrm{Tr}(q^{-s\alpha_{-n}\alpha_n}) = 1 + q^s + q^{2s} + q^{3s} + \cdots = \frac{1}{(1-q^s)}, \tag{5.88}$$

the partition function can be readily evaluated to obtain

$$Z(\beta) = \frac{1}{q}\prod_{s=1}^{\infty}\frac{1}{(1-q^s)^{24}}. \tag{5.89}$$

It is convenient to introduce a variable τ by $\beta := -2\pi i\tau$, so that $q := e^{2\pi i\tau}$. The function

$$\Delta(\tau) = q\prod_{s=1}^{\infty}(1-q^s)^{24}, \tag{5.90}$$

is the famous discriminant function. Under modular transformations

$$\tau \to \frac{a\tau+b}{c\tau+d} \quad a,b,c,d \in \mathbb{Z}, \quad with \quad ad-bc=1 \tag{5.91}$$

it transforms as a modular form of weight 12:

$$\Delta\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^{12}\Delta(\tau)\ . \tag{5.92}$$

This remarkable property allows us to relate high temperature ($\beta \to 0$) to low temperature ($\beta \to \infty$) and derive a simple explicit expression for the asymptotic degeneracies $d(n)$ for n very large.

5.5.2 Cardy Formula

The degeneracy $d(N)$ can be obtained from the canonical partition function by the inverse Laplace transform

$$d(N) = \frac{1}{2\pi i}\int d\beta e^{\beta N} Z(\beta). \tag{5.93}$$

We would like to evaluate this integral (5.93) for large N which corresponds to large worldsheet energy. Such an asymptotic expansion of $d(N)$ for large N is given by the 'Cardy formula' which utilizes the modular properties of the partition function.

For large N, we expect that the integral receives most of its contributions from high temperature or small β region of the integrand. To compute the large N asymptotics, we then need to know the small β asymptotics of the partition function. Now, $\beta \to 0$ corresponds to $q \to 1$ and in this limit the asymptotics of $Z(\beta)$ are very difficult to read off from (5.89) because its a product of many quantities that are becoming very large. It is more convenient to use the fact that $Z(\beta)$ is the inverse of $\Delta(\tau)$ which is a modular form of weight 12 we can conclude

$$Z(\beta) = (\beta/2\pi)^{12} Z(\frac{4\pi^2}{\beta}). \tag{5.94}$$

This allows us to relate the $q \to 1$ or high temperature asymptotics to $q \to 0$ or low temperature asymptotics as follows. Now, $Z(\tilde{\beta}) = Z\left(\frac{4\pi^2}{\beta}\right)$ asymptotics are easy to read off because as $\beta \to 0$ we have $\tilde{\beta} \to \infty$ or $e^{-\tilde{\beta}} = \tilde{q} \to 0$. As $\tilde{q} \to 0$

$$Z(\tilde{\beta}) = \frac{1}{\tilde{q}} \prod_{n=1}^{\infty} \frac{1}{(1-\tilde{q}^n)^{24}} \sim \frac{1}{\tilde{q}}. \tag{5.95}$$

This allows us to write

$$d(N) \sim \frac{1}{2\pi i} \int \left(\frac{\beta}{2\pi}\right)^{12} e^{\beta N + \frac{4\pi^2}{\beta}} d\beta. \tag{5.96}$$

This integral can be evaluated easily using saddle point approximation. The function in the exponent is $f(\beta) \equiv \beta N + \frac{4\pi^2}{\beta}$ which has a maximum at

$$f'(\beta) = 0 \quad \text{or} \quad N - \frac{4\pi^2}{\beta_c} = 0 \quad \text{or} \quad \beta_c = \frac{2\pi}{\sqrt{N}}. \tag{5.97}$$

The value of the integrand at the saddle point gives us the leading asymptotic expression for the number of states

$$d(N) \sim \exp(4\pi\sqrt{N}). \tag{5.98}$$

This implies that the ensemble of such BPS states of a given charge vector Q has nonzero statistical entropy that goes to leading order as

$$S_{stat}(Q) := \log(d(Q)) = 4\pi\sqrt{Q^2/2}. \tag{5.99}$$

We would now like to identify the black hole solution corresponding to this state and test if this microscopic entropy agrees with the macroscopic entropy of the black hole.

The formula that we derived for the degeneracy $d(N)$ is valid more generally in any 1+1 CFT. In a general CFT, the partition function is a modular form of weight $-k$

$$Z(\beta) \sim Z\left(\frac{4\pi^2}{\beta}\right)\beta^k.$$

which allows us to determine high temperature asymptotics from low temperature asymptotics for $Z(\tilde{\beta})$ once again because

$$\tilde{\beta} \equiv \frac{4\pi^2}{\beta} \to \infty \qquad \text{as} \quad \beta \to 0. \tag{5.100}$$

At low temperature only ground state contributes

$$\begin{aligned} Z(\tilde{\beta}) &= \mathrm{Tr}\exp(-\tilde{\beta}(L_0 - c/24)) \\ &\sim \exp(-E_0\tilde{\beta}) \sim \exp\left(\frac{\tilde{\beta}c}{24}\right), \end{aligned}$$

where c is the central charge of the theory. Using the saddle point evaluation as above we then find.

$$d(N) \sim \exp\left(2\pi\sqrt{\frac{cN}{6}}\right). \tag{5.101}$$

In our case, because we had 24 left-moving bosons, $c = 24$, and then (5.101) reduces to (5.98).

5.6 Spectrum of Quarter-BPS Dyons

In this section we consider the spectrum of quarter-BPS dyons in the simplest string compactification with $\mathcal{N} = 4$ in four spacetime dimensions. Surprisingly, the partition function for counting these dyons turns out to involve interesting mathematical objects called Siegel modular forms which are a natural generalizations for the group $Sp(2, \mathbb{Z})$ of usual modular forms of the group $Sp(1, \mathbb{Z}) \sim SL(2, \mathbb{Z})$. See Sect. 5.8.2.1 for a review of Siegel modular forms and related Jacobi modular forms.

5.6.1 Siegel Modular Forms and Dyons

Siegel forms occur naturally in the context of counting of quarter-BPS dyons. The partition function for these dyons depends on three (complexified) chemical potentials (σ, τ, z), conjugate to the three T-duality invariant integers

$$(P^2/2, Q^2/2, P \cdot Q) := (m, n, \ell)$$

respectively and is given by

$$\mathrm{Z}(\Omega) = \frac{1}{\Phi_{10}(\Omega)}. \tag{5.102}$$

Note that this is very analogous to the case of half-BPS states discussed in the tutorials where the partition function was

$$\mathrm{Z}(\tau) = \frac{1}{\Delta(\tau)}. \tag{5.103}$$

was the inverse of a modular form $\Delta(\tau)$ of weight 12 of the group $Sp(1, \mathbb{Z})$.

The product representation of the Igusa form is particularly useful for the physics application because it is closely related to the generating function for the elliptic genera of symmetric products of $K3$ introduced in the Appendix. This is a consequence of the fact that the multiplicative lift of the Igusa form is obtained starting with the elliptic genus of a single copy $K3$ as the input. The generating function for the elliptic genera of symmetric products of K3 is defined by

$$\widehat{Z}(\sigma, \tau, z) := \sum_{m=-1}^{\infty} \chi_{m+1}(\tau, z) p^m \tag{5.104}$$

where $\chi_m(\tau, z)$ is the elliptic genus of $\mathrm{Sym}^m(K3)$ with $\chi_0(\tau, z) \equiv 1$ and $\chi_1(\tau, z) \equiv \chi(\tau, z)$. A standard orbifold computation [31] gives

$$\widehat{Z}(\sigma, \tau, z) = \frac{1}{p} \prod_{s>0, t\geq 0, r} \frac{1}{(1 - p^s q^t y^r)^{C_0(4st - r^2)}} \tag{5.105}$$

in terms of the Fourier coefficients C_0 of the elliptic genus of a single copy of $K3$. As we will explain in the next section, this partition function captures the degeneracies of bound state of m D1-branes and a single D5-brane carrying momentum and spin.

Comparing the product representation for the Igusa form (5.234) with (5.105), we get the relation:

$$\mathrm{Z}(\Omega) = \frac{1}{\Phi_{10}(\sigma, \tau, z)} = \frac{\widehat{Z}(\sigma, \tau, z)}{\psi(\tau, z)}. \tag{5.106}$$

This relation of the Igusa form to the elliptic genera of symmetric products of $K3$ and the degeneracies of D1–D5 bound states has a deeper physical significance and allows for a microscopic derivation of the counting formula as we explain below.

The the logic of the derivation is as follows:

1. We derive the degeneracy for a special charge configuration in one corner of the moduli space.
2. Using constraints from wall-crossing, we extend this answer for the same set of charges to all over the moduli space.
3. Using duality symmetries, we extend this answer to all possible values of charges.

With this general strategy in mind, we turn to the derivation of the dyon partition function for a special representative set of charges in a certain weakly coupled region of the moduli space.

5.6.2 A Representative Charge Configuration

Consider four-dimensional BPS-states in Type IIB on $K3 \times S^1 \times \tilde{S}^1$ with the following charge configuration:

- 1 KK-monopole associated with the circle $\tilde{S}^1$.
- 1 D5-branes wrapping $K3 \times S^1$
- m D1-branes wrapping S^1

- n units of momentum along the circle S^1
- l units of momentum along the circle $\tilde{S}^1$

We would like to compute $d(m, n, l)$ which is the number of quantum states with these quantum numbers counting bosons with +1 and fermions with -1 . Let F be the spacetime fermion number then we could try to compute

$$\mathrm{Tr}_{m,n,l}\Big[(-1)^F\Big]. \tag{5.107}$$

However, this vanishes. If a state breaks $2n$ supersymmetries, then it has $2n$ real fermion zero modes which are the Goldstinoes of the broken symmetry. Quantization of each pair leads to Bose–Fermi degeneracy so the trace above vanishes. This can be remedied by inserting $(2h)^n$ where h is the 'helicity', that is, the third component of angular momentum in the rest frame. For states paired by a complex fermion the effect of this insertion is to 'soak up' the fermion zero mode since this mode has spin half. Thus, we compute

$$d(m, n, l) = \mathrm{Tr}_{m,n,l}\Big[(-1)^F (2h)^6\Big] \tag{5.108}$$

since for a quarter-BPS state, out of the 16 supersymmetries 12 are broken. In practice, this means we just ignore the 12 fermionic zero modes from broken supersymmetry and evaluate simply $\mathrm{Tr}(-1)^F$ over the remaining modes. The index thus defined receives contribution only from the BPS states.

It turns out that we can relate these unknown degeneracies $d(m, n, l)$ of 4d-states to known degeneracies of the D1–D5-P configuration in five dimensions which are much easier to compute. This is known as the 4–5d lift [32]. The main idea is to use the fact that the geometry of the Kaluza–Klein monopole (5.67) in the charge configuration above asymptotes to $\mathbb{R}^3 \times \tilde{S}^1$ at asymptotic infinity $r \to \infty$ but reduces to flat Euclidean space $\mathbb{R}^4$ near the core of the monopole at $r \to 0$. Thus at asymptotic infinity we have a KK-monopole in four-dimensional flat Minkowski spacetime which near the core looks like a five-dimensional flat Minkowski spacetime. Our charge configuration then reduces essentially to the *five-dimensional* Strominger–Vafa black hole [33] with angular momentum [34] discussed in the previous subsection.

Our strategy will be to compute the grand canonical partition function introducing chemical potentials (σ, τ, z) conjugate to the charges (m, n, l) and the 'fugacities'

$$p := e^{2\pi i\sigma}, \quad q := e^{2\pi i\tau}, \quad y := e^{2\pi i z}. \tag{5.109}$$

The partition function is then

$$Z(\sigma, \tau, z) = \sum_{m,n,l} p^m q^n y^l (-1)^l d(m, n, l). \tag{5.110}$$

The factor of $(-1)^l$ is introduced for convenience which can be absorbed by $z \to z + 1/2$.

Since $d(m, n, l)$ is a topological quantity protected from quantum corrections, the dyon partition function it does not depend on the coupling or the moduli such as the radius $\tilde{R}$. We can focus on the region near the core by taking the radius of the circle $\tilde{S}^1$ goes to infinity so that in this limit we have a weakly coupled problem. In this limit, the charge l corresponding to the momentum around this circle gets identified with the angular momentum l in five dimensions. The total partition function at weak coupling at large radius $\tilde{R}$ is thus a product of three factors

$$Z(\Omega) = Z_{D1}(p, q, y) Z_{KK}(q) Z_{CM}(q, y). \tag{5.111}$$

The three factors arise as follows:

1. The factor $Z_{D1}(\sigma, \tau, z)$ counts the bound states of the D1-brane bound to a single D5-brane, carrying arbitrary momentum and angular momentum.
2. The factor $Z_{KK}(\tau)$ counts the bound states of momentum n with the Kaluza–Klein monopole. The KK-monopole cannot carry any momentum along the $\tilde{S}^1$ directions nor does it carry any D1-brane charge. Hence the partition function depends only τ.
3. The factor $Z_{CM}(\tau, z)$ counts the bound states of the center of mass motion of the Strominger–Vafa black hole in the Kaluza–Klein geometry [35, 36]. It carries no D1-brane charge and hence depends only τ and z.

At weak coupling, these three systems reduce to decoupled bosonic and fermionic oscillators and our computation is reduced to something very similar to perturbative calculation described in the previous section. Each oscillator carries certain quantum numbers (s, t, r) which can contribute to the total charge (m, n, l) of our interest. Each bosonic oscillator contributes

$$\sum_{k=0}^{\infty} e^{2\pi i k(s\sigma, t\tau, rz)} = \left(1 - p^s q^t y^r\right)^{-1}. \tag{5.112}$$

Each fermionic oscillator contributes

$$\sum_{k=0}^{1} e^{2\pi i k(s\sigma, t\tau, rz)} (-1)^k = \left(1 - p^s q^t y^r\right) \tag{5.113}$$

where the $(-1)^k$ is present because of $(-1)^F$. The partition function will be thus of the general form

$$Z(\Omega) \sim \prod_{s,t,r} \frac{1}{(1 - p^s q^t y^r)^{f(s,t,r)}}, \tag{5.114}$$

where $f(s, t, r)$ is the difference between the number of bosonic oscillators and the number of fermionic oscillators for given charges (s, t, r). All physics is now contained in these numbers. In the remaining subsections we discuss systematically various contribution to the partition function to determine $f(s, t, r)$ for our system.

5.6.3 Bound States of D1-Branes and D5-Branes

As a warm up, let us first consider D1-brane (or fundamental Type-II string) in flat space wrapped around a circle S^1 or radius R with coordinate $y \sim y + 2\pi R$. The fluctuations of the D1-brane consists of 8 transverse bosons $\phi^i(t, y)$ as well as 8 left-chiral fermions $S^a(t + y)$ and 8 right-chiral fermions $\tilde{S}^a(t - y)$ where t is the time coordinate, $i = 1, \ldots, 8$, and $a = 1, \ldots, 8$. These constitute the field content of the 1+1 D CFT living on S^1. The fluctuations are of the form

$$\phi^i(t, y) = \phi^i_0 + p^i_0 t + \sum_{n>0} \phi^i_n e^{-\frac{n}{R}(t-y)} + \sum_{n>0} \tilde{\phi}^i_n e^{-\frac{n}{R}(t+y)} + c.c. \tag{5.115}$$

For the fermions we have similarly

$$S^a(t - y) = \sum_{n>0} S^a_n e^{-\frac{n}{R}(t-y)} + c.c. \tag{5.116}$$

$$\tilde{S}^a(t + y) = \sum_{n>0} \tilde{S}^a_n e^{-\frac{n}{R}(t+y)} + c.c. \tag{5.117}$$

We can quantize this system as usual. Then ϕ^i_n and $\tilde{\phi}^i_n$ are bosonic oscillators with frequencies n/R and occupation numbers N^i_n and $\tilde{N}^i_n$ respectively. Similarly, S^a_n and $\tilde{S}^a_n$ are fermionic oscillators with frequencies n/R and occupation numbers M^i_n and $\tilde{M}^i_n$ respectively. The total left-moving momentum along S^1 is

$$P = \frac{1}{R} \sum_{i=1}^{8} \sum_{n=1}^{\infty} n(N^i_n - \tilde{N}^i_n) + \frac{1}{R} \sum_{a=1}^{8} \sum_{n=1}^{\infty} n(M^a_n - \tilde{M}^a_n) \tag{5.118}$$

and the total energy is

$$E = \frac{1}{R} \sum_{i=1}^{8} \sum_{n=1}^{\infty} n(N^i_n + \tilde{N}^i_n) + \frac{1}{R} \sum_{a=1}^{8} \sum_{n=1}^{\infty} n(M^a_n + \tilde{M}^a_n) \tag{5.119}$$

To obtain a BPS state we want to minimize the energy given fixed momentum P. This implies

$$\tilde{N}^i_n = 0, \quad \tilde{M}^i_n = 0 \quad E = P. \tag{5.120}$$

We would like to know how many BPS states there are for a given charge P. This is a combinatorial problem of finding $d(P)$ which is the number of ways to choose a set of integers $\{N^i_n, M^a_n\}$ satisfying the constraint

$$\frac{1}{R} \left(\sum_{n=1}^{\infty} \left(\sum_{i=1}^{8} n N^i_n + \sum_{a=1}^{8} n(M^a_n \right) \right) = P. \tag{5.121}$$

As usual it is easier to pass to the canonical ensemble. Computing

$$Z(\tau) := \sum_{\{N_n^i, M_n^a\}} q^N \equiv \sum_P d(N) q^N, \quad q := e^{2\pi i \tau}, \tag{5.122}$$

ignoring the constraint. Here we have use for convenience $N = RP$ which is an integer or equivalently can absorb R into τ. One can then obtain $d(N)$ by inverse Laplace transform using

$$Z(\tau) := \sum_P d(N) q^N, \quad d(N) = \int_0^1 e^{-2\pi i N \tau} Z(\tau) d\tau. \tag{5.123}$$

The partition function is readily evaluated and is given by

$$Z(\tau) = \frac{\prod_{n=1}^{\infty}(1+q^n)^8}{\prod_{n=1}^{\infty}(1-q^n)^8} \tag{5.124}$$

From this one can find that

$$d(N) \sim e^{2\pi\sqrt{2N}}, \tag{5.125}$$

which follows also from the Cardy formula applied to the worldsheet CFT living on the circle, using the fact that for 8 free bosons and 8 free fermions the central charge is 12.

After this warm-up exercise, let us turn to the problem of motion of m D1-branes bound to a single D5-brane. Now, *a priori* the D1-brane can again oscillate in all 8 transverse directions. However, if we switch on a 2-form field along 2-cycles of $K3$, then open strings connecting D1-branes and D5-branes become tachyonic. Condensation into ground state binds the D1-branes to the D5-branes and as a result they can oscillate only along the directions along the $K3$.

We are interested in a configuration with m units of D1-brane charge n units of momentum, and l units of angular momentum. If m is divisible by s then we have to consider both the configuration with m D1-branes winding number 1 as well as the configuration with m/s D1-branes with winding number s. Similarly, the momentum and angular momentum can be shared among these m or m/s D1-branes. As usual, it is more convenient to relax all constraints on the charges and compute instead the (grand) canonical partition function. So, we introduce chemical (complexified) chemical potentials σ, τ, z conjugate to the integers m, n, l and compute the unrestricted sum by summing over all possible charges (r, s, t). The degeneracies $d_{D1}(m, n, l)$ can then be extracted by an inverse Fourier transform.

Consider a D1-brane wound r times along the S^1, carrying momentum s along the S^1 with angular momentum $J_L = t/2$. Let

$$Z_{D1} = \frac{1}{p} \prod_{s>0, t\geq 0, r} \frac{1}{(1 - p^s q^t y^r)^{c(s,t,r)}}. \tag{5.126}$$

Now, a D1-brane wrapping s times around a circle R is like a D1-brane wrapping once on a circle of effective radius $R_e = 2\pi Rs$. If we want it to carry physical momentum t, then since

$$\frac{t}{R} = \frac{ts}{nR} = \frac{ts}{R_e} \tag{5.127}$$

Because of conformal invariance, the partition function does not depend on the overall scale R. We thus conclude that the partition function for winding s and physical momentum t is the same as the partition function for winding 1 and physical momentum st. In other words,

$$c(s, t, r) = c_0(st, r). \tag{5.128}$$

These coefficients are nothing but the $c_0(n, l)$ defined in (5.232) of the elliptic genus $\chi(\tau, z)$ of a single copy of K_3. Hence $c(s, t, r) = c_0(st, r) = C_0(4st - r^2)$ from (5.233). Indeed, our computation of Z_{D1} is one way to derive the generating function $\hat{Z}$ for the elliptic genera of symmetric products of $K3$. In summary,

$$Z_{D1}(\sigma, \tau, z) = \hat{Z}(\sigma, \tau, z). \tag{5.129}$$

Comment: The problem of counting microstates of m D1-branes bound to a D5-brane is the counting problem that arises in computing the microstates of the well-known Strominger–Vafa black hole in five dimensions [33]. The microscopic configuration there consists of Q_5 D5-branes wrapping $K3 \times S^1$, Q_1 D1-branes wrapping the S^1, with total momentum n along the circle. We have chosen $Q_5 = 1$ and $Q_1 = m$ but more generally, we can simply replace m by $Q_1 Q_5$. The bound states are described by an effective string wrapping the circle carrying left-moving momentum n. The central charge of the system can be computed at weak coupling and is given by $6m$. In this system, the leading order entropy at large charge can be computed by applying the Cardy formula provided we operate in a certain regime in moduli and charge space. We work in a region of moduli space where the K3 is small compared to the S1. In such a situation, the dynamics of the D1–D5 system are encapsulated in a 1+1 D CFT living on S^1. The D1–D5-P configuration can then be regarded as a state in this CFT with the right moving oscillators fixed to their ground state and the left moving excitation number or CFT temperature proportional to n. Then in the limit of $n \gg Q1Q5$, the Cardy formula for the high temperature expansion of the CFT can be used to compute the leading order degeneracy of the state. Applying Cardy's formula therefore, gives,

$$d_m(n) = \exp(2\pi\sqrt{mn}). \tag{5.130}$$

This implies a microscopic entropy $S = \log d = 2\pi\sqrt{Q_1 Q_5 n}$. The corresponding BPS black hole solutions with three charges in five dimensions can be found in supergravity and the resulting entropy matches precisely with the macroscopic entropy [33].

5.6.4 Dynamics of the KK-Monopole

In the previous subsection we have worked out the low-energy massless fluctuations of the KK-monopole. If we excite only the left-movers then we have 24 bosons carrying momentum t. The KK-monopole cannot support any momentum along the S^1 circle. Summing over all momenta gives rise to the partition function

$$Z_{KK}(\tau) = \frac{1}{q} \prod_{t=1}^{\infty} \frac{1}{(1-q^t)^{24}} = \frac{1}{\eta^{24}(\tau)} \tag{5.131}$$

The factor of $1/q$ comes because the ground state carries some 'zero point' momentum -1. Altogether, we recognize this as precisely the partition function of the left-moving BPS oscillations of the heterotic string as expected from duality.

5.6.5 D1–D5 Center-of-Mass Oscillations

Now it remains for us to find the contribution to the partition function from the oscillations of the center of mass of the D1–D5 system moving in the background the KK-monopole. This is easy to evaluate using the fact that for large radius near the center of the KK-monopole, the Taub-NUT space is essentially flat Euclidean space $\mathscr{R}^4$. The partition function of four bosons and four fermions is simply

$$Z_{CM}(\tau, z) = \frac{\eta^6(\tau)}{\theta_1^2(\tau, z)}. \tag{5.132}$$

Putting this all together we find the desired result

$$Z(\Omega) = \frac{\hat{Z}(\sigma, \tau, z)}{\psi(\tau, z)} = \frac{1}{\Phi_{10}(\Omega)}. \tag{5.133}$$

5.6.6 Wall-Crossing and Contour Prescription

Given the partition function (5.103), one can extract the black hole degeneracies from the Fourier coefficients. However, there is one complication that also turns out to have interesting physical implications. The Igusa cusp form has double zeros at $z = 0$ and its images. The partition function is therefore a *meromorphic* Siegel form (5.225) of weight -10 with double poles at these divisors. As a result, different Fourier contours would give different answers for the degeneracies and there appears to be an ambiguity in the choice of the Fourier contour.

This ambiguity turns out to have a very nice physical interpretation. The spectrum of quarter-BPS dyons actually has a moduli dependence. For a given charge vector Γ, there are single-centered black hole solutions that exist everywhere in the moduli space. However, in addition, there can be two-centered solutions such that one center carries charge Γ_1 and the other Γ_2 with $\Gamma = \Gamma_1 + \Gamma_2$. A simple example is when one charge center has charge $(Q, 0)$ and the other has charge $(0, P)$. The distance between these two centers is fixed in terms of the charges and the moduli fields.

As one changes the moduli, the distance between the two centers can go to infinity and the two-centered solution can decay at certain walls i.e. surfaces of co-dimension one. Thus, on one side of the wall, we have only a single-centered black hole whereas on the other side we have the single-centered black hole as well as the two-centered black hole. Hence the degeneracy on one side of the wall is different from the degeneracy on the other side of the all. Upon crossing the wall, the degeneracy jumps. This phenomenon is known as the 'wall- crossing phenomenon'. The moduli space is thus divided up into chambers separated by walls. The degeneracy is different from chamber to chamber.

This dependence of the degeneracy on the chamber in the moduli space is nicely captured by the dependence of the Fourier coefficients on the choice of the contour. As we will explain below, the choice of the contour depends on the moduli in a precise way. As the moduli are varied, the contour is deformed. The dependence of the contour on the moduli is such that as the moduli hit a wall in the moduli space, the contour hits a pole of the partition function. The poles are thus nicely correlated with the walls. Crossing the wall in the moduli space corresponds to crossing a pole in the contour space. The jump in the degeneracy upon crossing the wall is given by the residue at the pole that is crossed by the contour.

To see this more precisely, note that the three quadratic T-duality invariants of a given dyonic state can be organized as a 2×2 symmetric matrix

$$\Lambda = \begin{pmatrix} Q \cdot Q & Q \cdot P \\ Q \cdot P & P \cdot P \end{pmatrix} = \begin{pmatrix} 2n & \ell \\ \ell & 2m \end{pmatrix}, \tag{5.134}$$

where the dot products are defined using the $O(22, 6; \mathbb{Z})$ invariant metric Λ. The matrix Ω in (5.102) and (5.222) can be viewed as the matrix of complex chemical potentials conjugate to the charge matrix Λ. The charge matrix Λ is manifestly T-duality invariant. Under an S-duality transformation (5.53), it transforms as

$$\Lambda \to \gamma \Lambda \gamma^t \tag{5.135}$$

There is a natural embedding of this physical S-duality group $SL(2, \mathbb{Z})$ into $Sp(2, \mathbb{Z})$:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} (\gamma^t)^{-1} & \mathbf{0} \\ \mathbf{0} & \gamma \end{pmatrix} = \begin{pmatrix} d & -c & 0 & 0 \\ -b & a & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{pmatrix} \in Sp(2, \mathbb{Z}). \tag{5.136}$$

The embedding is chosen so that $\Omega \to (\gamma^T)^{-1} \Omega \gamma^{-1}$ and $\mathrm{Tr}(\Omega \cdot \Lambda)$ in the Fourier integral is invariant. This choice of the embedding ensures that the physical degeneracies extracted from the Fourier integral are S-duality invariant if we appropriately transform the moduli at the same time as we explain below.

To specify the contours, it is useful to define the following moduli-dependent quantities. One can define the matrix of right-moving T-duality invariants

$$\Lambda_R = \begin{pmatrix} Q_R \cdot Q_R & Q_R \cdot P_R \\ Q_R \cdot P_R & P_R \cdot P_R \end{pmatrix}. \tag{5.137}$$

which depends both on the integral charge vectors N, M as well as the T-moduli μ. One can then define two matrices naturally associated to the S-moduli $\lambda = \lambda_1 + i\lambda_2$ and the T-moduli μ respectively by

$$\mathscr{S} = \frac{1}{\lambda_2}\begin{pmatrix} |\lambda|^2 & \lambda_1 \\ \lambda_1 & 1 \end{pmatrix}, \quad \mathscr{T} = \frac{\Lambda_R}{|\det(\Lambda_R)|^{\frac{1}{2}}}. \tag{5.138}$$

Both matrices are normalized to have unit determinant. In terms of them, we can construct the moduli-dependent 'central charge matrix'

$$\mathscr{Z} = |\det(\Lambda_R)|^{\frac{1}{4}}\left(\mathscr{S} + \mathscr{T}\right), \tag{5.139}$$

whose determinant equals the BPS mass

$$M_{Q,P} = |\det \mathscr{Z}|. \tag{5.140}$$

We define

$$\tilde{\Omega} \equiv \begin{pmatrix} \sigma & -z \\ -z & \tau \end{pmatrix} \tag{5.141}$$

related to Ω by an $SL(2,\mathbb{Z})$ transformation

$$\tilde{\Omega} = \hat{S}\Omega\hat{S}^{-1} \quad where \quad \hat{S} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{5.142}$$

so that, under a general S-duality transformation γ, we have the transformation $\tilde{\Omega} \to \gamma\tilde{\Omega}\gamma^T$ as $\Omega \to (\gamma^T)^{-1}\Omega\gamma^{-1}$.

With these definitions, Λ, Λ_R, $\mathscr{Z}$ and $\tilde{\Omega}$ all transform as $X \to \gamma X \gamma^T$ under an S-duality transformation (5.53) and are invariant under T-duality transformations. The moduli-dependent Fourier contour can then be specified in a duality-invariant fashion by [37]

$$\mathscr{C} = \{\mathrm{Im}\,\tilde{\Omega} = \varepsilon^{-1}\mathscr{Z}; \quad 0 \le \mathrm{Re}(\tau), \mathrm{Re}(\sigma), \mathrm{Re}(z) < 1\}, \tag{5.143}$$

where $\varepsilon \to 0^+$. For a given set of charges, the contour depends on the moduli λ, μ through the definition of the central charge vector (5.139). The degeneracies

$d(m, n, l)|_{\lambda,\mu}$ of states with the T-duality invariants (m, n, l), at a given point (λ, μ) in the moduli space are then given by[5]

$$d(m, n, l)|_{\lambda,\mu} = \int_{\mathscr{C}} e^{-i\pi \mathrm{Tr}(\Omega\cdot\Lambda)} Z(\Omega) d^3\Omega. \tag{5.144}$$

This contour prescription thus specifies how to extract the degeneracies from the partition function for a given set of charges and in any given region of the moduli space. In particular, it also completely summarizes all wall-crossings as one moves around in the moduli space for a fixed set of charges. Even though the indexed partition function has the same functional form throughout the moduli space, the spectrum is moduli dependent because of the moduli dependence of the contours of Fourier integration and the pole structure of the partition function. Since the degeneracies depend on the moduli *only* through the dependence of the contour $\mathscr{C}$, moving around in the moduli space corresponds to deforming the Fourier contour.

With this understanding of the wall crossing and the contour prescription, we have completely specified how to extract dyon degeneracies from the Fourier coefficients of the partition function. The partition function in turn is constructed explicitly in terms of Fourier coefficients of known objects such as ψ or χ. We will not here analyze wall-crossing in any further detail which can be found in [24, 37, 38].

5.6.7 Asymptotic Expansion

Given the exact formula for the degeneracies, one can try to extract the asymptotic degeneracies in the limit where *m, n* are both large and positive. Since the Fourier integral now involves three variables, the calculation is more involved than the Cardy formula that we encountered for modular forms of single variable. The answer however, is simple. The statistical entropy $\log(d)$ is obtained by minimizing the following function with respect to λ

$$\mathscr{E}_B(\lambda) = \frac{\pi}{2\lambda_2}|Q + \lambda P|^2 - 64\pi^2\phi(\lambda, \bar{\lambda}) + O(Q^{-2}), \tag{5.145}$$

where $\phi(\lambda, \bar{\lambda})$:

$$\phi(\lambda, \bar{\lambda}) = -\frac{1}{64\pi^2}\left\{12\log\left[-2i(\lambda - \bar{\lambda})\right] + 24\log\left[\eta(\lambda)\right] + 24\log\left[\eta(\bar{\lambda})\right]\right\}. \tag{5.146}$$

For a detailed description of the expansion, see [36, 39].

[5] The physical degeneracies have an additional multiplicative factor of $(-1)^{\ell+1}$ which we omit here for simplicity of notation in later chapaters.

5.7 Quantum Black Holes

Now we turn to the black holes in string theory that corresponds to the ensembles of the BPS quantum microstates. Such dyonic BPS black holes are essentially generalizations of the Reissner–Nordström black hole but now with both electric and magnetic charges under several different $U(1)$ gauge fields. They are solutions of the effective action of string theory which contains many more terms compared to the Einstein–Maxwell action (5.1).

To view a black hole as an ensemble of states, it is important to find full the black hole solution of the effective action that connects the near horizon region that we analyze below to an asymptotically flat spacetime. For the leading two-derivative effective action of toroidally compactified heterotic string theory, such exact interpolating solutions for dyonic BPS black holes are known [40, 41]. The black hole geometry exhibits the attractor mechanism: the values of scalar fields get 'attracted' to their atttactor values at the horizon that are determined entirely by the charges of the black hole and independent of their values at asymptotic infinity [42–44]. Incorporating the effect higher-derivative terms in the effective action for the interpolating solutions is in general much more complicated and can be found in [45–48].

For our purposes, we are only interested in the near-horizon properties of the black hole such as its entropy and the attractor values of various scalar fields at the horizon. This can be analyzed much more simply using the entropy function formalism developed in Sect. 5.3.6.

In Sect. 5.7.1 we discuss the near horizon solution and the entropy for the leading two-derivative effective action and consider the correction to the Wald entropy to the next subleading order in Sect. 5.7.2. They compare beautifully with statistical entropy given by the logarithm of the microscopic degeneracies computed in the Sect. 5.6.

The case of black holes corresponding to the half-BPS states is in some ways more interesting which we discuss in Sect. 5.7.3. In this case, the entropy is actually zero to leading order because the geometry has a null singularity instead of a smooth horizon. The area of the event horizon is thus zero to leading order. Subleading quantum corrections modify the geometry so that the corrected geometry has a string scale horizon. The Wald entropy associated with this horizon precisely matches with the statistical entropy computed in Sect. 5.5.

5.7.1 Wald Entropy to Leading Order

For a state with electric charge vector q and magnetic charge vector p, the fields near the horizon take the form[6]

[6] For an extensive description of this computation see [49].

$$\begin{aligned} ds^2 =& \frac{v_1}{16}\left(-(\sigma^2-1)d\tau^2 + \frac{d\sigma^2}{\sigma^2-1}\right) + \frac{v_2}{16}\left(d\theta^2 + \sin^2\theta d\phi^2\right) \\ F^{(i)}_{\sigma\tau} =& \frac{1}{4}e_i, \quad F^{(i)}_{\theta\phi} = \frac{1}{16\pi}p_i, \quad M_{ij} = u_{ij}, \quad S = u_s, \quad a = u_a. \end{aligned} \tag{5.147}$$

Substituting into the action (5.64) we get

$$\begin{aligned} f(u_S, u_a, u_M, \mathbf{v}, \mathbf{e}, \mathbf{p}) \equiv& \int d\theta d\phi \sqrt{-\det G}\mathscr{L} \\ =& \frac{1}{8}v_1 v_2 u_S \left[-\frac{2}{v_1} + \frac{2}{v_2} + \frac{2}{v_1^2} e_i (L u_M L)_{ij} e_j \right. \\ & \left. - \frac{1}{8\pi^2 v_2^2} p_i (L u_M L)_{ij} p_j + \frac{u_a}{\pi u_S v_1 v_2} e_i L_{ij} p_j \right]. \end{aligned} \tag{5.148}$$

Hence the entropy function becomes

$$\begin{aligned} \mathscr{E}(q, u_S, u_a, u_M, v, e, p) :=& 2\pi \left(e_i q_i - f(u_S, u_a, u_M, v, e, p)\right) \\ =& 2\pi \left[e_i q_i - \frac{1}{8} v_1 v_2 u_S \left\{ -\frac{2}{v_1} + \frac{2}{v_2} + \frac{2}{v_1^2} e_i (L u_M L)_{ij} e_j \right.\right. \\ & \left.\left. - \frac{1}{8\pi^2 v_2^2} p_i (L u_M L)_{ij} p_j + \frac{u_a}{\pi u_S v_1 v_2} e_i L_{ij} p_j \right\}\right]. \end{aligned} \tag{5.149}$$

Eliminating e_i from (5.46) using the equation $\partial\mathscr{E}/\partial e_i = 0$ we get:

$$\begin{aligned} & \mathscr{E}(q, u_S, u_a, u_M, v, e(u, v, q, p), p) \\ &= 2\pi \left[\frac{u_S}{4}(v_2 - v_1) + \frac{v_1}{v_2 u_S} q^T u_M q + \frac{v_1}{64\pi^2 v_2 u_S}(u_S^2 + u_a^2) p^T L u_M L p \right. \\ & \left. - \frac{v_1}{4\pi v_2 u_S} u_a q^T u_M L p \right]. \end{aligned}$$

We can simplify the formulæ by defining new charge vectors:

$$Q_i = 2q_i, \qquad P_i = \frac{1}{4\pi} L_{ij} p_j, \tag{5.150}$$

which are normalized so that they are integral and satisfy the Dirac quantization condition. In terms of $\mathbf{Q}$ and $\mathbf{P}$ the entropy function $\mathscr{E}$ is given by:

$$\mathscr{E} = \frac{\pi}{2}\left[u_S(v_2 - v_1) + \frac{v_1}{v_2 u_S}\left(Q^T u_M Q + (u_S^2 + u_a^2) P^T u_M P - 2u_a Q^T u_M P \right) \right]. \tag{5.151}$$

Substituting (5.159) into (5.151) and using (5.155, 5.156), we get:

$$\mathscr{E} = \frac{\pi}{2}\left[u_S(v_2 - v_1) + \frac{v_1}{v_2}\left\{\frac{Q^2}{u_S} + \frac{P^2}{u_S}(u_S^2 + u_a^2) - 2\frac{u_a}{u_S}Q \cdot P\right\}\right]. \tag{5.152}$$

Note that we have expressed the right hand side of this equation in an T-duality invariant form. Written in this manner, Eq. 5.152 is valid for general **P, Q** satisfying

$$P^2 > 0, \quad Q^2 > 0, \quad (Q \cdot P)^2 < Q^2 P^2. \tag{5.153}$$

We now need to find the extremum of $\mathscr{E}$ with respect to $u_S,\ u_a,\ u_{Mij},\ v_1$ and v_2. In general this leads to a complicated set of equations. We can simplify the analysis by using the $O(22, 6; \mathbb{R})$ symmetries (5.60) of the two-derivative action (5.64) which induces the following transformations on the various parameters:

$$\begin{aligned} e_i \to \Omega_{ij} e_j, \quad p_i \to \Omega_{ij} p_j, \quad u_M \to \Omega u_M \Omega^T, \\ q_i \to (\Omega^T)^{-1}_{ij} q_j, \quad Q_i \to (\Omega^T)^{-1}_{ij} Q_j, \quad P_i \to (\Omega^T)^{-1}_{ij} P_j. \end{aligned} \tag{5.154}$$

The entropy function (5.151) is invariant under these transformations. Since at its extremum with respect to u_{Mij} the entropy function depends only on **P, Q**, $v_1,\ v_2,\ u_S$ and u_a it must be a function of the $O(22, 6)$ invariant combinations:

$$Q^2 = Q_i L_{ij} Q_j, \quad P^2 = P_i L_{ij} P_j, \quad Q \cdot P = Q_i L_{ij} P_j, \tag{5.155}$$

besides $v_1,\ v_2,\ u_S$ and u_a. Let us for definiteness take $Q^2 > 0,\ P^2 > 0$, and $(Q \cdot P)^2 < Q^2 P^2$. In that case with the help of an $SO(22, 6)$ transformation we can make

$$(I_r - L)_{ij} Q_j = 0, \quad (I_r - L)_{ij} P_j = 0, \tag{5.156}$$

where I_r denotes the $r \times r$ identity matrix. This is most easily seen by diagonalizing L to the form

$$\begin{pmatrix} -I_{22} & 0_6 \\ 0_{22} & I_6 \end{pmatrix}. \tag{5.157}$$

In this case Q and P satisfying (5.156) will have

$$Q_i = 0, \quad P_i = 0, \quad \text{for } 1 \le i \le 22. \tag{5.158}$$

Let us now see that for P and Q satisfying this condition, every term in (5.151) is extremized with respect to u_M for

$$u_M = I_r. \tag{5.159}$$

Clearly a variation δu_{Mij} with either i or j in the range $[7, r]$ will give vanishing contribution to each term in $\delta\mathscr{E}$ computed from (5.151). On the other hand due to the constraint (5.58) on M, any variation δM_{ij} (and hence δu_{Mij}) with $1 \leq i, j \leq 6$ must vanish, since in this subspace satisfying (5.58) requires M to be both symmetric and orthogonal. Thus each term in $\delta\mathscr{E}$ vanishes under all allowed variations of u_M.

We should emphasize that (5.159) is not the only possible value of u_M that extremizes $\mathscr{E}$. Any u_M related to (5.159) by an $O(22, 6)$ transformation that preserves the vectors $\mathbf{Q}$ and $\mathbf{P}$ will extremize $\mathscr{E}$. Thus there is a family of extrema representing flat directions of $\mathscr{E}$. However, as we have argued in Sect. 5.3.4, the value of the entropy is independent of the choice of u_M.

It remains to extremize $\mathscr{E}$ with respect to v_1, v_2, u_S and u_a. Extremization with respect to v_1 and v_2 give:

$$v_1 = v_2 = u_S^{-2}\left(Q^2 + P^2(u_S^2 + u_a^2) - 2u_a Q\cdot P\right). \tag{5.160}$$

Substituting this into (5.152) gives:

$$\mathscr{E} = \frac{\pi}{2}\frac{1}{u_S}\left\{Q^2 - 2u_a Q\cdot P + P^2(u_S^2 + u_a^2)\right\}. \tag{5.161}$$

It is convenient to write it in a manifestly $SL(,\mathbb{Z})$ invariant way as

$$\mathscr{E} = \frac{\pi}{2}\frac{1}{\lambda_2}|Q + \lambda P|^2. \tag{5.162}$$

if we write $\lambda = u_a + iu_S := \lambda_1 + i\lambda_2$.

Finally, extremizing with respect to u_a, u_S we get

$$u_S = \frac{\sqrt{Q^2P^2 - (Q\cdot P)^2}}{P^2}, \quad u_a = \frac{Q\cdot P}{P^2}, \quad v_1 = v_2 = 2P^2. \tag{5.163}$$

The black hole entropy, given by the value of $\mathscr{E}$ for this configuration, is

$$S_{BH} = \pi\sqrt{Q^2P^2 - (Q\cdot P)^2}. \tag{5.164}$$

To get an idea about orders of magnitude let us take $Q\cdot P = 0$ for simplicity. Then from (5.164) the radius r_H of the horizon of the black hole scales as

$$r_H^2 \sim \sqrt{Q^2P^2}\,\ell_4^2 \tag{5.165}$$

where ℓ_4 four-dimensional planck length. The four dimensional string coupling g_4^2 at the horizon can be read off from the attractor value of the dilaton in (5.163):

$$g_4^2 = \frac{1}{u_S} = \sqrt{\frac{P^2}{Q^2}}. \tag{5.166}$$

We see that string loop corrections are small if $P^2 \ll Q^2$. The string length ℓ_s is related the Planck length by

$$\ell_4 = g_4 \ell_s. \tag{5.167}$$

Hence the α' corrections are small if the radius curvature is large in string units, that is, if

$$r_H^2/\ell_s^2 \sim P^2 \gg 1. \tag{5.168}$$

Hence if we take $Q^2 \gg P^2 \gg 1$, we can compute the Wald entropy in a systematic expansion in $1/Q^2$ keeping both the α' and string loop corrections small.

5.7.2 Subleading Corrections to the Wald Entropy

The asymptotic expansion in Sect. 5.6.7 is obtained in the regime when all charges scale the same way and are much larger than one. In other words,

$$Q^2 \sim P^2 \gg 1. \tag{5.169}$$

We have already computed the leading order entropy for in section (5.7.1). We would now like to see how to take the effects of higher order corrections. Let us suppose the Lagrangian is of the form

$$\mathscr{L} = \mathscr{L}_0 + \varepsilon \mathscr{L}_1, \tag{5.170}$$

where the term of order ε is a small correction from higher-derivative terms. The entropy function defined using this Lagrangian will also be of the form

$$\mathscr{E} = \mathscr{E}_0 + \varepsilon \mathscr{E}_1. \tag{5.171}$$

The solutions of the extremization equations will also have an expansion

$$\begin{aligned} e^*(q,p) &= e^*_{(0)} + \varepsilon e^*_{(1)} + \ldots; \\ u^*(q,p) &= u^*_{(0)} + \varepsilon u^*_{(1)} + \ldots; \quad v^*(q,p) = v^*_{(0)} + \varepsilon v^*_{(1)} + \ldots. \end{aligned} \tag{5.172}$$

To compute the entropy we have to compute the value of the entropy function $\mathscr{E}^*$ at the extermum

$$\mathscr{E}^*(q,p) = \mathscr{E}_0(q,u^*,v^*,e^*,p) + \varepsilon \mathscr{E}_1(q,u^*,v^*,e^*,p). \tag{5.173}$$

If we are interested in the first subleading correction to order ε we simply expand these functions to obtain

$$\mathscr{E}^*(q,p) = \mathscr{E}_0(q,u^*_0,v^*_0,e^*_0,p) + \varepsilon \mathscr{E}_1(q,u^*_0,v^*_0,e^*_0,p) + O(\varepsilon^2). \tag{5.174}$$

The important point is that to $O(\varepsilon)$ one could have had terms like

$$\frac{\partial \mathscr{E}_0}{\partial e}, \quad \frac{\partial \mathscr{E}_0}{\partial v}, \quad \frac{\partial \mathscr{E}_0}{\partial u}, \tag{5.175}$$

evaluated at the leading order extremum values u_0^*, v_0^*, e_0^*. However, these all vanish because to the leading order, the extremum values of near horizon fields are found precisely by setting all terms in (5.175) to zero. Hence, to find the first subleading correction, it is not necessary to solve the extermization equations all over again. It suffices to evaluate the correction to the entropy $\mathscr{E}_1$ at the extremum values found using the zeroth order entropy function $\mathscr{E}_0$. This greatly simplify practical computations.

To illustrate these ideas, we apply them to the heterotic action for the dyonic black holes of our interest. The heterotic supergravity action (5.64) is only the leading 2-derivative supergravity approximation to the full string effective action. The theory has a 4-derivative correction to the effective action given by the lagrangian

$$\Delta\mathscr{L} = \phi(\lambda, \bar{\lambda})\left(R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} - 4R_{\mu\nu}R^{\mu\nu}\right), \tag{5.176}$$

where $\phi(\lambda, \bar{\lambda})$ is a nontrivial function of axion-dilaton $\lambda := a + iS$:

$$\phi(\lambda, \bar{\lambda}) = -\frac{1}{64\pi^2}\left[12\log(S) + 24\log\left(\eta(a - iS)\right) + 24\log\left(\eta(a + iS)\right)\right]. \tag{5.177}$$

Note that this is exactly the same function $\phi(\lambda, \bar{\lambda})$ introduced in (5.146). It is easy to check that addition of this term induces a correction to the entropy function of the form

$$\mathscr{E}_1 = 64\pi^2\phi(\lambda, \bar{\lambda}). \tag{5.178}$$

Consequently, the Wald entropy corrected to this order is then given by

$$S_{wald} = \pi\sqrt{Q^2P^2 - (Q\cdot P)^2} + 64\pi^2\phi\left(a = \frac{Q\cdot P}{P^2}, S = \frac{\sqrt{Q^2P^2 - (Q\cdot P)^2}}{P^2}\right) + \ldots \tag{5.179}$$

As a result, the thermodynamic Wald entropy given by (5.179) matches beautifully with the statistical entropy given by (5.145) not only to the leading order but also the next subleading order. As mentioned in the preface, the subleading finite size corrections have much more structure than the leading Bekenstein–Hawking entropy and involve a rather nontrivial modular function ϕ.

We should emphasize that the origin of this function in the two computations is of totally different. In the computation of the Wald entropy $S_{wald}(Q, P)$, it arises from specific terms in the effective action of massless fields in string theory. In the computation of the statistical entropy $\log(d(Q, P))$, on the other hand, it arises from the asymptotic expansion of the Fourier coefficients of the partition function for quarter-BPS dyons which for some reason is related the Igusa cusp form. This thus points to a highly nontrivial internal consistency in the structure of string theory and gives us some confidence that we may be on the right track in the search for a quantum theory of gravity.

5.7.3 Wald Entropy of Small Black Holes

For half-BPS black holes, we can choose a duality frame in which they are purely perturbative with electric charge vector Q and no magnetic charge, or $P = 0$. In this case, it follows from (5.163) and (5.164) that the near horizon solution of the leading order two derivative action is singular. In particular, the area of the horizon goes to zero and the attractor value of the string coupling constant goes to zero. Thus, in this case it is not sensible to study the effects of higher derivative terms as small corrections to the leading order solution. Rather, one must consider the full entropy function and find the near horizon geometry by extremizing it. It turns out that upon the inclusion of α' corrections, the near horizon geometry is no longer singular but has a horizon with area of order one in string units. Such black holes with a small string scale horizon have been termed 'small' black holes [50, 51]. Moreover, the Wald entropy of this horizon precisely agrees with the statistical entropy [52, 53]. This is an interesting phenomenon which illustrates that quantum corrections within string theory can modify classical geometry to generate a horizon whose properties are in accordance with the microscopic theory.

To illustrate how this works out, let us analyze for simplicity the effect of the following four-derivative term in the string effective action

$$\Delta\mathscr{L} = \frac{S}{64\pi^2}\left(R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta} - 4R_{\mu\nu}R^{\mu\nu}\right), \tag{5.180}$$

Now for the total entropy function, instead of (5.162), one obtains

$$\mathscr{E} = \frac{\pi}{2}\left(\frac{Q^2}{u_S} + 8u_S\right). \tag{5.181}$$

Extremizing with respect to u_S, we obtain the attractor value of the dilaton field

$$u_S^* = \sqrt{Q^2/8}, \tag{5.182}$$

and hence the Wald entropy is given by

$$S_{Wald} := \mathscr{E}^*(Q) := \mathscr{E}(u_S^*(Q)) = 4\pi\sqrt{Q^2/2}, \tag{5.183}$$

which matches beautifully with the statistical entropy (5.99).

We should remember though that since the horizon area is of order one in string units, all α' corrections are of the same order and hence the effect of all higher-derivative terms must be included at once. It turns out, however, that even upon including the effect of all supersymmetrized F-type terms [52, 53] one obtains the same results.[7]

A general scaling argument [54] shows that up to an over all constant, the Wald entropy must have the same form as (5.183) even after all α' corrections are included up to. Moreover, by viewing the four-dimensional small black hole as an excitation of a five-dimensional black string it has been shown in [55, 56] the Wald entropy is related to the coefficient of five-dimensional Chern-Simons terms. Since Chern-Simons terms are topological in nature, their coefficient is not renormalized even after including higher quantum correction. Together, these results strongly indicate that Wald entropy of small black holes upon including stringy all α' corrections will agree with the statistical entropy.

The agreement above and also for the entropy of quarter-BPS dyons in Sect. 5.7.2 is obtained using only the F-type terms in the string effective action. This strongly suggests a nonrenormalization theorem that other D-terms do not renormalize the Wald entropy. For a subclass of D-type terms such a nonrenormalization theorem has recently been proven [57]. It would be interesting to see how it can be generalized to all possible D-terms in this context.

5.8 Mathematical Background

5.8.1 $\mathcal{N} = 4$ Supersymmetry

We summarize here some facts about the representation of the $\mathcal{N} = 4$ superalgebra. For more details see for example [58].

5.8.1.1 Massless Supermultiplets

There are two massless representations that will be of interest to us.

1. Supergravity multiplet:
   It contains the metric $g_{\mu\nu}$, six vectors $A_\mu^{(ab)}$, and two gravitini $\psi^a_{\mu\alpha}$.
2. Vector Multiplet:
   It contains a vector A_μ, six scalar fields $X^{(ab)}$, and the gaugini χ^a_α.

The low energy massless spectrum of a supergravity theory consists of the supergravity multiplet and n_v vector multiplets. Supersymmetry then completely fixes the form of the two derivative action. The compactification of heterotic string theory on T^6 leads to a theory in four spacetime dimensions with $\mathcal{N} = 4$ supersymmetry and 28 abelian gauge fields which corresponds to $28 - 6 = 22$ vector multiplets.

---

7 F-type terms can be written as chiral integrals on superspace.

5.8.1.2 General BPS Representations

In the rest frame of the dyon, the $\mathcal{N} = 4$ supersymmetry algebra takes the form

$$\{Q^a_\alpha, Q^{\dagger b}_{\dot\beta}\} = M\delta_{\alpha\dot\beta}\delta^{ab}, \quad \{Q^a_\alpha, Q^b_\beta\} = \varepsilon_{\alpha\beta}Z^{ab}, \quad \{Q^{\dagger a}_{\dot\alpha}, Q^{\dagger b}_{\dot\beta}\} = \varepsilon_{\dot\alpha\dot\beta}\bar{Z}^{ab} \tag{5.184}$$

where $a, b = 1, \ldots 4$ are $SU(4)$ R-symmetry indices and α, β are Weyl spinor indices. In a given charge sector, the central charge matrix encodes information about the charges and the moduli. To write it explicitly, we first define a central charge vector in $\mathscr{C}^6$

$$Z^m(\Gamma) = \frac{1}{\sqrt{\tau_2}}(Q^m_R - \tau P^m_R), \quad m = 1, \ldots 6, \tag{5.185}$$

which transforms in the (complex) vector representation of $Spin(6)$. Using the equivalence $Spin(6) = SU(4)$, we can relate it to the antisymmetric representation of Z_{ab} by

$$Z_{ab}(\Gamma) = \frac{1}{\sqrt{\tau_2}}(Q_R - \tau P_R)^m \lambda^m_{ab}, \quad m = 1, \ldots 6 \tag{5.186}$$

where λ^m_{ab} are the Clebsch–Gordon matrices. Since $Z(\Gamma)$ is antisymmetric, it can be brought to a block-diagonal form by a $U(4)$ rotation

$$\tilde{Z} = UZU^T, \; U \in U(4), \quad \tilde{Z}_{ab} = \left(\begin{array}{c|c} Z_1\varepsilon & 0 \\ \hline 0 & Z_2\varepsilon \end{array}\right), \quad \varepsilon = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \tag{5.187}$$

where Z_1 and Z_2 are non-negative real numbers. A $U(2)$ rotation in the 12 plane and another $U(2)$ rotation in the 34 plane will not change the block diagonal form. Since ε is the invariant tensor of $SU(2)$, the $U(2) \times U(2)$ transformation can only change independently the phases of Z_1 and Z_2. We will therefore treat more generally Z_1 and Z_2 as complex numbers.

We now split the $SU(4)$ index as $a = (r, i)$, where $r, i = 1, 2$ and i represents the block number. Defining the following fermionic oscillators

$$\mathscr{A}^i_\alpha = \frac{1}{\sqrt{2}}(\mathscr{Q}^{1i}_\alpha + \varepsilon_{\alpha\beta}\mathscr{Q}^{\dagger 2i}_\beta), \quad \mathscr{B}^i_\alpha = \frac{1}{\sqrt{2}}(\mathscr{Q}^{1i}_\alpha - \varepsilon_{\alpha\beta}\mathscr{Q}^{\dagger 2i}_\beta), \quad \mathscr{Q}^a = U^a_b Q^b \tag{5.188}$$

the supersymmetry algebra takes the form

$$\{\mathscr{A}^{i\dagger}_{\dot\alpha}, \mathscr{A}^j_\beta\} = (M + Z_i)\delta_{\dot\alpha\beta}\delta^{ij}, \quad \{\mathscr{B}^{i\dagger}_{\dot\alpha}, \mathscr{B}^j_\beta\} = (M - Z_i)\delta_{\dot\alpha\beta}\delta^{ij} \tag{5.189}$$

with all other anti-commutators being zero.

Let us conclude by giving an explicit representation for λ^m_{ab}. An $SU(4)$ rotation which rotates the supercharges, $Q' = UQ$, acts on the Clebsch–Gordon matrices as

$$U\lambda^m U^T = R^m{}_n(U)\lambda^m \tag{5.190}$$

where $R^m{}_n$ is an $SO(6)$ rotation matrix. The Clebsch–Gordon matrices λ^m_{ab} are given by the components $(C\Gamma^m)_{ab}$ where Γ^m are the Dirac matrices of $Spin(5)$ in the Weyl basis satisfying the Clifford algebra $\{\Gamma^m, \Gamma^n\} = 2\delta^{mn}$, and C is the charge conjugation matrix. The Gamma matrices are given explicitly in terms of Pauli matrices by

$$\Gamma^1 = \sigma_1 \times \sigma_1 \times 1, \quad \Gamma^4 = \sigma_2 \times 1 \times \sigma_1 \tag{5.191}$$

$$\Gamma^2 = \sigma_1 \times \sigma_2 \times 1, \quad \Gamma^5 = \sigma_2 \times 1 \times \sigma_2 \tag{5.192}$$

$$\Gamma^3 = \sigma_1 \times \sigma_3 \times 1, \quad \Gamma^6 = \sigma_2 \times 1 \times \sigma_3, \tag{5.193}$$

where the charge conjugation matrix is defined by $C\Gamma^m C^{-1} = -\Gamma^{m*}$

$$C = \sigma_1 \times \sigma_2 \times \sigma_2, \quad \Gamma = \sigma_3 \times 1 \times 1, \quad C\Gamma^m = \begin{pmatrix} \lambda^m_{ab} & 0 \\ 0 & \bar{\lambda}^m_{\dot{a}\dot{b}} \end{pmatrix} \tag{5.194}$$

where the un-dotted indices transform in the spinor representation of $Spin(6)$ or the 4 of $SU(4)$ whereas the the dotted indices transform in the conjugate spinor representation of $Spin(6)$ or the $\bar{4}$ of $SU(4)$. The matrices λ^m_{ab} thus defined have the required antisymmetry and transform properties as in (5.190).

5.8.2 *Modular Cornucopia*

We assemble here together some properties of modular forms, Jacobi forms, and Siegel modular forms.

5.8.2.1 Modular Forms

Let $\mathbb{H}$ be the upper half plane, i.e., the set of complex numbers τ whose imaginary part satisfies $\mathrm{Im}(\tau) > 0$. Let $SL(2, \mathbb{Z})$ be the group of matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5.195}$$

with integer entries such that $ad - bc = 1$.

A *modular form* $f(\tau)$ of weight k on $SL(2, \mathbb{Z})$ is a holomorphic function on $\mathscr{H}$, that transforms as

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k f(\tau) \quad \forall \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}), \tag{5.196}$$

for an integer k (necessarily even if $f(0) \neq 0$). It follows from the definition that $f(\tau)$ is periodic under $\tau \to \tau + 1$ and can be written as a Fourier series

$$f(\tau) = \sum_{n=-\infty}^{\infty} a(n)q^n, \quad q := e^{2\pi i\tau}, \tag{5.197}$$

and is bounded as $\mathrm{Im}(\tau) \to \infty$. If $a(0) = 0$, then the modular form vanishes at infinity and is called a *cusp form*. Conversely, one may weaken the growth condition at ∞ to $f(\tau) = \mathcal{O}(q^{-N})$ rather than $\mathcal{O}(1)$ for some $N \geq 0$; then the Fourier coefficients of f have the behavior $a(n) = 0$ for $n < -N$. Such a function is called a *weakly holomorphic modular form*.

The vector space over $\mathbb{C}$ of holomorphic modular forms of weight k is usually denoted by M_k. Similarly, the space of cusp forms of weight k and the space of weakly holomorphic modular forms of weight k are denoted by S_k and $M_k^!$ respectively. We thus have the inclusion

$$S_k \subset M_k \subset M_k^!. \tag{5.198}$$

The growth properties of Fourier coefficients of modular forms are known:

1. $f \in M_k^! \Rightarrow a_n = \mathcal{O}(e^{C\sqrt{n}})$ as $n \to \infty$ for some $C > 0$;
2. $f \in M_k \Rightarrow a_n = \mathcal{O}(n^{k-1})$ as $n \to \infty$;
3. $f \in S_k \Rightarrow a_n = \mathcal{O}(n^{k/2})$ as $n \to \infty$.

Some important modular forms on $SL(2, \mathbb{Z})$ are:

1. The *Eisenstein series* $E_k \in M_k$ ($k \geq 4$). The first two of these are

$$E_4(\tau) = 1 + 240 \sum_{n=1}^{\infty} \frac{n^3 q^n}{1-q^n} = 1 + 240q + \ldots, \tag{5.199}$$

$$E_6(\tau) = 1 - 504 \sum_{n=1}^{\infty} \frac{n^5 q^n}{1-q^n} = 1 - 504q + \ldots. \tag{5.200}$$

2. The *discriminant function* Δ. It is given by the product expansion

$$\Delta(\tau) = q \prod_{n=1}^{\infty} (1-q^n)^{24} = q - 24q^2 + 252q^3 + \ldots \tag{5.201}$$

or by the formula $\Delta = \left(E_4^3 - E_6^2\right)/1728$.

The two forms E_4 and E_6 generate the ring of modular forms, so that any modular form of weight k can be written (uniquely) as a sum of monomials $E_4^\alpha E_6^\beta$ with $4\alpha + 6\beta = k$. We also have $M_k = \mathbf{C} \cdot E_k \oplus S_k$ and $S_k = \Delta \cdot M_{k-12}$, so that any $f \in M_k$ also has a unique expansion as $\sum_{0 \leq n \leq k/12} \alpha_n E_{k-12n} \Delta^n$ (with $E_0 = 1$). From either representation, we see that a modular form is uniquely determined by its weight and first few Fourier coefficients.

5.8.2.2 Jacobi Forms

Consider a holomorphic function $\varphi(\tau, z)$ from $\mathbb{H} \times \mathbb{C}$ to $\mathbb{C}$ which is "modular in τ and elliptic in z" in the sense that it transforms under the modular group as

$$\varphi\left(\frac{a\tau+b}{c\tau+d}, \frac{z}{c\tau+d}\right) = (c\tau+d)^k e^{\frac{2\pi i m c z^2}{c\tau+d}} \varphi(\tau, z), \quad \forall \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2;\mathbb{Z}) \tag{5.202}$$

and under the translations of z by $\mathbb{Z}\tau + \mathbb{Z}$ as

$$\varphi(\tau, z+\lambda\tau+\mu) = e^{-2\pi i m(\lambda^2\tau+2\lambda z)}\varphi(\tau, z), \quad \forall \quad \lambda, \mu \in \mathbb{Z}, \tag{5.203}$$

where k is an integer and m is a positive integer.

These equations include the periodicities $\varphi(\tau+1, z) = \varphi(\tau, z)$ and $\varphi(\tau, z+1) = \varphi(\tau, z)$, so φ has a Fourier expansion

$$\varphi(\tau, z) = \sum_{n,r} c(n, r) q^n y^r \qquad (q := e^{2\pi i\tau}, \ y := e^{2\pi i z}). \tag{5.204}$$

Equation (5.203) is then equivalent to the periodicity property

$$c(n, r) = C(4nm - r^2; r) \quad \text{where } C(d; r) \text{ depends only on } r(\text{mod } 2m). \tag{5.205}$$

The function $\varphi(\tau, z)$ is called a *holomorphic Jacobi form* (or simply a *Jacobi form*) of weight k and index m if the coefficients $C(d;r)$ vanish for $d < 0$, i.e. if

$$c(n, r) = 0 \quad \text{unless} \quad 4mn \geq r^2. \tag{5.206}$$

It is called a *Jacobi cusp form* if it satisfies the stronger condition that $C(d; r)$ vanishes unless d is strictly positive, i.e.

$$c(n, r) = 0 \quad \text{unless} \quad 4mn > r^2, \tag{5.207}$$

and conversely, it is called a *weak Jacobi form* if it satisfies the weaker condition

$$c(n, r) = 0 \quad \text{unless} \quad n \geq 0 \tag{5.208}$$

rather than (5.206).

5.8.2.3 Theta Functions

In this section, we collect definitions and useful properties of theta function. The Jacobi theta function is defined by

$$\vartheta\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right](v|\tau) = \sum_{n\in\mathbb{Z}} q^{\frac{1}{2}(n-a)^2} e^{2\pi i(v-b)(n-a)}, \tag{5.209}$$

where a, b are real and $q = e^{2\pi i\tau}$. It satisfies the modular properties

$$\vartheta[^a_b](v|\tau+1) = e^{-i\pi a(a-1)}\vartheta[^{a}_{a+b-\frac{1}{2}}](v|\tau) \tag{5.210}$$

$$\vartheta[^a_b]\left(\frac{v}{\tau}\middle|-\frac{1}{\tau}\right) = e^{2i\pi ab+i\pi\frac{v^2}{\tau}}\vartheta[^a_b](v|\tau) \tag{5.211}$$

The Jacobi–Erderlyi theta functions are the values at half periods,

$$\begin{aligned}&\vartheta_1(z|\tau) = \vartheta[^{\frac{1}{2}}_{\frac{1}{2}}](z|\tau), \quad \vartheta_2(z|\tau) = \vartheta[^{\frac{1}{2}}_{0}](z|\tau), \quad \vartheta_3(z|\tau) = \vartheta[^0_0](z|\tau),\\ &\vartheta_4(z|\tau) = \vartheta[^{0}_{\frac{1}{2}}](z|\tau)\end{aligned} \tag{5.212}$$

In particular,

$$\vartheta_1(v/\tau, -1/\tau) = i\sqrt{-i\tau}e^{i\pi v^2/\tau}\vartheta_1(v,\tau) \tag{5.213}$$

The Dedekind η function is defined as

$$\eta(\tau) = q^{\frac{1}{24}}\prod_{n=1}^{\infty}(1-q^n). \tag{5.214}$$

It satisfies the modular property

$$\eta\left(-\frac{1}{\tau}\right) = \sqrt{-i\tau}\,\eta(\tau) \tag{5.215}$$

It is related to the Jacobi–Erderlyi theta functions by the identities

$$\frac{\partial}{\partial v}\vartheta_1(v)|_{v=0} = 2\pi\ \eta^3(\tau) \tag{5.216}$$

$$\vartheta_2(0|\tau)\vartheta_3(0|\tau)\vartheta_4(0|\tau) = 2\eta^3 \tag{5.217}$$

The partition function of a single left-moving boson is given by

$$Z_{boson}(\tau) := \mathrm{Tr}(q^{L_0}) = \frac{1}{\eta(\tau)}. \tag{5.218}$$

5.8.2.4 Siegel Modular Forms

Let $Sp(2,\mathbb{Z})$ be the group of (4×4) matrices g with integer entries satisfying $gJg^t = J$ where

$$J \equiv \begin{pmatrix} 0 & -I_2 \\ I_2 & 0 \end{pmatrix} \tag{5.219}$$

is the symplectic form. We can write the element g in block form as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{5.220}$$

where A, B, C, D are all (2×2) matrices with integer entries. Then the condition $gJg^t = J$ implies

$$AB^t = BA^t, \quad CD^t = DC^t, \quad AD^t - BC^t = \mathbf{1}, \tag{5.221}$$

Let $\mathbb{H}_2$ be the (genus two) Siegel upper half plane, defined as the set of (2×2) symmetric matrix Ω with complex entries

$$\Omega = \begin{pmatrix} \tau & z \\ z & \sigma \end{pmatrix} \tag{5.222}$$

satisfying

$$\mathrm{Im}(\tau) > 0, \quad \mathrm{Im}(\sigma) > 0, \quad \det(\mathrm{Im}(\Omega)) > 0. \tag{5.223}$$

An element $g \in Sp(2, \mathbb{Z})$ of the form (5.220) has a natural action on $\mathbb{H}_2$ under which it is stable:

$$\Omega \to (A\Omega + B)(C\Omega + D)^{-1}. \tag{5.224}$$

The matrix Ω can be thought of as the period matrix of a genus two Riemann surface[8] on which there is a natural symplectic action of $Sp(2, \mathbb{Z})$.

A Siegel form $F(\Omega)$ of weight k is a holomorphic function $\mathbb{H}_2 \to \mathbb{C}$ satisfying

$$F[(A\Omega + B)(C\Omega + D)^{-1}] = \{\det(C\Omega + D)\}^k F(\Omega). \tag{5.225}$$

A Siegel modular form can be written in terms of its Fourier series

$$F(\Omega) = \sum a(n, r, m) q^n y^r p^m. \tag{5.226}$$

The Siegel modular form which makes its appearance in the present physics problem of counting $\mathcal{N} = 4$ dyons is the Igusa form Φ_{10} which is the unique (cusp) form[9] of weight 10. This Siegel modular form is a very interesting mathematical object and has a number of useful properties directly relevant for the present physical application. In particular, it can be constructed very explicitly in two different ways in terms of familiar modular forms and theta functions by using two different 'lifts.' These constructions are called lifts because they allow us to construct the Igusa cusp form which is a function of three variables using the Fourier expansions of a weak Jacobi forms which are functions of only two variables.

---

8 See [35, 59, 60] for a discussion of the connection with genus-two Riemann surfaces.

9 It is a 'cusp' form because it vanishes at 'cusps' which correspond to $z=0$ and its images.

- *Additive lift*
  Consier the function $\psi(\tau, z)$

$$\psi(\tau, z) = \eta^{18}(\tau)\vartheta_1^2(\tau, z) \;, \tag{5.227}$$

which is a weak Jacobi form of weight 1 and index 10 (see Sect. 5.8.2.2 for definitions). It admits a Fourier expansion

$$\psi(\tau, z) = \sum_{n,r} c_{10}(n, r) q^n y^r \quad q := e^{2\pi i\tau} y := e^{2\pi iz}. \tag{5.228}$$

From the properties of weak Jacobi forms, it follows that the Fourier coefficients $c_{10}(n, r)$ depend only on the combination $4n - r^2$ and hence we can write $c_{10}(n, r) = C_{10}(4n - r^2)$ for some function C_{10}. The additive lift then gives the Fourier expansion of the Igusa cusp form in terms of the Fourier coefficients of $\psi(\tau, z)$ as

$$\Phi_{10}(\Omega) = \sum_{n,m,l} a(m, n, l) p^m q^n y^l, \quad p := e^{2\pi i\sigma} \;, \tag{5.229}$$

where $a(m, n, l)$ are defined by

$$a(n, r, m) = \sum_{\substack{d|(n,r,m)\\ d\geq 1}} d^{k-1} C_{10}\left(\frac{4mn - r^2}{d^2}\right) \;. \tag{5.230}$$

This lift is ‘additive’ in that it gives a sum representation of the Igusa form.

- *Multiplicative lift*
  Consider the function $\chi(\tau, z)$

$$\chi(\tau, z) = 8\left(\frac{\vartheta_2(\tau, z)^2}{\vartheta_2(\tau)^2} + \frac{\vartheta_3(\tau, z)^2}{\vartheta_3(\tau)^2} + \frac{\vartheta_4(\tau, z)^2}{\vartheta_4(\tau)^2}\right), \tag{5.231}$$

which is a weak Jacobi form of weight 0 and index 1 with a Fourier expansion

$$\chi(\tau, z) = \sum_{n,r} c_0(n, l) q^n y^l \quad q := e^{2\pi i\tau}, \quad y := e^{2\pi iz}. \tag{5.232}$$

This function arises in physics applications as the elliptic genus of the $K3$ surface (see Sect. 5.8.3 for details). Once again, $c_0(n, l)$ depend only on the combination $d := 4n - l^2$ and hence we can write

$$c_0(n, l) = C_0(4n - l^2) \tag{5.233}$$

which defines the function $C_0(d)$. The multiplicative lift gives a product representation of the Igusa cusp form in terms of $C_0(d)$:

$$\Phi_{10}(\Omega) = pqy \prod_{(s,t,r)>0} (1 - p^s q^t y^r)^{C_0(4st-r^2)}, \tag{5.234}$$

in terms of C_0 given by (5.231, 5.232). Here the notation $(s, t, r) > 0$ means that either $s > 0$, $t, r \in \mathbb{Z}$, or $s = 0$, $t > 0$, $r \in \mathbb{Z}$, or $s = t = 0$, $r < 0$.
This lift is 'multiplicative' in that it gives a product representation of the Igusa form.

5.8.3 A Few Facts About K3

5.8.3.1 K3 as an Orbifold

"Kummer's third surface" or K3 has played an important role in many developments concerning duality. Let us recall some of its properties. $K3$ is a four dimensional manifold which has $SU(2)$ holonomy. To understand what this means, consider a generic 4d real manifold. If you take a vector in the tangent space at point P, parallel transport it, and come back to point P, then, in general, it will be rotated by an $SO(4)$ matrix:

$$V_i(P) \to O_{ij}\ V_i(P) \quad O_{ij} \in SO(4). \tag{5.235}$$

Such a manifold is then said to to have $SO(4)$ holonomy. In the case of K3, the holonomy is a subgroup of $SO(4)$, namely $SU(2)$. The smaller the holonomy group, the more "symmetric" the space. For example, for a torus, the holonomy group consists of just the identity because the space is flat and Riemann curvature is zero; so, upon parallel transport along a closed loop, a vector comes back to itself. For a K3, there *is* nonzero curvature but it is not completely arbitrary: the Riemann tensor is non-vanishing but the Ricci tensor R_{ij} vanishes. Therefore, K3 can alternatively be defined as the manifold of compactification that solves the vacuum Einstein equations.

Only other thing about K3 that we need to know is the topological information. A surface can have nontrivial cycles which cannot be shrunk to a point. For example, a torus has two nontrivial 1-cycles. The number of nontrivial k-cycles which cannot be smoothly deformed into each other is given by the kth Betti number b_k of the surface. The number of non-trivial k-cycles is in one to one correspondence with the number of harmonic k-forms on the surface given by the kth de-Rham cohomology [5, 6]. A harmonic k-form F_k satisfies the Laplace equation, or equivalently satisfies the equations

$$d^* F_k = 0, \quad dF_k = 0 \tag{5.236}$$

A manifold always has a harmonic 0-form, *viz.*, a constant, and a harmonic 4-form, *viz.*, the volume from, assuming we can integrate on it. K3 has no harmonic 1-forms or 3-forms, but has 22 harmonic 2-forms. So, the Betti numbers for K3 are:

$$b_0 = 1, \quad b_1 = 0, \quad b_2 = 22, \quad b_3 = 0, \quad b_4 = 1. \tag{5.237}$$

Out of the 22 2-forms, 19 are anti-self-dual, and 3 are self-dual. In other words,

$$b_2^s = 3, \quad b_2^a = 19. \tag{5.238}$$

This is all the information one needs to compute the massless spectrum of compactifications on K3.

K3 has a simple description as a $\mathbf{Z}_2$ orbifold of a 4-torus. Let (x_1, x_2, x_3, x_4) be the real coordinates of the torus $\mathbf{T}^4$. Let us further take the torus to be a product $\mathbf{T}^4 = \mathbf{T}^2 \times \mathbf{T}^2$. Let us introduce complex coordinates (z_1, z_2), $z_1 = x_1 + ix_2$ and $z_2 = x_3 + ix_4$. The 2-torus with coordinate z_1 is defined by the identifications $z_1 \sim z_1 + 1 \sim z_1 + i$, and similarly for the other torus. The tangent space group is $Spin(4) \equiv SU(2)_1 \times SU(2)_2$, and the vector representation is $\mathbf{4v} \equiv (\mathbf{2}, \mathbf{2})$. If we take a subgroup $SU(2)_1 \times U(1)$ of $Spin(4)$, then the vector decomposes as

$$\mathbf{4v} = \mathbf{2}_+ \oplus \bar{\mathbf{2}}_-. \tag{5.239}$$

The coordinates (z_1, z_2) transform as the doublet $\mathbf{2}_+$ and $(\bar{z}_1, \bar{z}_2)$ as the $\bar{\mathbf{2}}_-$. The $\mathbf{Z}_2 = \{1, I\}$ is generated by

$$I : (z_1, z_2) \to (-z_1, -z_2). \tag{5.240}$$

This $\mathbf{Z}_2$ is a subgroup and in fact the center of $SU(2)_1$. Consequently, as we shall see, the resulting manifold has $SU(2)$, indeed a $\mathbf{Z}_2$ holonomy. For a torus coordinatized by z_1, there are 4 fixed points of $z_1 \to -z_1$ Altogether, on $\mathbf{T}^4/\mathbf{Z}_2$, there are 16 fixed points.

Let us calculate the number of harmonic forms on this orbifold. To begin with, we have on the torus $\mathbf{T}^4$, the following harmonic forms:

$$\begin{array}{ll} 1 & 1 \\ 4 & dx^i \\ 6 & dx^i \wedge dx^j \\ 4 & dx^i \wedge dx^j \wedge dx^l \\ 1 & dx^i \wedge dx^j \wedge dx^k \wedge dx^l. \end{array} \tag{5.241}$$

The first column gives the number of forms indicated in the second column where the indices *i, j, k, l* take values $1, \cdots 4$. Under the reflection *I*, only the even forms $1, dx^i \wedge dx^j$, and $dx^i \wedge dx^j \wedge dx^k \wedge dx^l$ survive.

$$\begin{array}{lccc} 0-\text{form} & 1 & & 1 \\ 1 & 4 & & 0 \\ 2 & 6 & \xrightarrow{\frac{1+I}{2}} & 6, \\ 3 & 4 & & 0 \\ 4 & 1 & & 1 \end{array} \tag{5.242}$$

where the second column give the number of forms on the torus and the third column the number of forms that survive the projection. Let us look at the 2-forms from the torus that survive the $\mathbf{Z}_2$ projection. By taking the combinations

$$dx^i \wedge dx^j \pm \frac{1}{2}\varepsilon^{ijkl} dx^k \wedge dx^l$$

we see that three of these 2-forms are self-dual and the remaining three are anti-self-dual.

At the fixed point of the orbifold symmetry there is a curvature singularity. The singularity can be repaired as follows. We cut out a ball of radius R around each point, which has a boundary $S^3/\mathbf{Z}_2$, replace it with a noncompact smooth manifold that is also Ricci flat and has a boundary $S^3/\mathbf{Z}_2$, and then take the limit $R \to 0$. The required noncompact Ricci-flat manifold with boundary $S^3/\mathbf{Z}_2$ is known to exist and is called the Eguchi–Hanson space. The Betti number of the Eguchi–Hanson space are $b_0 = b_4 = 1$ and $b_2^a = 1$. Therefore, each fixed point contributes an anti-self-dual 2-form which corresponds to a nontrivial 2-cycle in the Eguchi–Hanson space that would be stuck at the fixed point in the limit $R \to 0$.

Altogether, we get $b_0 = 1$, $b_2^s = 3$, $b_2^a = 3 + 16 = 19$, $b_4 = 1$, and $b_1 = b_3 = 0$ giving us the cohomology of K3. It obviously has $SU(2)$ holonomy. Away from the fixed point, a parallel transported vector goes back to itself, because all the curvature is concentrated at the fixed points. As we go around the fixed point a vector is returned to its reflected image (for instance, $(dz_1, dz_2) \to -(dz_1, dz_2)$), i.e., transformed by an element of $SU(2)$.

In string theory there is no need to repair the singularity by hand. We shall see in Sects. 5.3 and 5.4 that the twisted states in the spectrum of Type-II string moving on an orbifold automatically take care of the repairing. The twisted states somehow know about the Eguchi–Hanson manifold that would be necessary to geometrically repair the singularity.

5.8.3.2 Elliptic Genus of $K3$

Consider a two-dimensional superconformal field theories (SCFT) with (2, 2) or more worldsheet supersymmetry.[10] We denote the superconformal field theory by $\sigma(\mathscr{M})$ when it corresponds to a sigma model with a target manifold $\mathscr{M}$. Let H be the Hamiltonian in the Ramond sector, and J be the left-moving $U(1)$ R-charge. The elliptic genus $\chi(\tau, z; \mathscr{M})$ is then defined [61–63] as a trace over the Hilbert space $\mathscr{H}_R$ in the Ramond sector

$$\chi(\tau, z; \mathscr{M}) = \mathrm{Tr}_{\mathscr{H}_R}\left(q^H y^J (-1)^F\right) \tag{5.243}$$

where F is the fermion number. An elliptic genus so defined satisfies the modular transformation property (5.202) as a consequence of modular invariance of the

[10] An SCFT with (r, s) supersymmetries has r left-moving and s right-moving supersymmetries.

path integral. Similarly, it satisfies the elliptic transformation property (5.203) as a consequence of spectral flow. Furthermore, in a unitary SCFT, the positivity of the Hamiltonian implies that the elliptic genus is a weak Jacobi form.

A particularly useful example in the present context is $\sigma(K3)$, which is a (4, 4) SCFT whose target space is a $K3$ surface. The elliptic genus is a topological invariant and is independent of the moduli of the $K3$. Hence, it can be computed at some convenient point in the $K3$ moduli space, for example, at the orbifold point where the $K3$ is the Kummer surface. At this point, the $\sigma(K3)$ SCFT can be regarded as a $\mathbb{Z}_2$ orbifold of the $\sigma(T^4)$ SCFT which is an SCFT with a torus T^4 as the target space. A simple computation using standard techniques of orbifold conformal field theory yields [64] the formula for the elliptic genus we introduced earlier in (5.231):

$$\chi(\tau, z) = 8\left(\frac{\vartheta_2(\tau, z)^2}{\vartheta_2(\tau)^2} + \frac{\vartheta_3(\tau, z)^2}{\vartheta_3(\tau)^2} + \frac{\vartheta_4(\tau, z)^2}{\vartheta_4(\tau)^2}\right). \tag{5.244}$$

The first term can be seen to arise from the untwisted projected partition function, the second from the twisted, unprojected partition function and the third from the twisted, projected partition function.

Note that for $z = 0$, the trace (5.243) reduces to the Witten index of the SCFT and correspondingly the elliptic genus reduces to the Euler character of the target space manifold. In our case, one can readily verify from (5.231) that $\chi(\tau, 0; K3)$ equals 24 which is the Euler character of $K3$.

References

1. Carroll, S.M.: Spacetime and Geometry: An Introduction to General Relativity, p. 513. Addison-Wesley, San Francisco (2004)
2. Wald, R.M.: General Relativity, pp. 491. University of Chicago Press, Chicago (1984)
3. Misner, C.W., Thorne, K.S., Wheeler, J.A., Gravitation, p. 1279. San Francisco (1973)
4. Birrell, N.D., Davies, P.C.W.: Quantum Fields in Curved Space, p. 340. Cambridge University Press, Cambridge (1982)
5. Green, M.B., Schwarz, J.H., Witten, E.: Superstring Theory, vol. 1: Introduction, p. 469. Cambridge University Press, Cambridge (1987) (Cambridge Monographs On Mathematical Physics)
6. Green, M.B., Schwarz, J.H., Witten, E.: Superstring Theory, vol. 2: Loop Amplitudes, Anomalies and Phenomenology, p. 596. Cambridge University Press, Cambridge (1987) (Cambridge Monographs On Mathematical Physics).
7. Polchinski, J.: String Theory. vol. 1, Cambridge University Press, Cambridge (1998)
8. Polchinski, J.: String Theory. Vol. 2: Superstring Theory and Beyond, p. 531. Cambridge University Press, Cambridge (1998)
9. Sen, A.: Black Hole Entropy Function, Attractors and Precision Counting of Microstates. Gen. Rel. Grav. **40**, 2249–2431 (2008) ([arXiv:0708.1270 [hep-th]])
10. Einstein, A.: PAW, p. 844 (1915)
11. Schwarzschild, K.: PAW, p. 189 (1916)
12. Bekenstein, J.D.: Black holes and entropy. Phys. Rev. **D7**, 2333–2346 (1973)
13. Hawking, S.W.: Particle creation by black holes. Commun. Math. Phys. **43**, 199–220 (1975)

14. Wald, R.M.: Black hole entropy is the noether charge. Phys. Rev. **D48**, 3427–3431 (1993) ([gr-qc/9307038])
15. Iyer, V., Wald, R.M.: Some properties of noether charge and a proposal for dynamical black hole entropy. Phys. Rev. **D50**, 846–864 (1994) ([gr-qc/9403028])
16. Jacobson, T., Kang, G., Myers, R.C.: Black hole entropy in higher curvature gravity. ([gr-qc/9502009])
17. Sen, A.: Entropy function and AdS(2)/CFT(1) correspondence. JHEP **0811**, 075 (2008) ([arXiv:0805.0095 [hep-th]])
18. Sen, A.: Quantum entropy function from AdS(2)/CFT(1) correspondence. ([arXiv:0809.3304 [hep-th]])
19. Hull, C.M., Townsend, P.K.: Unity of superstring dualities. Nucl. Phys. **B438**, 109–137 (1995) ([hep-th/9410167])
20. Witten, E.: String theory dynamics in various dimensions. Nucl. Phys. **B443**, 85–126 (1995) ([hep-th/9503124])
21. Sen, A.: Dyon-monopole bound states, selfdual harmonic forms on the multi-monopole moduli space, and sl(2, z) invariance in string theory. Phys. Lett. **B329**, 217–221 (1994) ([hep-th/9402032])
22. Sen, A.: Strong–weak coupling duality in four-dimensional string theory. Int. J. Mod. Phys. **A9**, 3707–3750 (1994) ([arXiv:hep-th/9402002])
23. Witten, E., Olive, D.I.: Supersymmetry algebras that include topological charges. Phys. Lett. **B78**, 97 (1978)
24. Dabholkar, A., Gaiotto, D., Nampuri, S.: Comments on the spectrum of CHL dyons. JHEP **01**, 023 (2008) ([arXiv:hep-th/0702150])
25. Banerjee, S. Sen, A.: Duality orbits, dyon spectrum and gauge theory limit of heterotic string theory on T^6. ([arXiv:0712.0043 [hep-th]])
26. Banerjee, S., Sen, A.: S-duality action on discrete T-duality invariants. ([arXiv:0801.0149 [hep-th]])
27. Banerjee, S., Sen, A., Srivastava, Y.K.: Partition functions of torsion >1 dyons in heterotic string theory on T^6. ([arXiv:0802.1556 [hep-th]])
28. Dabholkar, A., Gomes, J., Murthy, S.: Counting all dyons in N = 4 string theory. ([arXiv:0803.2692 [hep-th]])
29. Dabholkar, A., Harvey, J.A.: Nonrenormalization of the superstring tension. Phys. Rev. Lett. **63**, 478 (1989)
30. Dabholkar, A., Gibbons, G.W., Harvey, J.A., Ruiz Ruiz, F.: Superstrings and solitons. Nucl. Phys. **B340**, 33–55 (1990)
31. Dijkgraaf, R., Moore, G.W., Verlinde, E.P., Verlinde, H.L.: Elliptic genera of symmetric products and second quantized strings. Commun. Math. Phys. **185**, 197–209 (1997) ([hep-th/9608096])
32. Gaiotto, D., Strominger, A., Yin, X.: 5D black rings and 4D black holes. JHEP **02**, 023 (2006) ([arXiv:hep-th/0504126])
33. Strominger, A., Vafa, C.: Microscopic origin of the Hekenstein–Hawking entropy. Phys. Lett. **B379**, 99–104 (1996) ([hep-th/9601029])
34. Breckenridge, J.C., Myers, R.C., Peet, A.W., Vafa, C.: D-branes and spinning black holes. Phys. Lett. **B391**, 93–98 (1997) ([hep-th/9602065])
35. Gaiotto, D.: Re-recounting dyons in N = 4 string theory. ([hep-th/0506249])
36. David, J.R., Sen, A.: CHL dyons and statistical entropy function from D1–D5 system. JHEP **0611** , 072 (2006) ([hep-th/0605210])
37. Cheng, M.C.N., Verlinde, E.: Dying dyons don't count. ([arXiv:0706.2363 [hep-th]])
38. Sen, A.: Walls of marginal stability and dyon spectrum in N = 4 supersymmetric string theories. JHEP **05**, 039 (2007) ([hep-th/0702141])
39. Lopes Cardoso, G., de Wit, B., Kappeli, J., Mohaupt, T.: Asymptotic degeneracy of dyonic N = 4 string states and black hole entropy. JHEP **12**, 075 (2004) ([hep-th/0412287])

40. Sen, A.: Black hole solutions in heterotic string theory on a torus. Nucl. Phys. **B440**, 421–440 (1995) ([hep-th/9411187])
41. Cvetic, M., Youm, D.: Dyonic bps saturated black holes of heterotic string on a six torus. Phys. Rev. **D53**, 584–588 (1996) ([hep-th/9507090])
42. Ferrara, S., Kallosh, R., Strominger, A.: N = 2 extremal black holes. Phys. Rev. **D52**, 5412–5416 (1995) ([hep-th/9508072])
43. Ferrara, S., Kallosh, R.: Supersymmetry and attractors. Phys. Rev. **D54**, 1514–1524 (1996) ([hep-th/9602136])
44. Strominger, A.: Macroscopic entropy of $n=2$ extremal black holes. Phys. Lett. **B383** 39–43 (1996) ([hep-th/9602111])
45. Lopes Cardoso, G., de Wit, B., Mohaupt, T.: Corrections to macroscopic supersymmetric black-hole entropy. Phys. Lett. **B451**, 309–316 (1999) ([hep-th/9812082])
46. Lopes Cardoso, G., de Wit, B., Mohaupt, T.: Deviations from the area law for supersymmetric black holes. Fortsch. Phys. **48**, 49–64 (2000) ([hep-th/9904005])
47. Lopes Cardoso, G., de Wit, B., Mohaupt, T.: Area law corrections from state counting and supergravity. Class. Quant. Grav. **17**, 1007–1015 (2000) ([hep-th/9910179])
48. Lopes Cardoso, G., de Wit, B., Kappeli, J., Mohaupt, T.: Stationary bps solutions in n = 2 supergravity with r**2 interactions. JHEP **12**, 019 (2000) ([hep-th/0009234])
49. Sen, A.: Entropy function for heterotic black holes. ([hep-th/0508042])
50. Dabholkar, A., Denef, F., Moore, G.W. Pioline, B.: Exact and asymptotic degeneracies of small black holes. JHEP **08**, 021 (2005) ([hep-th/0502157])
51. Dabholkar, A., Denef, F., Moore, G.W., Pioline, B.: Precision counting of small black holes. JHEP **10**, 096 (2005) ([arXiv:hep-th/0507014])
52. Dabholkar, A.: Exact counting of black hole microstates. Phys. Rev. Lett. **94**, 241301 (2005) ([hep-th/0409148])
53. Dabholkar, A., Kallosh, R., Maloney, A.: A stringy cloak for a classical singularity. JHEP **12**, 059 (2004) ([hep-th/0410076])
54. Sen, A.: Extremal black holes and elementary string states Mod. Phys. Lett. **A10**, 2081–2094 (1995) ([hep-th/9504147])
55. Kraus, P., Larsen, F.: Microscopic black hole entropy in theories with higher derivatives. JHEP **09**, 034 (2005) ([hep-th/0506176])
56. Kraus, P., Larsen, F.: Holographic gravitational anomalies. JHEP **01**, 022 (2006) ([hep-th/0508218])
57. de Wit, B., Katmadas, S., van Zalk, M.: New supersymmetric higher-derivative couplings: Full N = 2 superspace does not count! ([arXiv:1010.2150 [hep-th]])
58. Kiritsis, E.: Introduction to non-perturbative string theory. ([hep-th/9708130])
59. Dabholkar, A., Nampuri, S.: Spectrum of dyons and black holes in CHL orbifolds using borcherds lift. JHEP **11**, 077 (2007) ([arXiv:hep-th/0603066])
60. Banerjee, S., Sen, A., Srivastava, Y.K.: Genus two surface and quarter BPS dyons: the contour prescription. JHEP **03**, 151 (2009) ([arXiv:0808.1746 [hep-th]])
61. Witten, E.: Elliptic genera and quantum field theory. Commun. Math. Phys. **109**, 525 (1987)
62. Alvarez, O., Killingback, T.P., Mangano, M.L., Windey, P.: String theory and loop space index theorems. Commun. Math. Phys. **111**, 1 (1987)
63. Ochanine, S.: Sur les genres multiplicatifs definis par des integrales elliptiques. Topology **26**, 143 (1987)
64. Ginsparg, P.H.: Applied conformal field theory. ([hep-th/9108028])

Chapter 6
Lectures on Topological String Theory

Hirosi Ooguri

6.1 Topological Sigma-Models

In string theory, one typically studies embeddings of the string worldsheet Σ (which is a Riemann surface) into a 10-dimensional manifold of the form $X : \Sigma \to \mathbb{R}^{1,3} \times M$, with M a compact 6-dimensional manifold. In the following, we will restrict our attention to the study of supersymmetric sigma models onto the 6 compact dimensions. In particular, we will assume M to be a Calabi–Yau threefold so that the theory possesses $\mathcal{N} = (2, 2)$ worldsheet superconformal symmetry.

Let us start by reviewing the $\mathcal{N} = 2$ superconformal algebra (SCA). This is generated by the following fields:

$$T(z) = \sum_{n\in\mathbb{Z}} L_n z^{-n-2}, \quad G^{\pm}(z) = \sum_{n\in\mathbb{Z}} G^{\pm}_{n+a} z^{-(n\pm a)-3/2}, \quad J(z) = \sum_{n\in\mathbb{Z}} J_n z^{-n-1}. \tag{6.1}$$

with the following set of OPE's:

$$T(z)T(w) \sim \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial_w T(w)}{z-w} \tag{6.2}$$

$$T(z)G^{\pm}(w) \sim \frac{3/2}{(z-w)^2} G^{\pm}(w) + \frac{\partial_w G^{\pm}(w)}{z-w} \tag{6.3}$$

H. Ooguri (✉)
California Institute of Technology,
452-48, Pasadena, CA 91125, USA
e-mail: ooguri@theory.caltech.edu

Institute for the Physics and Mathematics of the Universe,
Todai Institutes for Advanced Study, University of Tokyo,
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_6,

$$T(z)J(w) \sim \frac{J(w)}{(z-w)^2} + \frac{\partial_w J(w)}{z-w} \tag{6.4}$$

$$G^+(z)G^-(w) \sim \frac{2c/3}{(z-w)^3} + \frac{2J(w)}{(z-w)^2} + \frac{2T(w) + \partial_w J(w)}{z-w} \tag{6.5}$$

$$J(z)G^\pm(w) \sim \pm\frac{G^\pm(w)}{z-w} \tag{6.6}$$

$$J(z)J(w) \sim \frac{c/3}{(z-w)^2} \tag{6.7}$$

which reveal that J and G are primary fields with conformal weights 1 and $3/2$, respectively, and that $G^\pm$ has $U(1)$ charge ± 1 under J_0. In the following, we will restrict ourselves to the Ramond sector (i.e. $a = 0$ in (6.1)).

The superconformal algebra possesses two distinguished vector space isomorphisms:

$$T(z) \to T(z) + \frac{1}{2}\partial J(z), \quad J(z) \to J(z), \quad G^\pm(z) \to G^\pm(z), \tag{6.8}$$

and

$$T(z) \to T(z) - \frac{1}{2}\partial J(z), \quad J(z) \to -J(z), \quad G^\pm(z) \to G^\pm(z). \tag{6.9}$$

The above transformations are known as *topological twists*.

We shall now analyse the twist (6.8). The OPE's in the new basis are:

$$T(z)T(w) \sim \frac{2T(w)}{(z-w)^2} + \frac{\partial_w T(w)}{z-w} \tag{6.10}$$

$$T(z)G^-(w) \sim \frac{2G^-(w)}{(z-w)^2} + \frac{\partial_w G^-(w)}{z-w} \tag{6.11}$$

$$T(z)G^+(w) \sim \frac{G^+(w)}{(z-w)^2} + \frac{\partial_w G^+(w)}{z-w} \tag{6.12}$$

$$T(z)J(w) \sim -\frac{\hat{c}}{(z-w)^3} + \frac{J(w)}{(z-w)^2} + \frac{\partial_w J(w)}{z-w} \tag{6.13}$$

$$G^+(z)G^-(w) \sim \frac{2\hat{c}}{(z-w)^3} + \frac{2J(w)}{(z-w)^2} + \frac{2T(w)}{z-w} \tag{6.14}$$

$$J(z)G^\pm(w) \sim \pm\frac{G^\pm(w)}{z-w} \tag{6.15}$$

$$J(z)J(w) \sim \frac{\hat{c}}{(z-w)^2} \tag{6.16}$$

where we introduced $\hat{c} = c/3$. The weights of the supercurrents G^+ and G^- with respect to T undergo the following transformations:

$$G^+ : 3/2 \to 1, \quad G^- : 3/2 \to 2. \tag{6.17}$$

Therefore, the charge Q corresponding to G^+:

$$Q_\gamma = \int_\gamma G^+ \tag{6.18}$$

has weight 0. Here γ is one of the $2g$ generators of $H_1(\Sigma_g, \mathbb{Z})$. It is important to notice that the commutator of Q with a local field on Σ is independent of the choice of γ. The supercharge Q is to be thought of as a BRST operator. It is indeed nilpotent, i.e., $Q^2 = 0$. Moreover the modified $T(z)$ satisfies the following identity:

$$T(z) = \frac{1}{2}\{Q, G^-(z)\}. \tag{6.19}$$

Now we can give meaning to the topological nature of the twists (6.8), (6.9), by interpreting the modified $T(z)$ as the energy momentum tensor of a 2-dimensional $\mathcal{N} = 2$ superconformal field theory. Then (6.19) is the statement that the energy momentum tensor is BRST trivial. Topological invariance of the theory, i.e., the independence of its correlation functions on the worldsheet metric η:

$$\frac{1}{\sqrt{\eta}}\frac{\delta}{\delta\eta}\langle\mathcal{O}\rangle = \langle T\ \mathcal{O}\rangle = \langle\{Q, G^-\}\mathcal{O}\rangle = 0 \tag{6.20}$$

is then achieved by restricting the space of observables $\mathcal{O}$ to $[\cdot, Q]$-cohomology H_Q, the space of chiral primary fields. These in turn correspond to the highest weights of positive energy representations of the $\mathcal{N} = 2$ SCA.

In the process of redefining the energy momentum tensor, it is crucial to understand that the action will undergo a non trivial transformation. Let S_0 denote the action of the superconformal field theory with energy momentum tensor T prior to the twist, then S is related to S_0 via:

$$S = S_0 + \int_\Sigma \bar{A} J, \tag{6.21}$$

where $A = (i/2)\omega$ and ω is the spin connection on Σ. It is left as an exercise to the reader to verify that this transformation is tantamount to the topological twist (6.8).

6.1.1 The Non-linear Sigma Model

We will now specialize our discussion to the class of $\mathcal{N} = (2, 2)$ superconformal field theories with action:

$$S = \int_{\Sigma^{(2|2)}} d^4\theta d^2 z\ K(\Phi^i, \Phi^{\bar{i}}), \tag{6.22}$$

where $\Sigma^{(2|2)}$ is the supermanifold with bosonic part a closed Riemann surface Σ and real spinorial representation $S \oplus \overline{S}$ with S the irreducible spinorial representation of $U(1)$. The function K is the Kähler potential on the target manifold M, i.e., it locally defines a Kähler metric on M by $g_{i\bar{j}} := \partial_i \partial_{\bar{j}} K$. The Φ^i's are chiral superfields of the form

$$\Phi^i(y^\pm, \theta^\pm) = X^i(y^\pm) + \psi^i(y^\pm)\theta^+ + \overline{\psi}^i(y^\pm)\theta^- + F^i(y^\pm)\theta^+\theta^-, \tag{6.23}$$

where $y^+ = z - i\theta^+\bar{\theta}^+$ and $y^- = \bar{z} - i\theta^-\bar{\theta}^-$, and

$$X \in \Gamma(\Sigma, M) \tag{6.24}$$

$$\psi \in \Gamma(\Sigma, X^* TM^{(1,0)} \otimes S) \tag{6.25}$$

$$\chi \in \Gamma(\Sigma, X^* TM^{(0,1)} \otimes S) \tag{6.26}$$

$$F \in \Gamma(\Sigma, X^* TM^{(1,0)} \otimes S \otimes \overline{S}) \tag{6.27}$$

$$(\psi \leftrightarrow \overline{\psi}, \chi \leftrightarrow \overline{\chi}, S \leftrightarrow \overline{S}). \tag{6.28}$$

In terms of the component fields, S assumes the following form:

$$S = \int_\Sigma d^2 z \Big[\sqrt{\eta} g_{i\bar{j}} \eta^{\mu\nu} \partial_\mu X^i \partial_\nu X^{\bar{j}} - i g_{i\bar{j}} \overline{\chi}^{\bar{j}} D_z \overline{\psi}^i + i g_{i\bar{j}} \chi^{\bar{j}} D_{\bar{z}} \psi^i \\ -\frac{1}{2} R_{i\bar{j}k\bar{l}} \psi^i \overline{\psi}^k \overline{\chi}^{\bar{j}} \chi^{\bar{l}} + g_{i\bar{j}} (F^i - \Gamma^i_{jk} \psi^j \overline{\psi}^k)(F^{\bar{j}} - \Gamma^{\bar{j}}_{\bar{k}\bar{l}} \overline{\chi}^{\bar{k}} \chi^{\bar{l}})\Big]. \tag{6.29}$$

The fields F^i and $F^{\bar{i}}$ are non-dynamical (auxiliary) and the last term in (6.29) can be eliminated yielding a classically equivalent theory. Here D_z is covariant both with respect to Σ and to the target M, i.e., $D_z = \nabla_z + \iota_{\partial_z X}\Gamma$, where ∇_z is the connection on the spin bundle $X^* TM^{(1,0)} \otimes S \to \Sigma$.

The classical theory possesses a superconformal symmetry with Noether currents:

$$\mathscr{J} = g_{i\bar{j}} (\psi^i \chi^{\bar{j}} + \overline{\psi}^i \overline{\chi}^{\bar{j}}) \tag{6.30}$$

$$\mathscr{G}^+ = g_{i\bar{j}} (\psi^i \partial_z X^{\bar{j}} + \overline{\psi}^i \partial_{\bar{z}} X^{\bar{j}}) \tag{6.31}$$

$$\mathscr{G}^- = g_{i\bar{j}} (\chi^{\bar{j}} \partial_z X^i + \overline{\chi}^{\bar{j}} \partial_{\bar{z}} X^i) \tag{6.32}$$

$$\mathscr{T} = g_{i\bar{j}} (\partial_z X^i \partial_{\bar{z}} X^{\bar{j}} + i \chi^{\bar{j}} \partial_z \psi^i + i \overline{\chi}^{\bar{j}} \partial_{\bar{z}} \overline{\psi}^i) \tag{6.33}$$

which are clearly only conserved along the physical trajectories where they decompose as $\mathscr{T} = T + \overline{T}$ and similarly for $\mathscr{G}^\pm$ and $\mathscr{J}$, i.e., as a sum of generators of the left and right classical superconformal algebra, respectively. For curved Riemann surfaces Σ, $\mathscr{G}^+$ and $\mathscr{G}^-$ are not conserved, as their derivation hinges on the existence of a covariantly constant spinor, which in this case is absent. After twisting the theory, two of the spinors become scalars, thus restoring half of the supersymmetry.

Upon quantizing the theory, the axial $U(1)_A$ together with the conformal symmetry will generically become anomalous. In the following, we will provide a detailed analysis of the $U(1)_A$ anomaly. For a thorough discussion of the conformal anomaly we refer to [1].

6.1.1.1 The Axial Anomaly

Classically, $U(1)_A$ is generated by $J(z) - \overline{J}(\bar{z})$, in particular, the global $U(1)_A$ transformations have the form:

$$\psi^i \to e^{-i\alpha}\psi^i, \quad (\psi^i \leftrightarrow \chi^{\bar{i}}, \alpha \leftrightarrow -\alpha), \tag{6.34}$$

$$\overline{\psi}^i \to e^{+i\alpha}\overline{\psi}^i, \quad (\overline{\psi}^i \leftrightarrow \overline{\chi}^{\bar{i}}, \alpha \leftrightarrow -\alpha). \tag{6.35}$$

Let us analyze the $U(1)_A$ transformation of the path integral measure:

$$\mathscr{D}X\mathscr{D}\psi\mathscr{D}\chi\mathscr{D}\overline{\psi}\mathscr{D}\overline{\chi}e^{iS} . \tag{6.36}$$

Given that the action S and X are invariant under $U(1)_A$, the only potentially anomalous term is

$$\mathscr{D}\psi\mathscr{D}\chi\mathscr{D}\overline{\psi}\mathscr{D}\overline{\chi} . \tag{6.37}$$

In order to see the appearance of the anomaly explicitly, let us express the ψ fields in terms of the eigenfunctions of the hermitian operators $D_{\bar{z}}^\dagger D_{\bar{z}}$ and $D_z^\dagger D_z$. The Hilbert space in question is the completion of $\Gamma(\Sigma, (X^*TM^{1,0} \oplus X^*TM^{0,1}) \otimes (S \oplus \bar{S}))$ with scalar product

$$(\Psi_1, \Psi_2) = \int_\Sigma d^2z\, g(\overline{\Psi}_1, \Psi_2) = \int_\Sigma d^2z\, g_{i\bar{j}}(\chi_1^{\bar{j}}\overline{\psi}_2^i + \overline{\chi}_1^{\bar{j}}\psi_2^i) , \tag{6.38}$$

therefore $D_z^\dagger = -D_{\bar{z}}$. The free part of the action for the Ψ fields is then written as

$$S[\Psi]_{free} = i(\Psi, \not{D}\Psi). \tag{6.39}$$

Therefore, D_z and $D_{\bar{z}}$ have the following domains and ranges:

$$D_{\bar{z}} : \Gamma(\Sigma, E^{(1,0)} \oplus \overline{E}^{(0,1)}) \to \Gamma(\Sigma, \overline{E}^{(1,0)} \oplus E^{0,1}), \tag{6.40}$$

$$D_z : \Gamma(\Sigma, \overline{E}^{(1,0)} \oplus E^{0,1}) \to \Gamma(\Sigma, E^{(1,0)} \oplus \overline{E}^{(0,1)}), \tag{6.41}$$

where we have denoted $X^*TM^{(1,0)} \otimes \bar{S}$ as $\overline{E}^{(1,0)}$, and similarly for the other subbundles. The fields $\overline{\psi}, \chi \in \Gamma(\Sigma, \overline{E}^{(1,0)} \oplus E^{0,1})$ are decomposed in terms of eigenfunctions of $-D_{\bar{z}} D_z$ as:

$$\overline{\psi} = \sum_{l=1}^{k} c_{0l} \overline{\phi}^{0l} + \sum_{n=1}^{\infty} c_n \overline{\phi}^n, \tag{6.42}$$

$$\chi = \sum_{l=1}^{k} \tilde{c}_{0l} \xi^{0l} + \sum_{n=1}^{\infty} \tilde{c}_n \xi^n. \tag{6.43}$$

On the other hand, $\psi, \overline{\chi} \in \Gamma(\Sigma, E^{(1,0)} \oplus \overline{E}^{(0,1)})$ have the following spectral decomposition with respect to $-D_z D_{\bar{z}}$:

$$\psi = \sum_{n=1}^{\infty} \tilde{b}_n \phi^n, \tag{6.44}$$

$$\overline{\chi} = \sum_{n=1}^{\infty} b_n \overline{\xi}^n. \tag{6.45}$$

Without loss of generality we take

$$k := \dim \mathrm{Ker} D_{\bar{z}} D_z - \dim \mathrm{Ker} D_z D_{\bar{z}} \geq 0. \tag{6.46}$$

Due to the general properties of Dirac operators, the left and right moving modes of $D_{\bar{z}} D_z$ and $D_z D_{\bar{z}}$, respectively, have the same eigenvalues. Moreover, the eigenmodes of $D_{\bar{z}} D_z$ and $D_z D_{\bar{z}}$ with non-zero eigenvalues are paired up: let ψ be an eigenfunction of $D_{\bar{z}} D_z$ with non-zero eigenvalue, then $D_z \psi \neq 0$ and it is an eigenfunction of $D_z D_{\bar{z}}$ to the same eigenvalue. It follows that

$$\mathscr{D}\Lambda := \mathscr{D}\psi \mathscr{D}\chi \mathscr{D}\overline{\psi} \mathscr{D}\overline{\chi} = \prod_{l=1}^{k} dc_{0l} d\tilde{c}_{0l} \prod_{n=1}^{\infty} db_n dc_n d\tilde{b}_n d\tilde{c}_n \ . \tag{6.47}$$

In particular, an infinitesimal $U(1)_A$ transformation yields

$$\frac{d}{d\alpha} \mathscr{D}\Lambda \Big|_{\alpha=0} = 2k \ \mathscr{D}\Lambda. \tag{6.48}$$

Thus, we observe that unless $k = 0$, the $U(1)_A$ symmetry of the classical theory is broken upon quantization. In order to give a geometric meaning to this statement, we first notice that

$$k \equiv \dim \mathrm{Ker} D_{\bar{z}} D_z - \dim \mathrm{Ker} D_z D_{\bar{z}} = \dim \mathrm{Ker} D_z - \dim \mathrm{Ker} D_{\bar{z}} =: \mathrm{Ind} D_z. \tag{6.49}$$

This follows by observing that $D_z^\dagger \psi \in \mathrm{Ker} D \iff D_z^\dagger \psi = 0$, therefore $\dim \mathrm{Ker} D_{\bar{z}} D_z = \dim \mathrm{Ker} D_z$. By the Atyiah-Singer Index theorem it then follows

$$k = \int_\Sigma c_1(X^* TM \otimes (S \oplus \bar{S})). \tag{6.50}$$

This simplifies further by the property of the Chern character $Ch(E_1 \otimes E_2) = Ch(E_1)Ch(E_2)$, to

$$k = \mathrm{rank}(X^* TM) \int_\Sigma c_1(S \oplus \bar{S}) + \mathrm{rank}(S \oplus \bar{S}) \int_\Sigma c_1(X^* TM) \tag{6.51}$$

$$= \dim(M)(2g-2) + 4 \int_{X(\Sigma)} c_1(TM), \tag{6.52}$$

where g is the genus of Σ and in the last step we used Gauss–Bonnet and the naturality property of characteristic classes. We require that for a given target M, the sigma model should preseve $N = (2, 2)$ supersymmetry on a genus $g = 1$ Riemann surface Σ. We hence need

$$\int_{X(\Sigma)} c_1(TM) = 0 \quad \forall X \in \Gamma(\Sigma, M) \tag{6.53}$$

$$\implies c_1(TM) = 0. \tag{6.54}$$

That is, we require M to be a Calabi–Yau manifold.

6.1.1.2 Calabi–Yau Manifolds

Let us briefly recall the notion of a Calabi–Yau manifold. A Calabi–Yau manifold M is defined as a Ricci flat Kähler manifold. The Ricci tensor of a Kähler manifold with Kähler metric g can be expressed in the form

$$R_{i\bar{j}} = \partial_i \partial_{\bar{j}} \log \det g. \tag{6.55}$$

The solution to $R_{i\bar{j}} = 0$, requiring the metric to be real, is

$$\det g = e^{f(x)} e^{\bar{f}(\bar{x})}, \tag{6.56}$$

where f is a holomorphic function on M. It follows that $e^{f(x)} dx^1 \wedge \ldots \wedge dx^n$ is a nowhere vanishing holomorphic top-form, which is ensured by the non-degeneracy of the metric. We have hence showed (M, g) Calabi–Yau $\implies h^{n,0}(M) = 1$, in particular, $c_1(TM) = 0$. Yau's Theorem asserts that also the converse to this statement is true.

Theorem *Let (M, J) be a compact complex manifold admitting Kähler metrics, with $c_1(TM) = 0$. Then for each Kähler class $[\omega] \in H^{1,1}(M, \mathbb{R})$, there is a unique representative ω whose associated Kähler metric is Ricci-flat.*

6.1.2 The A-Twist

Consider the non-linear sigma model on a flat Riemann surface with a Calabi–Yau target M, then the model is $\mathcal{N} = (2,2)$ superconformally invariant. The A-twist consists in applying (6.8) to both the left- and right-moving sectors of the $\mathcal{N} = (2,2)$ sigma model. Restricting to the left-moving part, the fermionic fields undergo the following transformation of conformal weights:

$$\psi^i : (1/2, 0) \to (0, 0), \tag{6.57}$$

$$\chi^{\bar{i}} : (1/2, 0) \to (1, 0). \tag{6.58}$$

Therefore we see that the fermions are not sections of the same bundles as in the original sigma model, but rather

$$\psi \in \Gamma(\Sigma, X^* TM^{1,0}), \tag{6.59}$$

$$\chi \in \Gamma(\Sigma, X^* TM^{0,1} \otimes \Omega^{1,0}(\Sigma)) \ , \tag{6.60}$$

and similarly for the right-moving fields:

$$\overline{\psi} \in \Gamma(\Sigma, X^* TM^{1,0} \otimes \Omega^{0,1}(\Sigma)), \tag{6.61}$$

$$\overline{\chi} \in \Gamma(\Sigma, X^* TM^{0,1}). \tag{6.62}$$

The action has precisely the same form as (6.29) except that D_z and $D_{\bar{z}}$ are now covariant derivatives on $\Gamma(\Sigma, X^* TM^{1,0})$. This is precisely the effect of (6.21) on both the left and right moving sectors. The BRST operator is then the sum of the left-moving and right-moving BRST operators:

$$Q_{BRST} = \oint G^+ dz + \oint \overline{G}^+ d\bar{z} \ . \tag{6.63}$$

Note that for genus $g \neq 1$, the non-linear sigma model is no longer superconformally invariant, nevertheless, we can replace the operation of twisting by redefining the bundles as above, yielding a well defined Q_{BRST}.

We will now restrict attention to the large volume limit where the non-trivial BRST transformations of the fields are:

$$\begin{array}{ll} \delta X^i = \varepsilon \psi^i & \delta X^{\bar{i}} = \varepsilon \overline{\chi}^{\bar{i}} \\ \delta \psi^i = 0 & \delta \chi^{\bar{i}} = \varepsilon \partial X^{\bar{i}} \\ \delta \overline{\psi}^i = \varepsilon \bar{\partial} X^i & \delta \overline{\chi}^{\bar{i}} = 0 \ . \end{array} \tag{6.64}$$

We can now express the action of the A-twisted non-linear sigma model as:

$$S = it\, \delta \left(\int_\Sigma d^2 z V \right) + t \int_\Sigma d^2 z \ X^* \omega, \tag{6.65}$$

where $V = g_{i\bar{j}} \left(\psi_z^{\bar{j}} \partial_{\bar{z}} X^i + \partial_z X^{\bar{j}} \overline{\psi}_{\bar{z}}^i \right)$, and ω is the Kähler form on M. It is hence clear that under a deformation of the complex structure of M, the action will be mapped to itself plus a BRST trivial term. Along the same lines of the argument in (6.20), one can then show that the correlation functions are independent of the complex structure moduli. Another interesting feature of the model is that it localizes on holomorphic maps. This can be shown as follows. Consider a general correlator of fields $\mathscr{O}_i$:

$$\langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = \int \mathscr{D}X \mathscr{D}\Psi e^{iS} O_1 \cdots O_n \,. \tag{6.66}$$

The integral over X can be split as:

$$\int \mathscr{D}X = \sum_{\beta \in H_2(M)} \int_{[X^*(\Sigma)]=\beta} \mathscr{D}X. \tag{6.67}$$

Using expression (6.65), we can then rewrite the correlator as:

$$\langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = \sum_{\beta \in H_2(M)} e^{it\omega\cdot\beta} \int_{[X^*(\Sigma)]=\beta} \mathscr{D}X \mathscr{D}\Psi e^{-t\delta(\int_\Sigma d^2 z V)} O_1 \cdots O_n, \tag{6.68}$$

where we have denoted $\omega \cdot \beta := \int_\Sigma X^* \omega$. We now observe that the integral over each class β is independent of t, in particular we can evaluate it in the limit $t \to \infty$ (classical limit), where the only field configurations contributing to the integral are the solutions to $\delta \left(\int_\Sigma d^2 z V = 0 \right)$. One can then show that those in turn are the solutions to the equations of motion. In particular, the A-model localizes on maps satisfying $\bar{\partial} X = 0$, i.e., holomorphic maps.

A general operator is given by a general combination of all the fields in the theory. Physical operators are, however, in Q_{BRST}-cohomology. Let $\mathscr{O}$ denote a general operator expressed as a sum of operators O_i, with distinct conformal weights (n_i, m_i):

$$\mathscr{O} = \sum_{i=1}^{k} \mathscr{O}_i. \tag{6.69}$$

Then, by degree reasons, if $[Q, \mathscr{O}] = 0$, then $[Q, \mathscr{O}_j] = 0$ for all j. Now let $\mathscr{O}_j$ have non-zero conformal weight, say $n_j \neq 0$, then

$$\mathscr{O}_j = \frac{1}{n_j} [L_0, \mathscr{O}_j] = \frac{1}{n_j} [[Q_{BRST}, G_{-1}^-], \mathscr{O}_j] = \frac{1}{n_j} [Q_{BRST}, [G_{-1}^-, \mathscr{O}_j]]. \tag{6.70}$$

From the above argument we can restrict attention to operators of conformal weight $(0, 0)$. These are of the form

$$\omega = \omega_{i_1,\ldots,i_p,\bar{j}_1,\ldots,\bar{j}_q}(X) \psi^{i_1} \ldots \psi^{i_p} \overline{\chi}^{\bar{j}_1} \ldots \overline{\chi}^{\bar{j}_q} \,. \tag{6.71}$$

The action of the BRST differential is then

$$\delta\omega = \varepsilon(\psi^i \frac{\partial \omega}{\partial X^i} + \chi^{\bar{j}} \frac{\partial \omega}{\partial X^{\bar{j}}}). \tag{6.72}$$

This indicates that we can identify δ with the de Rahm differential $d = \partial + \bar{\partial}$ of M and

$$\{\text{space of } (0,0)\text{-operators}\} = \Omega_d^{\bullet}(M), \tag{6.73}$$

therefore

$$\{\text{space of physical operators}\} = H_{Q_{BRST}} = H_d^{\bullet}(M) = \bigoplus_{p,q=0}^{n} H_d^{p,q}(M), \tag{6.74}$$

where (p, q) are the $(-J, \bar{J})$ charges. On the other hand, the unitarity constraints on the highest weight states corresponding to the chiral primary fields impose:

$$H_{Q_{BRST}} \cong \bigoplus_{p,q=0}^{\hat{c}} \text{Harmonics}^{p,q}(M). \tag{6.75}$$

This provides further evidence for the identity $\hat{c} = \dim(M)$, which can also be derived by just analysing the field content of the theory.

6.1.3 The B-Twist

The B-twist consists in redefining the spinors as sections of the following vector bundles:

$$\psi \in \Gamma(\Sigma, X^*TM^{1,0} \otimes \Omega^{1,0}(\Sigma)), \tag{6.76}$$

$$\chi \in \Gamma(\Sigma, X^*TM^{0,1}), \tag{6.77}$$

and similarly for the right moving spinors:

$$\overline{\psi} \in \Gamma(\Sigma, X^*TM^{1,0} \otimes \Omega^{0,1}(\Sigma)), \tag{6.78}$$

$$\overline{\chi} \in \Gamma(\Sigma, X^*TM^{0,1}). \tag{6.79}$$

The BRST operator is given by:

$$Q_{BRST} = \oint G^- dz + \oint \overline{G}^+ d\bar{z}, \tag{6.80}$$

and in the large volume limit transforms the component fields as follows:

$$\begin{aligned} \delta X^i &= 0 \qquad \delta X^{\bar{i}} = \varepsilon(\chi^{\bar{i}} + \overline{\chi}^{\bar{i}}) \\ \delta\psi^i &= \varepsilon\partial X^i \;\; \delta\chi^{\bar{i}} = 0 \\ \delta\overline{\psi}^i &= \varepsilon\bar{\partial} X^i \;\; \delta\overline{\chi}^{\bar{i}} = 0 \; . \end{aligned} \tag{6.81}$$

In terms of the BRST transformation δ, the action of the B-twisted non-linear sigma model reads:

$$S = it\,\delta\left(\int_\Sigma V\right) + t\,W, \tag{6.82}$$

where $V = g_{i\bar{j}}\left(\psi_z^i \partial_{\bar{z}} X^{\bar{j}} + \partial_z X^{\bar{j}} \overline{\psi}_{\bar{z}}^i\right)$, and

$$W = \int_\Sigma \left(i\theta_i \left(D_z \overline{\psi}_{\bar{z}}^i - D_{\bar{z}} \psi_z^i\right) + \frac{1}{2} R_{i\bar{j}k}^{l}\, \overline{\psi}_{\bar{z}}^i \psi_z^k \eta^{\bar{j}} \theta_l \right) \;, \tag{6.83}$$

with the fields $\eta^{\bar{i}}$ and θ_i being defined as:

$$\eta^{\bar{i}} := \chi^{\bar{i}} + \overline{\chi}^{\bar{i}}, \quad \theta_i := g_{i\bar{j}}(\overline{\chi}^{\bar{j}} - \chi^{\bar{j}}) \; . \tag{6.84}$$

It can be shown that W is invariant, up to a BRST-trivial term, under deformations of the Kähler form of M. This then implies that the correlation functions are independent of Kähler deformations. The B-model is particularly simple as it localizes on constant maps, turning the path integral over maps X into an integral over the target manifold M. This follows from the fact that V does not depend on θ, while W depends on it linearly, allowing to absorb the factor of t in front of W by a redefinition of θ. More precisely:

$$\langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = \int \mathscr{D}X \mathscr{D}\Psi \mathscr{D}\eta \mathscr{D}(t^{-1}\theta) e^{-t\,\delta(\int_\Sigma V) + iW} \prod_i O_i\left(X, \Psi, \eta, t^{-1}\theta\right). \tag{6.85}$$

Assuming the O_i fields are homogeneous in θ, we obtain:

$$\langle \mathscr{O}_1 \cdots \mathscr{O}_n \rangle = \mathscr{N}(t) \int \mathscr{D}X \mathscr{D}\Psi \mathscr{D}\eta \mathscr{D}\theta e^{-t\,\delta(\int_\Sigma V) + iW} \prod_i O_i\,(X, \Psi, \eta, \theta)\,. \tag{6.86}$$

The above integral is independent of t, and we can thus choose to evaluate it in the classical limit $t \to \infty$, where the only field configurations contributing are the ones solving the equation $\delta \int_\Sigma V = 0$. In particular, the path integral over X localizes on constant maps.

The general form of a conformal weight (0, 0) field is:

$$B = B_{\bar{i}_1,\cdots,\bar{i}_p}^{j_1,\cdots,j_q} \eta^{\bar{i}_1} \cdots \eta^{\bar{i}_p} \theta_{j_1} \cdots \theta_{j_p}. \tag{6.87}$$

We can now identify:

$$[Q_{BRST}, .] \leftrightarrow \bar{\partial} \tag{6.88}$$

$$\eta^{\bar{i}} \leftrightarrow \mathrm{d}\bar{X}^{\bar{i}} \tag{6.89}$$

$$\theta_j \leftrightarrow \frac{\partial}{\partial X^j} \tag{6.90}$$

Hence, BRST-cohomology can be identified with the Dolbeault cohomology of the target M:

$$H_{Q_{BRST}} = \bigoplus_{p,q=0}^{n} H_{\bar{\partial}}^{(0,p)}(M, \bigwedge^{q} T^{(1,0)}M) \tag{6.91}$$

6.1.4 Deformations

The non-linear sigma model can be deformed by marginal operators, i.e., operators with left and right conformal weights equal to 1, and R-charge 0. Those correspond to complex structure and Kähler deformations of the target space M. In the A-twist of the model, the marginal operators implementing complex structure deformations become trivial, while in the B-twist the ones implementing Kähler deformations.

6.1.4.1 A-Type Deformations

Deformations of the A-twisted model are provided by operators in BRST-cohomology of R-charges (1,1), i.e., by operators of the form:

$$\phi^a = k^a_{i\bar{j}} \psi^i \chi^{\bar{j}}, \tag{6.92}$$

where $k^a \in H^{(1,1)}(M, \mathbb{R})$. The associated marginal operator is then:

$$(\phi^a)^{(1,1)} = G^-_{-1} \overline{G}^-_{-1} \phi^a = k^a_{i\bar{j}} \partial X^i \bar{\partial} X^{\bar{j}} + \ldots \tag{6.93}$$

Given a basis $\{k^a\}_{a=1,\ldots,h^{1,1}}$ of $H^{(1,1)}(M, \mathbb{R})$, we can deform the action as follows:

$$S \mapsto S + \left(\sum_{a=1}^{h^{1,1}} t_a (\phi^a)^{1,1} + h.c. \right), \tag{6.94}$$

where the t_a's are complex numbers called complexified Kähler moduli. The hermitian conjugate term in the above expression has the form:

$$\sum_{a=1}^{h^{1,1}} \bar{t}^a \int_\Sigma G_0^+ \overline{G}_0^+ \overline{\phi}^a . \tag{6.95}$$

This is a BRST trivial term and thus does not contribute to the correlation functions. It is convenient to express the deformation of the action as follows:

$$(t_a + \bar{t}_a) \int_\Sigma k^a_{i\bar{j}} \left(\partial X^i \bar{\partial} X^{\bar{j}} + \partial X^{\bar{j}} \bar{\partial} X^i \right) + (t_a - \bar{t}_a) \int_\Sigma k^a_{i\bar{j}} \left(\partial X^i \bar{\partial} X^{\bar{j}} - \partial X^{\bar{j}} \bar{\partial} X^i \right) + \dots \tag{6.96}$$

The term multiplying $(t_a + \bar{t}_a)$ implements the following deformation to the original action:

$$g(\cdot, \cdot) \mapsto g(\cdot, \cdot) + (t_a + \bar{t}_a) k(I \cdot, \cdot), \tag{6.97}$$

where I is the complex structure. The term multiplying $(t_a - \bar{t}_a)$ instead deforms the imaginary part of the Kähler form, also known as the B-field.

Within this formalism, we can recover the statement that the A-model localizes on holomorphic maps as follows. The triviality of the deformations parametrized by $\bar{t}$ allows us to evaluate correlation functions in the limit $\bar{t} \to \infty$. The bosonic part of the trivial deformation has the form

$$\int_\Sigma k^a_{i\bar{j}} \partial X^{\bar{j}} \bar{\partial} X^i . \tag{6.98}$$

Therefore, the only bosonic field configurations contributing to the correlation functions satisfy $\bar{\partial} X = 0$, i.e., are holomorphic maps.

6.1.4.2 B-Type Deformations

In direct analogy to the A-model, deformations of the B-model are provided by physical fields of R-charges $(1, 1)$, i.e., of the form:

$$\phi^a = (b^a)^j_{\bar{i}} \, \eta^{\bar{i}} \, \theta_j , \tag{6.99}$$

where $b^a \in H^{(0,1)}(M, T^{(1,0)}M)$. The associated marginal operator is

$$(\phi^a)^{(1,1)} = G^+_{-1} \overline{G}^-_{-1} \phi^a . \tag{6.100}$$

6.1.5 The Chiral Anomaly Revisited

In (Sect. 6.1.1), we discovered that the non-linear sigma model possesses an anomaly of the axial $U(1)$ transformation after quantization. The twisted theories (A and

B-models) can be viewed as possessing the same fields as the NLSM but having a shifted action (6.21). The shift in the action is, however, invariant under axial $U(1)$ transformations, therefore we expect to obtain the same anomaly. If we wish to use the natural fields in the formulation of the twisted theories, then, following the analysis in (Sect. 6.1.1), we find that the axial anomaly is measured by the discrepancy in the number of zero modes of the Dolbeault Laplacian operator on holomorphic X^*TM valued one-forms and X^*TM valued zero forms. Indeed, expanding the one forms and zero forms in the eigenbasis of the Dolbeault Laplacian $\Delta_\partial := \partial\partial^\dagger + \partial^\dagger\partial$, we find that the non-zero modes are paired up. Let's denote $\Gamma(\Sigma, X^*TM^{0,1} \oplus \Omega^{(1,0)}) =: \Omega^1$ and $\Gamma(\Sigma, X^*TM^{0,1}) =: \Omega^0$. Let $f \in \Omega^0$ s.t. $\Delta_\partial f = \lambda f$ with $\lambda \neq 0$, then since $[\Delta_\partial, \partial] = 0$, $\Delta_\partial \partial f = \lambda \partial f$. While for $\alpha \in \Omega^1$ s.t. $\Delta_\partial \alpha = \lambda\alpha$ with $\lambda \neq 0$, it follows (α is a top form) $\alpha = \partial(\frac{1}{\lambda}\partial^\dagger\alpha)$ and $\Delta_\partial(\partial^\dagger\alpha) = \lambda(\partial^\dagger\alpha)$. From this discussion, plugging the expansions in the path integral measure one finds that the anomaly is equal to:

$$k = 2(\dim \mathrm{Ker}\Delta|_{\partial\Omega^1} - \dim \mathrm{Ker}\Delta|_{\partial\Omega^0}) \tag{6.101}$$

$$= 2(h^1(\Sigma, X^*TM) - h^0(\Sigma, X^*TM)) = 2\chi(X^*TM), \tag{6.102}$$

where χ denotes the arithmetic genus. By the Hirzebruch-Riemann-Roch Theorem this can be expressed as

$$k = 2\int_\Sigma ch(X^*TM) \wedge td(T\Sigma) = 2\beta \cap c_1(T^*M) + 2\dim_{\mathbb{C}}(M)(1-g), \tag{6.103}$$

where $\beta = [X(\Sigma)] \in H_2(M)$.

6.1.6 Topological D-Branes on Calabi–Yau Manifolds

So far, our discussion was restricted to field theories defined on closed Riemann surfaces. In order to extend these theories to worldsheets with boundaries, we have to study the possible boundary conditions that we can impose on the fields. The presence of a boundary will, in general, break part of the symmetries of the original theory. In particular, translational invariance in the direction normal to the boundary will be broken, as a consequence of which we lose some, or all, of the supersymmetries (since $\{Q, \overline{Q}\} \sim P$). The boundary conditions we will be interested in are the ones that preserve half of the supercharges, namely Q_A or Q_B, so that the resulting theory is compatible with the A-/B-twist. In the language of string theory, such boundary conditions are called, correspondingly, A-type/B-type D-branes.

In this section we will focus on the description of D-branes in nonlinear sigma models with a Calabi–Yau target. We will be working in the large volume region of the Kähler moduli space, where the sigma model can be studied perturbatively, and where notions of classical geometry apply. In this limit, the D-brane can be interpreted as specifying a certain submanifold L of the target space, on which the open string has to end (i.e. $X(\partial\Sigma) \subset L$). Of course, this picture will break down in the non-geometric regime of the Kähler moduli space, and one has to look for a more appropriate description of D-branes there.

Since the subject of the present section is rather involved, we shall only provide a simplified account of the main results, leaving the various technical details to [1, 2], and references therein.

6.1.6.1 Boundary Conditions for the $\mathcal{N} = 2$ SCA

In order to preserve the A-/B-type BRST operators one has to impose an appropriate boundary condition on the supercurrents. Recall that in the A-model, the BRST operator is given by the supercharge $Q_A = \int G^+ dz + \int \overline{G}^+ d\bar{z}$. For this to be preserved, at $z = \bar{z}$, the supercurrents have to satisfy:

$$G^+ = \pm\overline{G}^+, \quad G^- = \pm\overline{G}^-, \tag{6.104}$$

where the sign corresponds to the usual ambiguity associated with fermions. Similarly, the BRST operator in the B-model, $Q_B = \int G^+ dz + \int \overline{G}^- d\bar{z}$, is preserved under the boundary condition:

$$G^+ = \pm\overline{G}^-, \quad G^- = \pm\overline{G}^+. \tag{6.105}$$

Notice that both (6.104) and (6.105) preserve the $\mathcal{N} = 1$ subalgebra:

$$T = \overline{T}, \quad G = \pm\overline{G} \tag{6.106}$$

with $G = G^+ + G^-$.

In this section, it is convenient to choose a slightly different notation for the fermionic fields: $\psi^I := (\psi^i, \chi^{\bar{i}})$ and similarly for $\overline{\psi}^I$. In general, a boundary condition relates the left-moving and right-moving sector as

$$\partial X^I = R^I_J(X)\bar{\partial}X^J, \quad \psi^I = R^I_J(X)\overline{\psi}^J, \tag{6.107}$$

for some matrix R, where the indices I, J run through both the holomorphic and anti-holomorphic coordinates. It is easy to see that for (6.106) to be satisfied, R has to fulfill:

$$g_{IJ} R^I_K R^J_L = g_{KL}, \tag{6.108}$$

i.e. R has to be an orthogonal matrix with respect to the Kähler metric.

6.1.6.2 A-Branes

Let us now analyze what kind of submanifolds can an A-type D-brane define. Notice first that consistency of (6.107) with the BRST transformations of the A-twist (6.64) requires $R^i_{\ j} = R^{\bar{i}}_{\ \bar{j}} = 0$. Let us now take an eigenvector v of R with eigenvalue $+1$. In this direction we obtain a Neumann boundary condition, and hence, v is a vector tangent to L. Since the complex structure J is diagonal in holomorphic coordinates:

$$J^m_n = i\delta^m_n, \quad J^{\bar{m}}_{\bar{n}} = -i\delta^{\bar{m}}_{\bar{n}}, \tag{6.109}$$

it is easy to see that the vector Jv has eigenvalue -1 with respect to R, and thus, it defines a direction normal to L. But $J^2 v = -v$, which means that tangent and normal directions are paired, so L must be necessarily of middle dimension. Moreover, since for any two tangent vectors v, w, the vector Jw is orthogonal to v, the Kähler form vanishes on L, which makes it a Lagrangian submanifold.

In general, one can introduce a gauge field A on L, which can be included in the action by adding the term:

$$S_{\partial\Sigma} = t \int_{\partial\Sigma} \phi^*(A). \tag{6.110}$$

This is to be interpreted as coupling the gauge field to the open string endpoint. Requiring BRST invariance of the action gives a constraint on A. Namely, the field strength $F = dA$ has to be vanishing, or in other words, the gauge bundle has to be flat.

We have thus shown that A-branes correspond to Lagrangian submanifolds with flat gauge bundles. Let us note that what we have described are, more precisely, *topological* A-branes, which are only required to preserve $\mathcal{N} = 2$ *worldsheet* supersymmetry. The *physical* (stable) D-branes, on the other hand, have to preserve *spacetime* supersymmetry, and it turns out that this is achieved when the A-brane corresponds to a *special* Lagrangian submanifold.

6.1.6.3 B-Branes

Let us now turn to the discussion of B-branes. Proceeding similarly as in the case of A-branes, we find that the boundary condition (6.107) is consistent with the B-twist only if $R^{\bar{i}}_{\ j} = R^i_{\ \bar{j}} = 0$. The tangent and the normal bundle of L are now invariant under the complex structure J, which means that L is a complex submanifold. Furthermore, inclusion of a gauge field is possible only if the field strength is a (1,1)-form (again by BRST invariance of the action). We thus arrive at the result that B-type D-branes correspond to complex submanifolds with a holomorphic vector bundle.

It turns out that the description of B-branes in terms of vector bundles is, in general, not adequate. One way to see this is to consider bound states of B-branes of different dimensions. Such a configuration can no longer be represented by a vector bundle, and a more appropriate notion is that of a *sheaf*. Even more generally, D-branes can be viewed as objects of a *category*, with the morphisms being the open string stretching between them. B-branes are then objects of the *derived category of coherent sheaves*, while A-branes live in the *Fukaya* category. We refer to [3, 4] for a discussion of these topics.

6.2 Coupling to Gravity

We start this section with a brief reminder of the essential features involved in coupling the bosonic string to gravity. The symmetry group that leaves the action invariant is $G := \mathit{Diff} \times \mathit{Weyl}$. The path integral measure consists in a measure on the space of metrics times a measure on the space of maps from the Riemann surface to the target divided by the action of G. Then one resorts to the Fadeev-Popov procedure to express the path integral as an integral over the moduli space $\mathscr{M}_g$ of the Riemann surface (of genus g) times the path integral of the "matter + ghost" system. The "matter + ghost" system is an $\mathscr{N} = 2$ superconformal theory where the BRST operator is a combination of left and right supercharges and the energy momentum tensor is BRST trivial.

Coming back to the A and B models, we define coupling to gravity in direct analogy to the bosonic string, by viewing both models as a "matter + ghost" system.

6.2.1 The Measure on $\mathscr{M}_g$

$\mathscr{M}_g$ is equivalent to the moduli space of complex structures on a Riemann surface Σ of genus g. Infinitesimal variations of the complex structure are parametrized by elements $\eta \in H^{(1,0)}(\Sigma, T^{(0,1)}\Sigma)$ called Beltrami differentials. More precisely, these are first order solutions to the Maurer–Cartan equation for the Dolbeault differential. One way of computing the (complex) dimension $h^1(\Sigma, T\Sigma)$ of the space of Beltrami differentials is via the now familiar Hirzebruch–Riemann–Roch formula:

$$h^1(\Sigma, T\Sigma) - h^0(\Sigma, T\Sigma) = \int_\Sigma ch(T\Sigma) \wedge td(T\Sigma) \tag{6.111}$$

$$= \int_\Sigma c_1(T\Sigma) + \dim_{\mathbb{C}}(\Sigma)(1-g) \tag{6.112}$$

$$= 3g - 3. \tag{6.113}$$

In the case of genus 1, we can use the holomorphic top form to establish $H^{(0,0)}(\Sigma, T^{(0,1)}\Sigma) = H^{0,0}(\Sigma) = \mathbb{C}$ as the surface is compact, therefore $h^1(\Sigma, T\Sigma) = 1$.

For genus $g \geq 2$, $h^0(\Sigma, T\Sigma) = 0$. In terms of the dimension of $\mathscr{M}_g$, $g = 0, 1$ are special. For $g = 1$ we recover the familiar answer $\dim(\mathscr{M}_1) = 1$, which is indeed the dimension of the fundamental domain of the upper half plane $\mathbb{H}/SL(2, \mathbb{Z})$. For genus $g = 0$ instead, $\dim(\mathscr{M}_g) = h^1(\Sigma, T\Sigma) = 0$.

In direct analogy with the bosonic string, we now turn to defining the measure on $\mathscr{M}_g$. First we restrict attention to Riemann surfaces with genus $g > 1$. Let $\eta_1, \ldots, \eta_{3g-3} \in H^{(1,0)}(\Sigma, T^{(0,1)}\Sigma)$ be a basis of $T\mathscr{M}_g$, and choose coordinates $m_1, \ldots, m_{3g-3}$. In the case of the A-model we define the measure on $\mathscr{M}_g$ as:

$$\omega_g = \bigwedge_{i=1}^{3g-3} dm_i \wedge d\overline{m}_i \left\langle \prod_{i=1}^{3g-3} G^-, \eta_i \prod_{i=1}^{3g-3} \left(\overline{G}^-, \overline{\eta}_i\right) \right\rangle, \tag{6.114}$$

where

$$(G^-, \eta_i) := \int_\Sigma G^-_{zz} (\eta_i)^z_{\bar{z}}\, dz \wedge d\bar{z}, \tag{6.115}$$

and G^- is identified with b, the “b” ghost of the bosonic string. For central charge $\hat{c} = \dim_{\mathbb{C}} M = 3$, the insertion of G^- and $\overline{G}^-$ exactly cancels the $U(1)_A$ anomaly (6.103). The genus $g > 1$ partition function of the A-model coupled to gravity is then given by

$$F_g = \int_{\mathscr{M}_g} \omega_g \ . \tag{6.116}$$

In the case of the B-model one has to substitute $\overline{G}^-$ with $\overline{G}^+$.

For $g = 1$ we again follow the analysis of the bosonic string, and it is natural to express the answer in the operator formalism:

$$\omega_1 = \frac{1}{2}\mathrm{Tr}\left[(-1)^F F_L F_R q^{H_L} \bar{q}^{H_R}\right], \tag{6.117}$$

where $F_L (F_R)$ and $H_L (H_R)$ denote the left (right) fermion number and Hamiltonian operators respectively, while $q = e^\tau$ with τ parametrizing the complex structure of the torus. Here the prefactor of $1/2$ reflects the $\mathbb{Z}_2$ symmetry of the torus. Up to a constant prefactor, ω_1 is essentially the only index for $N = (2, 2)$ theories, which for sigma models amounts to the statement that it is invariant under trivial deformations of the metric of the target Calabi–Yau manifold and is consequently well defined on its moduli space. The measure on $\mathscr{M}_1$ is the one induced by the $SL(2, \mathbb{R})$ invariant measure on $\mathbb{H}$. We thus obtain

$$F_1 = \int_{\mathscr{M}_1} \frac{d^2\tau}{\mathrm{Im}(\tau)} \omega_1. \tag{6.118}$$

For $g = 0$ coupling to gravity is trivial, and the correlation functions are the ones of the topological conformal field theory. The correlator is anomaly free only if the total left and right $U(1)$ charge of the inserted local operators is equal to 3. If we only allow local operators whose integrated descendants are marginal, then the only well defined correlator is:

$$\langle \phi_i(0)\phi_j(1)\phi_k(\infty)\rangle =: C_{ijk}. \tag{6.119}$$

The C_{ijk} are called Yukawa couplings, as in the effective gauge theory description, they correspond to the interaction vertices for the vector multiplets. We will now analyse the 3-point correlation function for the A and B-model.

6.2.2 The Genus Zero Generating Function

6.2.2.1 A-Model

In the A-model, the path integral localizes on holomorphic maps. It is therefore convenient to consider the moduli space $\mathscr{M}(\beta)$ of holomorphic maps $X : \Sigma \to M$, with $[X(\Sigma)] = \beta \in H_2(M, \mathbb{Z})$. By careful inspection of the path integral, after integrating over non-zero modes Ψ, the 3-point correlator reduces to

$$C_{ijk} = \langle \phi_i(0)\phi_j(1)\phi_k(\infty)\rangle = \sum_{\beta\in H_2(M,\mathbb{Z})} e^{it\omega\cdot\beta}\int_{\mathscr{M}(\beta)} ev_1^*(\omega_1)\wedge ev_2^*(\omega_2)\wedge ev_3^*(\omega_3) \tag{6.120}$$

Here ev_i denotes the evaluation map $ev_i : \mathscr{M}(\beta) \to M$ at the point $z_i \in \mathbb{CP}^1$ (here $z_i \in \{0, 1, \infty\}$), and $\omega_i \in H^{1,1}(M)$ is the Kähler form corresponding to ϕ_i. Although intuitive, it requires further effort to show that

$$C_{ijk} = \int_M \omega_1\wedge\omega_2\wedge\omega_3 + \sum_{0\neq\beta\in H_2(M,\mathbb{Z})} e^{it\omega\cdot\beta} N_\beta(\omega_1\cap\beta)(\omega_1\cap\beta)(\omega_1\cap\beta), \tag{6.121}$$

where N_β is the number of primitive rational curves of class β, also known as Gromov-Witten invariants. We define $q_i := e^{2\pi i t_i}$ and decompose ω and β as follows:

$$\omega = \sum_{l=1}^{h^{1,1}} t^l e_l, \quad \beta = \sum_{l=1}^{h^{1,1}} r_l (e^l)^\vee. \tag{6.122}$$

Recall that $t\omega\cdot\beta := t\int_{\mathbb{CP}^1} X^*\omega = k\omega\cap\beta$ where k is the arbitrary degree of X. Then we obtain:

$$C_{ijk} = \int_M \omega_1\wedge\omega_2\wedge\omega_3 + \sum_r r_i r_j r_k N_{0,r} \sum_{k\geq 1} q^{kr_1}\cdots q^{kr_n} \tag{6.123}$$

$$= \int_M \omega_1 \wedge \omega_2 \wedge \omega_3 + \sum_r r_i r_j r_k N_{0,r} \frac{q^{r_1} \cdots q^{r_n}}{1 - q^{r_1} \cdots q^{r_n}} \,. \tag{6.124}$$

Here $n = h^{1,1}$, and we have chosen to denote N_β by $N_{0,r}$ to emphasize that these numbers count genus zero curves.

6.2.2.2 B-Model

In the B-model, the path integral localizes on constant maps, therefore the path integral reduces to an integral over the Calabi–Yau target M. In particular, one can show:

$$C_{ijk} = \langle \phi_i(0)\phi_j(1)\phi_k(\infty)\rangle = \int_M \langle \mu_i \wedge \mu_j \wedge \mu_k, \Omega\rangle \wedge \Omega, \tag{6.125}$$

where $\mu_i \in H^{(0,1)}(M, X^* T^{(1,0)}M)$ is the Beltrami differential associated to the local operator ϕ_j. Up to rescaling, Ω is the unique holomorphic top-form on the Calabi–Yau. The above has the following geometric description. Let $\mathscr{L}$ denote the line bundle of holomorphic top forms on the moduli space of complex structures $\mathscr{M}_I$ of M. $\mathscr{L}$ is endowed with the fiberwise metric:

$$g(\Omega, \Omega) = \int_M \overline{\Omega} \wedge \Omega = e^{-K}. \tag{6.126}$$

The above is actually not well defined on the line bundle, as under a change of gauge $\Omega \mapsto f\Omega$, with f holomorphic, $K \mapsto K - f - \bar{f}$. K is however, the Kähler potential for a Kähler metric on $\mathscr{M}_I$. Indeed

$$g_{a\bar{b}} = -\partial_a \bar{\partial}_{\bar{b}} K = -\partial_a \frac{\int_M \bar{\partial}_{\bar{b}} \overline{\Omega} \wedge \Omega}{\int_M \overline{\Omega} \wedge \Omega} \tag{6.127}$$

$$= -\frac{\int_M \bar{\partial}_{\bar{b}} \overline{\Omega} \wedge \partial_a \Omega}{\int_M \overline{\Omega} \wedge \Omega} + \frac{\int_M \bar{\partial}_{\bar{b}} \overline{\Omega} \wedge \Omega \int_M \overline{\Omega} \wedge \partial_a \Omega}{(\int_M \overline{\Omega} \wedge \Omega)^2} \tag{6.128}$$

is well defined. The labels a and $\bar{b}$ stand for holomorphic and anti-holomorphic coordinates on $\mathscr{M}_I$ yet to be defined. $\partial_a \Omega$ corresponds to a first order deformation of the complex structure of M parametrized by Ω. A deformation of Ω arises from a deformation of the holomorphic coordinate one form $\delta dz^i = \langle \mu_a, dz^i \rangle$, where $\mu_a \in H^{(0,1)}(M, X^* T^{(1,0)}M)$. In particular, $\partial_a \Omega \in H^{3,0}(M) \oplus H^{2,1}(M)$. g therefore defines a metric on the Hodge bundle on $\mathscr{M}_I$. This has fiber $H^{3,0}(M) \oplus H^{2,1}(M) \oplus H^{1,2}(M) \oplus H^{0,3}(M)$. The Hodge bundle furnishes natural coordinates on $\mathscr{M}_I$. These are defined as follows. Let $\{\alpha_I, \beta^I\}_{I=1,\ldots,1+h^{1,2}}$ be a symplectic basis of half dimensional cycles of M, i.e.

$$\alpha_I \cap \alpha_J = 0, \quad \alpha_I \cap \beta^J = \delta_I^J \,. \tag{6.129}$$

Consider

$$X^I := \int_{\alpha_I} \Omega, \quad F_J := \int_{\beta^J} \Omega. \tag{6.130}$$

The F^J's enjoy the following integrability property:

$$\frac{\partial}{\partial X^I} F_J = \frac{\partial}{\partial X^J} F_I. \tag{6.131}$$

This implies the existence of a local potential F:

$$F_J = \frac{\partial}{\partial X^J} F. \tag{6.132}$$

As the F_J's are homogeneous of degree 1 in the X^I's, we can choose F as

$$F = \frac{1}{2} X^I F_I. \tag{6.133}$$

The X^I's define holomorphic coordinates on $\mathscr{M}_I$. In terms of them we can express K as:

$$K = \bar{X}^I F_I - X^I \bar{F}_I. \tag{6.134}$$

F therefore deserves the name *prepotential*. Let μ_I denote the Beltrami differential corresponding to the deformation $\frac{\partial}{\partial X^I}\Omega$, then

$$C_{IJK} = \partial_I \partial_J \partial_K F. \tag{6.135}$$

6.2.3 Higher Genera Generating Functions

Contrary to the pre-potential F, the higher genus generating functions are not holomorphic, in particular the localization principle does not apply. Instead, the obstruction to this property is captured by the Holomorphic anomaly equation. Although of conceptual importance, this equation also turns out useful in the computation of F_g itself. Next we will consider the case of $g \geq 2$. Applying $\frac{\partial}{\partial \bar{t}^a}$ to F_g results in computing the expectation value of the conjugate to a marginal operator. The left (right) trivial supercharge together with a left (right) b-ghost insertion pair up to yield the product of a left and right energy momentum tensor. The latter acts as an exterior differential on the moduli space $\mathscr{M}_g$, as it implements infinitesimal variations of the metric on the worldsheet. In summary, we obtain:

$$\frac{\partial}{\partial \bar{t}^{\bar{a}}} F_g = \int_{\mathscr{M}_g} \partial \bar{\partial} \omega_g^i. \tag{6.136}$$

Since $\partial\bar{\partial} = d\bar{\partial}$, we have:

$$\frac{\partial}{\partial \bar{t}^{\bar{a}}} F_g = \int_{\partial \mathscr{M}_g} \bar{\partial}\omega_g^i. \tag{6.137}$$

The holomorphic anomaly therefore is partly due to the non-vanishing boundary of $\mathscr{M}_g$. A pictorial description of $\partial\mathscr{M}_g$ is as follows. Points of the boundary correspond to degenerations of the Riemann surface Σ_g. In particular, these fall in two categories. The first consists of surfaces where a cycle degenerates to divide Σ_g into $(\Sigma_{g-k}, \star) \cup (\Sigma_k, \star)$, where $\star$ denotes a marked point. The second consists of surfaces where a cycle degenerates to reduce Σ_g to $(\Sigma_{g-1}, \star_1, \star_2)$. The process of degeneration descends properly on BRST cohomology. Indeed, one can consider the degeneration as the development of an infinitely long and thin cylinder carrying therefore a time evolution for indefinite time. The only states surviving the evolution are the zero modes. Those, however, are the harmonic representatives of the chiral ring. The degeneration of the surfaces reflects the factorisation of the corresponding correlation functions to yield:

$$\frac{\partial}{\partial \bar{t}^{\bar{i}}} F_g = \frac{1}{2} e^{-2K} C_{\bar{i}}^{jl} \left(\sum_{k=1}^{g-1} D_j F_{g-k} D_l F_k + D_j D_l F_{g-1} \right). \tag{6.138}$$

Note that in the above expression we have introduced covariant derivatives D_i. These are the appropriate operators inserting chiral primaries. The corresponding connection is Levi–Civita with respect to the Kähler metric on the relevant moduli space of the Calabi–Yau. The details of such a construction are part of the well known tt^* equations, which however, are beyond the scope of this lecture.

The formula (6.138) is recursive in g. This suggests the possibility of solving it in perturbation theory. Indeed, at present this appears to be the most efficient way of computing topological string amplitudes. There are of course boundary conditions that (6.138) has to be supplemented with. Conceptually under control is the dependence on the anti-holomorphic coordinates $\bar{t}^i$. This is completely fixed by specifying F_g at $\bar{t} \to \infty$. Here the localization principle applies and one can derive explicit expressions. We will encounter later a way of reproducing these expressions via large N duality. Finally, one has to fix the holomorphic dependence of F_g. This, however, is harder to deal with and to date one has to deal with meticulous case study.

6.2.4 Examples of Calabi–Yau Manifolds

In Sect. 6.1.1.2 we defined the notion of a Calabi–Yau manifold as a Kähler manifold with vanishing first Chern class. We now provide some explicit examples of Calabi–Yau manifolds in complex dimension 3 that will be of importance later in these lectures.

6.2.4.1 Local $\mathbb{CP}^2$

The first example we will consider is the total space of the holomorphic line bundle $\mathcal{O}(-3) \to \mathbb{CP}^2$. This space admits a description as a quotient space

$$X = \left(\mathbb{C}^4 \setminus \{(z_1, z_2, z_3) = 0\}\right) / \sim \tag{6.139}$$

where the equivalence relation is defined by:

$$(x, z_1, z_2, z_3) \to (\lambda^{-3} x, \lambda z_1, \lambda z_2, \lambda z_3), \quad \lambda \in \mathbb{C}^* . \tag{6.140}$$

It is easy to check that the first Chern class of this space is zero:

$$c_1(X) = c_1(\mathbb{CP}^2) + c_1\left(\mathcal{O}(-3)\right) = 3 + (-3) = 0 , \tag{6.141}$$

so that the Calabi–Yau condition is satisfied.

In fact, this is an example of a non-compact Calabi–Yau threefold with 2 compact dimensions represented by the base $\mathbb{CP}^2$ and 1 non-compact dimension corresponding to the fibres of the line bundle. This non-compact space can, however, also arise in the study of compact Calabi–Yau manifolds containing a $\mathbb{CP}^2$, describing the local geometry in the vicinity of the $\mathbb{CP}^2$.

6.2.4.2 Local $\mathbb{CP}^1$

Another simple example is provided by the total space of the vector bundle $\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{CP}^1$. This can be again realized as a quotient space

$$X = \left(\mathbb{C}^4 \setminus \{(z_1, z_2) = 0\}\right) / \sim \tag{6.142}$$

with the equivalence relation being:

$$(x_1, x_2, z_1, z_2) \sim (\lambda^{-1} x_1, \lambda^{-1} x_2, \lambda z_1, \lambda z_2), \quad \lambda \in \mathbb{C}^* . \tag{6.143}$$

The computation of the first Chern class

$$c_1(X) = c_1(\mathbb{CP}^1) + c_1\left(\mathcal{O}(-1) \oplus \mathcal{O}(-1)\right) = 2 + (-2) = 0 \tag{6.144}$$

reveals that this is again a Calabi–Yau threefold, this time with 1 compact and 2 non-compact dimensions.

6.2.4.3 Conifolds

Of crucial importance in our later discussion of geometric transitions will be the so-called conifold geometry. This space is defined as the vanishing locus of the polynomial

$$p = xy - uv \tag{6.145}$$

in $\mathbb{C}^4$. In contrast to the previous examples, this space doesn't correspond to a smooth manifold. Indeed, one easily sees that the equations $p = 0$ and $dp = 0$ have a common solution at the origin, and thus the geometry will be singular at this point.

One way to smooth out the singularity is to deform the defining equation as:

$$xy - uv = \mu^2 \,, \tag{6.146}$$

with μ a real number, which can be naturally interpreted as a complex structure modulus. The resulting geometry is called the *deformed conifold*. Let us discuss this geometry in some more detail. In order to do this, it will be convenient to change coordinates in the following way:

$$z_1 = \frac{x+y}{2} \qquad z_2 = i\frac{x-y}{2} \tag{6.147}$$

$$z_3 = i\frac{u+v}{2} \qquad z_4 = \frac{u-v}{2}. \tag{6.148}$$

In these new variables Eq. (6.146) becomes:

$$z_1^2 + z_2^2 + z_3^2 + z_4^2 = \mu^2 \,. \tag{6.149}$$

Decomposing z_i into a real and imaginary part, $z_i = q_i + ip_i$, this equation yields:

$$\mathbf{q}^2 - \mathbf{p}^2 = \mu^2, \quad \mathbf{q} \cdot \mathbf{p} = 0 \,. \tag{6.150}$$

To get a better picture of the geometry, let us study the slices $\mathbf{q}^2 + \mathbf{p}^2 = r^2$ for $r \geq 0$. It is an easy exercise to show that each slice corresponds to an S^2 of radius $\sqrt{(r^2 - \mu^2)/2}$ fibered over an S^3 of radius $\sqrt{(r^2 + \mu^2)/2}$. Since any such fibration is trivial, we obtain $S^3 \times S^2$. Noting that for $r = \mu$ the size of the S^2 shrinks to zero, we can represent the geometry of the deformed conifold as in Fig. 6.1. Note also that the minimal radius of the S^3 is equal to μ, which means, in particular, that for the singular conifold ($\mu = 0$) both the S^2 and S^3 shrink to zero size at $r = 0$.

It is important to note that the deformed conifold is actually the total space of the cotangent bundle T^*S^3. This can be seen by defining $\mathbf{q}' = \mathbf{q}/\sqrt{\mu^2 + \mathbf{p}^2}$ in terms of which (6.150) becomes:

$$\mathbf{q}'^2 = 1, \quad \mathbf{q}' \cdot \mathbf{p} = 0 \,. \tag{6.151}$$

This is precisely the equation for the total space of T^*S^3.

There is also another possibility to remove the conifold singularity, namely, by replacing the singular point by the space $\mathbb{CP}^1$.[1] To be more precise, we consider the space $Z \subset \mathbb{C}^4 \times \mathbb{CP}^1$ defined by the equation:

$$\begin{pmatrix} x' & u' \\ v' & y' \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = 0, \tag{6.152}$$

[1] In mathematical literature this is referred to as *blowing-up* the singularity.

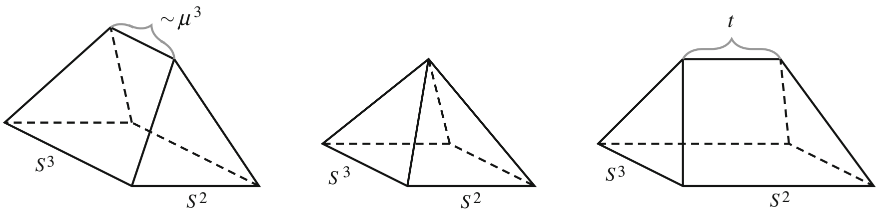


Fig. 6.1 From *left* to *right*: the deformed, singular, and resolved conifold

with $(x', y', u', v') \in \mathbb{C}^4$, and $(\xi_1, \xi_2) \in \mathbb{CP}^1$. It is not difficult to show that this space is isomorphic to the singular conifold for $(x', y', u', v') \neq 0 \neq (x, y, u, v)$. At the point $(x', y', u', v') = 0$, we can have (ξ_1, ξ_2) arbitrary. In effect, we thus replace the singularity at the origin by a $\mathbb{CP}^1 \cong S^2$. The resulting geometry, called the *resolved conifold*, can be represented as in Fig. 6.1. Note that the size of the $\mathbb{CP}^1$ gives rise to a Kähler modulus t, and the singular conifold is recovered in the limit $t = 0$.

Let us finally note that the geometry of the resolved conifold is precisely that of local $\mathbb{CP}^1$, which we discussed earlier. To see this, we decompose Z into two patches, $U_1 = \{\xi_1 \neq 0\}$ and $U_2 = \{\xi_2 \neq 0\}$. Using (6.152) we obtain the coordinates (s, u', y') with $s = \xi_2/\xi_1$ on U_1, and (t, x', v') with $t = \xi_1/\xi_2$ on U_2. On the overlap we have $u' = -s^{-1}x'$ and $y' = -t^{-1}v'$, which are precisely the transformation rules for $\mathscr{O}(-1) \oplus \mathscr{O}(-1)$.

6.2.4.4 Toric Calabi–Yau Threefolds

A large class of Calabi–Yau threefolds can be constructed with the methods of toric geometry. Instead of giving a thorough introduction into this subject (which can be found, e.g., in [1, 5]), we provide only a simplified account which shall be sufficient for our purposes.

Our starting point will be the space $\mathbb{C}^{N+3} \setminus \{0\}$ which we divide by the action of the group $U(1)^N$. In order to do this, we have to assign $U(1)$ charges $Q_i^a,\ a = 1, \ldots N$ to the variables z_i, in terms of which the $U(1)^N$ action is defined as:

$$z_i \to e^{i \sum_i Q_i^a \theta^a} z_i \ , \tag{6.153}$$

with $\theta^a \in [0, 2\pi]$. Furthermore, we impose the following constraint on the variables:

$$\sum_i Q_i^a |z_i|^2 = t^a \ . \tag{6.154}$$

This condition can be thought of as an analogue of the Gauss law constraint in gauge theory. Roughly speaking, we are fixing the phase and the modulus of N of the $N+3$

complex variables, so, in general, we expect to arrive at a complex three-dimensional manifold. It turns out that in order to obtain a Calabi–Yau manifold, one has to impose an additional condition on the charges:

$$\sum_i Q_i^a = 0 \; . \tag{6.155}$$

Let us note that this construction is similar to the way we defined the geometries of local $\mathbb{CP}^2$ and local $\mathbb{CP}^1$, except that in that case, we were dividing by the group $(\mathbb{C}^*)^N$. However, it is a standard fact in the theory of symplectic reduction that the two constructions yield the same result. Let us also note that the parameters t^a in (6.154) have an interpretation as Kähler moduli of the manifold.

We will now explain how the toric construction can be used to obtain a convenient representation of the geometries in terms of so-called *toric diagrams*. We illustrate this on the example of local $\mathbb{CP}^2$. Let us thus consider the four complex coordinates (z_0, z_1, z_2, z_3), and divide by the group $U(1)$ with $Q_0 = -3$, and $Q_i = 1,\; i = 1, 2, 3$, that is, we identify:

$$(z_0, z_1, z_2, z_3) \sim (e^{-3i\theta} z_0, e^{i\theta} z_1, e^{i\theta} z_2, e^{i\theta} z_3) \; . \tag{6.156}$$

The constraint (6.154) takes on the following form:

$$-3|z_0|^2 + |z_1|^2 + |z_2|^2 + |z_3|^2 = r \; . \tag{6.157}$$

Writing $z_i = \sqrt{p_i} e^{i\phi_i}$ with $p_i \geq 0$, and $\phi_i \in [0, 2\pi]$, we obtain from (6.154), (6.156):

$$(\phi_0, \phi_1, \phi_2, \phi_3) \sim (\phi_0 - 3\theta, \phi_1 + \theta, \phi_2 + \theta, \phi_3 + \theta) \; , \tag{6.158}$$

$$-3p_0 + p_1 + p_2 + p_3 = r \; . \tag{6.159}$$

These two relations can now be used to eliminate one of the phases and one of the p_i, for example, ϕ_0 and p_0. We are thus left with six independent (real) variables (ϕ_1, ϕ_2, ϕ_3), and (p_1, p_2, p_3). The p_i still have to satisfy:

$$p_i \geq 0, \tag{6.160}$$

$$p_1 + p_2 + p_3 \geq r, \tag{6.161}$$

where the second inequality comes from (6.159) and $p_0 \geq 0$. One can thus visualize the geometry of local $\mathbb{CP}^2$ as a T^3 fibration (corresponding to ϕ_1, ϕ_2, ϕ_3) over the region V in $\mathbb{R}^3$ defined by (6.160–6.161). Note that at the points where one of the $p_i = 0$, the corresponding S^1 degenerates. The toric diagram (6.2a) then represents the region V, with each plane corresponding to a locus where one of the S^1 degenerates (at the lines of intersection, we obtain higher degeneracy). It is also not difficult to see that the triangular region in the diagram corresponds to the base $\mathbb{CP}^2$.

The toric diagrams for other geometries can be obtained in a similar way. It is left as an easy exercise for the reader to show that the toric diagram for local $\mathbb{CP}^1$ has the form as shown in Fig. 6.2b.

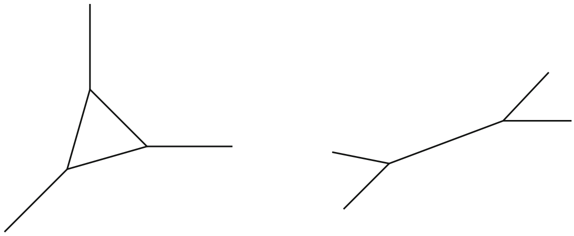


Fig. 6.2 The toric diagrams for: **a** local $\mathbb{CP}^2$, **b** local $\mathbb{CP}^1$

6.2.5 Geometric Transition and Large N Dualities

The aim of this section is to illustrate the power of large N dualities for the computation of topological string amplitudes. The basic notion, still widely conjectural, is the equivalence of pure open string theory and the closed string theory that governs the back reaction of the open string fields on the background. In [6] it was shown that open string field theory is governed by a formal Chern–Simons theory. By formal is meant that, given the Hilbert space of open string fields Ψ, the theory is specified by a skew-symmetric pairing $\langle\,,\rangle$, a differential Q, and an associative product $\star$, with respect to which Q is a derivation as

$$S_{CS} = \frac{1}{2}\langle \Psi, Q\Psi\rangle + \frac{1}{3}\langle \Psi, \Psi \star \Psi\rangle. \tag{6.162}$$

Q is really the BRST operator and $\star$ the operator product of vertex operators. The correlation functions of the string field theory, with the restrictions that the "in" and "out" states are in Q-cohomology, are the correlation functions of open string theory.

In the case of topological string theory, the corresponding string field theory is often manageable. In fact, in the case of the B-model, the Chern–Simons theory is an ordinary quantum field theory. Indeed, consider the example of a space-filling brane. This is specified by a holomorphic vector bundle E over M, the Hilbert space of string fields ψ is $\mathscr{H} = \Omega^{0,\bullet}(E^* \otimes E)$, $\star = \wedge$ and $Q = \bar{\partial}$, the Dolbeault differential on E. Finally, the action is given by:

$$S_B = \int_M \Omega \wedge \mathrm{tr}_E(\frac{1}{2}\psi \wedge \bar{\partial}\psi + \frac{1}{3}\psi \wedge \psi \wedge \psi). \tag{6.163}$$

This is the so-called holomorphic Chern–Simons theory. While in the case of the A-model, if there are no compact 2-cycles, or in the large volume limit, the open string field theory corresponding to a single brane, is the ordinary Chern–Simons theory on a Lagrangian submanifold with complex vector bundle E:

$$S_A = \int_L \mathrm{tr}_E\left(\frac{1}{2}\psi \wedge d\psi + \frac{1}{3}\psi \wedge \psi \wedge \psi\right). \tag{6.164}$$

We now turn to illustrate the very basics of these models in the simplest non-trivial examples. These are the deformed conifold for the A-model and the resolved conifold for the B-model. Let us pause to sketch the strategy. In the case of the A-model, a brane has support on a Lagrangian cycle. Its charge can be measured by linking L with a trivial 2-dimensional cycle $S = \partial C$, that is $C \cap L = 1$. The charge is then given by

$$c = \omega \cap S, \tag{6.165}$$

where ω is the Kähler flux generated by the brane. Now one might wonder whether ω can be interpreted as a chiral primary (with marginal operator as integrated descendant) for a purely closed topological A-model on a background where the volume of S is equal to c, in the spirit of the AdS/CFT correspondence. In the case of the B-model, a brane has support on a holomorphic cycle, therefore it can have varying dimension. However, the only one that has a non trivial back-reaction in terms of the closed B-model, is one that is 2-dimensional so that it can link with a trivial 3-dimensional cycle α. In that case the charge of the brane is given by

$$c = \eta \cap \alpha. \tag{6.166}$$

This open/closed duality conjecture was explored in great detail and confirmed for toric Calabi–Yau manifolds. We will sketch the results in what follows. We start with the open A-model on the deformed conifold.

6.2.5.1 Open A-Model on the Deformed Conifold

Recall that the deformed conifold is a variety in $\mathbb{C}^4$ defined via $xy - uv = \mu$. It can be viewed as T^*S^3, which can be interpreted as the phase space of a point particle on S^3. Indeed S^3 is a Lagrangian submanifold with respect to the Kähler metric induced by $\mathbb{C}^4$. This is the simplest example to test open-closed duality, as it has a single Lagrangian cycle and no compact 2-cycle. In [7] the following setting was analysed. Consider wrapping N branes on S^3 each of charge λ. The open string partition function is defined by

$$Z_{CS}(S^3) = \exp\left(- \sum_{g,n} F_{g,n} \lambda^{2g-2+n} N^n \right), \tag{6.167}$$

where $F_{g,n}$ is the generating function of connected Riemann surfaces of genus g with n boundary circles ending on S^3. If this theory were to have an A-model closed string description we would need to have a trivial S^3, so that no branes are present and a non-trivial S^2 with volume $t = N\lambda$. The natural guess is therefore the resolved conifold. In [7] it was shown that indeed, for

$$\lambda = \frac{2\pi}{k + N}, \tag{6.168}$$

where k is the level of the Chern–Simons theory, $F_g(t)$ defined by:

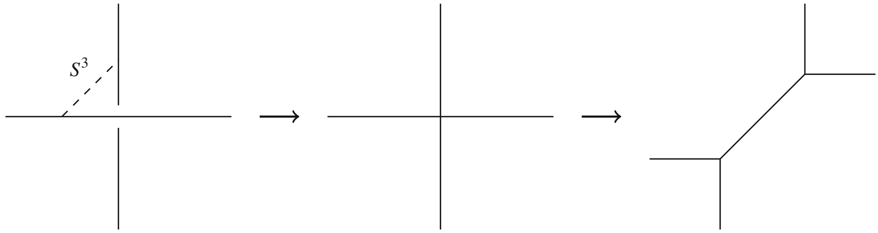


Fig. 6.3 The geometric transition from the deformed conifold (*left*) to the resolved conifold (*right*). The figure in the *center* represents the singular conifold

$$F_g(t) = \sum_n F_{g,n} t^n \tag{6.169}$$

coincides with the genus g generating function of the A-model on the resolved conifold for $N \to \infty$. This was achieved by explicit inspection of:

$$Z_{CS}(S^3) = \frac{e^{i\frac{\pi}{8}(N-1)N}}{(k+N)^{N/2}} \sqrt{\frac{k+N}{N}} \prod_{s=1}^{N-1} \left[2sin\left(\frac{s\pi}{k+N}\right)\right]^{N-s} \tag{6.170}$$

The transition from the deformed to the resolved conifold is called a geometric transition. It turns out to be very convenient to depict the transition using toric diagrams as in Fig. 6.3. The skew lines hide an S^3. Indeed on each line a distinct S^1 of T^2 degenerates, therefore a line segment connecting the two lines is a T^2 fibration equivalent to the Hopf fibration therefore equal to an S^3. This transition can be reversed to describe open-closed duality for the B-model. There we start with the B-model open theory with N branes of charge λ each, wrapped on the S^2 of the resolved conifold. As was shown in [8] the associated open string field theory reduces to a gauged gaussian Matrix Model:

$$Z = \frac{1}{\mathrm{vol}(U(N))} \int dM e^{-\frac{1}{2\lambda} trM^2} \tag{6.171}$$

where the integral is over $N \times N$ matrices M.

6.2.5.2 The Topological Vertex

The conifold transition can be employed to describe geometric transitions involving general toric threefolds. In fact the general case is essentially captured by the example involving the closed A-model on local $\mathbb{CP}^2$ [9] that we will briefly sketch. Consider the toric diagram (6.2a). The three semi-lines represent three S^2's of infinite volume. The trick is to modify (6.2a) by replacing these semi-lines with three segments representing spheres of finite moduli t_1, t_2 and t_3 respectively, and adjoining two half-lines at each end. In this way we recover the original geometry in the limit

$t_1, t_2, t_3 \to \infty$. The crucial observation is that the modified diagram is alternatively obtained by gluing together three toric diagrams, each representing a copy of the resolved conifold with respective Kähler modulus, so as to form a triangle in the centre. For each resolved conifold we can then perform a geometric transition so as to obtain a geometry with no Kähler deformations and three non-vanishing lagrangian S^3's. We then expect to recover the closed A-model on the local $\mathbb{CP}^2$ as the large N dual of the open theory on this geometry with stacks of N_1, N_2, N_3 branes wrapped on each lagrangian S^3 respectively. The situation in the present case is however, more involved than in the simple case of the conifold transition. In particular the string field theory in question is not simply the sum of three Chern Simons theories, one for each lagrangian brane. Instead there are interaction terms due to open-string instantons stretching between each pair of brane stacks. In particular the action is given by [10]:

$$S = \sum_{i=1}^{3} S_{CS}^{(i)} + \sum_{i<j}^{3} \sum_{\beta_{ij}} e^{-\omega\cdot\beta_{ij}} \mathrm{tr}\, U_{\mathscr{K}_i(\beta_{ij})} \mathrm{tr}\, U_{\mathscr{K}_j(\beta_{ij})} \tag{6.172}$$

The second summation above is over holomorphic curves β_{ij} with $\partial\beta_{ij} = C_i \sqcup C_j$ where C_i is an embedded circle (knot) in the ith lagrangian sphere. The matrix $U_{\mathscr{K}_i(\beta)}$ is the holonomy matrix or Wilson loop:

$$U_{\mathscr{K}_i(\beta_{ij})} := \mathscr{P} e^{i k_i \int_{C_i} A^{(i)}} \tag{6.173}$$

with $A^{(i)}$ the connection one-form taking values in the adjoint representation of $U(N_i)$. A careful study reveals that these instantons are (multicovers) of cylinders. Moreover two instantons starting from different branes and ending on the same brane will form a Hopf link on that brane. Let C_i denote a knot of degree 1, i.e. corresponding to a primitive instanton. The holonomy matrices corresponding to multicovers of such primitive instantons are powers of the primitive ones. The Frobenius formula allows to express the former as a linear combination of the latter holonomies in different representations R of the gauge group. The final formula for the open-string partition function is then:

$$Z = \sum_{R_1 R_2 R_3} e^{-t_1|R_1|} S_{R_1 R_2} e^{-t_2|R_2|} S_{R_2 R_3} e^{-t_3|R_3|} S_{R_3 R_1} \tag{6.174}$$

where $S_{R_1 R_2}$ is the expectation value of the Hopf link with representation R_1 on the first unknot and representation R_2 on the other, while $|R|$ denotes the number of boxes of the Young diagram associated to R.

The reasoning just used to employ the conifold transition for the case of local $\mathbb{CP}^2$ is clearly generalizable to other toric varieties, however, it comes with the unpleasant realization, that the large N limit has to be taken at the very end of the computation. A more general and conceptually sounder method is provided by the topological vertex developed in [11]. The basic idea behind it is that the toric diagrams of toric threefolds can be constructed by glueing copies of that of $\mathbb{C}^3$ viewed as the geometry underlying the open string A-model. In particular one has to keep track of three lagrangian submanifolds in $\mathbb{C}^3$ (and corresponding framing), one at every vertex of the corresponding toric diagram. Then the open string partition function on $\mathbb{C}^3$ in the large N limit, is expressed as

$$Z = \sum_{R_1,R_2,R_3} C_{R_1R_2R_3} \prod_{i=1}^{3} \mathrm{tr}_{R_i} V_i \tag{6.175}$$

where V_i denote degree one unknots in the ith lagrangian submanifold. Through this expression we have thus defined the amplitudes $C_{R_1R_2R_3}$. These are collectively called *topological vertex* and were originally derived by comparing the above expression for Z with one resulting from the geometric transition applied to one lagrangian submanifold of $\mathbb{C}^3$. This concludes the program of computing closed partition functions for general toric manifolds.

Acknowledgements I would like to thank the organizers of the summer school "Strings and Fundamental Physics" for the opportunity to give this set of lectures and for their hospitality. I am also grateful to Michael Kay and Daniel Plencner for helping with this manuscript.

References

1. Hori, K., Katz, S., Klemm, A., Pandharipande, R., Thomas, R., Vafa, C., Vakil, R., Zaslow, E.: Mirror Symmetry. American Mathematical Society, Providence (2003)
2. Ooguri, H., Oz, Y., Yin, Z.: D-branes on Calabi-Yau spaces and their mirrors. Nucl. Phys. B **477**(2), 407–430 (1996) ([hep-th/9606112])
3. Aspinwall, P.S.: D-branes on Calabi-Yau manifolds, [hep-th/0403166]
4. Sharpe, E.: Lectures on D-branes and Sheaves, [hep-th/0307245]
5. Bouchard, V.: Lectures on Complex Geometry, Calabi-Yau Manifolds and Toric Geometry, [hep-th/0702063]
6. Witten, E.: Chern-Simons gauge theory as a string theory. Prog. Math. **133**, 637–678 (1995) ([hep-th/9207094])
7. Gopakumar, R., Vafa, C.: On the gauge theory/geometry correspondence. Adv. Theor. Math. Phys. **3**, 1415–1443 (1999) ([hep-th/9811131])
8. Dijkgraaf, R., Vafa, C.: Matrix models, topological strings, and supersymmetric gauge theories. Nucl. Phys. B **644**, 3–20 (2002) ([hep-th/0206255v2])
9. Ooguri , H., Vafa, C.: Knot invariants and topological strings. Nucl. Phys. B **577**, 419–438 (2000) ([hep-th/9912123v3])
10. Mariño, M.: Chern-Simons Theory, Matrix Models, and Topological Strings. Oxford University Press, Oxford (2005)
11. Aganagic, M., Klemm, A., Mariño, M., Vafa, C.: The topological vertex. Commun. Math. Phys. **254**, 425–478 (2005) ([hep-th/0305132])

Chapter 7
Doubled Field Theory, T-Duality and Courant-Brackets

Barton Zwiebach

7.1 Introduction

These lecture notes are based on three lectures, each ninety minutes long, given by the author during the "International School on Strings and Fundamental Physics" which took place in Garching/Munich from July 25 to August 6, 2010. Aimed at graduate students, they require only a basic knowledge of string theory and give a simple introduction to double field theory. These notes were prepared by Marco Baumgartl and Nicolas Moeller.

We focus on making T-duality explicit in *field* theory Lagrangians. The 'T' in T-duality stands for 'toroidal'. T-duality is an old and still fascinating topic in string theory. We will develop some Lagrangians for T-dual field theories that are quite intriguing and may have interesting applications. The material covered here is based on joint work with Chris Hull and Olaf Hohm [1–4]. Earlier work in double field theory includes that of Siegel [5, 6] and Tseytlin [7]. These notes are informal and do not attempt to be comprehensive nor to provide complete references. They deal with the basics of the subject and do not describe any of the recent developments.

Theories implementing T-duality bring up mathematical constructions such as the Courant-brackets as well as elements of generalized geometry. There is plenty of mathematical work on these topics, much of it in the context of first-quantized string theory. In our double field theory context, Courant-brackets and ideas of generalized geometry appear in a very natural way and help construct the Lagrangians.

Courant-brackets are natural generalizations of the Lie brackets that govern general relativity. Courant-brackets should be relevant to the effective field theory of strings and we are beginning to see this. Before entering this fascinating topic

B. Zwiebach (✉)
Massachusetts Institute of Technology,
77 Massachusetts Avenue, Bldg. 6-305,
Cambridge, MA 02139, USA
e-mail: zwiebach@mit.edu

M. Baumgartl et al. (eds.), *Strings and Fundamental Physics*,
Lecture Notes in Physics 851, DOI: 10.1007/978-3-642-25947-0_7,

we will first talk about strings in toroidal backgrounds and some of their important properties.

7.2 String Theory in Toroidal Backgrounds

Consider a closed string living in a spacetime with a compactified coordinates. It is well known that upon quantization there will be momentum modes and winding modes for each compact direction. Let us denote the compact coordinates by x^a and the non-compact coordinates by x^μ, with $x^i = (x^a, x^\mu)$. The compact coordinates x^a give rise to string momentum excitations p_a. Since strings are extended objects, there are also winding quantum numbers w^a. These should in fact be associated to some new coordinates $\tilde{x}_a$. If one attempts to write down the complete field theory of closed strings in coordinate space it will include the x^a as well as the $\tilde{x}_a$. Thus, the arguments of all fields in such a theory will be doubled and we call it a double field theory (DFT). The doubled fields $\phi(x^a, \tilde{x}_a, x^\mu)$ are said to be functions of momentum *and* winding.

Since the field arguments are doubled, actions must include a suitable integration over the additional dual coordinates:

$$S = \int dx^a d\tilde{x}_a dx^\mu \mathscr{L}(x^a, \tilde{x}_a, x^\mu). \tag{7.1}$$

It is clear from the basic ideas of closed string field theory that the full string theory is described by a Lagrangian of this form. With an infinite number of fields included, however, it is very complicated. A simplification can be achieved by restricting to a subset of fields only. The natural restriction is to consider only the massless sector, which includes a dilaton ϕ, a metric g_{ij} with Riemann curvature $R(g)$, and a Kalb-Ramond field b_{ij} with field strength $H = db$.

The familiar low energy effective field theory of the bosonic closed string for these massless fields is given by

$$S_* = \int dx \sqrt{-g} e^{-2\phi} \left[R + 4(\partial\phi)^2 - \frac{1}{12} H^2 \right] + \dots \tag{7.2}$$

where the dots denote higher-derivative terms. In the light of the coordinate doubling on tori, what will this action become?

There will be quite some obstacles in finding the correct action. One leading principle which helps in its construction is generalized geometry. Generalized geometry is in fact a very mild generalization of geometry. Let us look at its gauge symmetry first. Its gauge symmetry parameters are vector fields $\xi^i \in T(M)$, which parametrize diffeomorphisms and live in the tangent bundle of the manifold, together with one-forms $\tilde{\xi}_i \in T^*(M)$, which describe gauge transformations of b_{ij} and live in the dual tangent bundle. Both are combined naturally in the setup of generalized geometry,

$$\xi + \tilde{\xi} \in T(M) \oplus T^*(M). \tag{7.3}$$

Generalized geometry does not double any coordinates. What it does achieve is to treat vectors and one-forms on an equal footing, so that it makes sense to add them to an object living in the sum of the tangent space and its dual.

In generalized geometry the Courant-bracket is the right extension of the Lie bracket. We will see that it will play a prominent role in our construction. Also, in generalized geometry and string theory the field $\mathscr{E}_{ij} = g_{ij} + b_{ij}$ appears repeatedly, and one also has the generalized metric $\mathscr{H}^{MN}$. The generalized metric is a key structure also in string theory. Up to now there were no actions written explicitly in terms of these fields.

In the following we will write down double field theories that are T-duality covariant versions of S_*. We will find that Courant-brackets, the field $\mathscr{E}_{ik}$, and the generalized metric $\mathscr{H}^{MN}$ will play an important role.

7.2.1 Sigma-Model Action

In order to construct a first-quantized action, we start with the usual sigma-model action for strings propagating in a background. It is given by

$$S = -\frac{1}{4\pi}\int_0^{2\pi} d\sigma \int_{-\infty}^{\infty} d\tau \left(\eta^{\alpha\beta}\partial_\alpha X^i \partial_\beta X^j G_{ij} + \varepsilon^{\alpha\beta}\partial_\alpha X^i \partial_\beta X^j B_{ij}\right), \tag{7.4}$$

where

$$\begin{aligned} &\eta^{\alpha\beta} = \mathrm{diag}(-1, 1), \qquad \varepsilon^{01} = -1, \qquad \partial_\alpha = (\partial_\tau, \partial_\sigma), \\ &X^i = (X^a, X^\mu) \qquad X^a \sim X^a + 2\pi, \qquad i = 0, \ldots, D-1. \end{aligned} \tag{7.5}$$

The X^a are periodic coordinates for the compact dimensions. The total number of dimensions is D. The closed string background fields G and B are $D \times D$ matrices and are taken to be constant with the following properties:

$$G_{ij} = \begin{pmatrix} \hat{G}_{\mathrm{ab}} & 0 \\ 0 & \eta_{\mu\nu} \end{pmatrix}, \qquad B_{ij} = \begin{pmatrix} \hat{B}_{ab} & 0 \\ 0 & 0 \end{pmatrix}, \qquad G^{ij}G_{jk} = \delta^i_k. \tag{7.6}$$

Both G and B can be combined into the field E defined by

$$E_{ij} = G_{ij} + B_{ij} = \begin{pmatrix} \hat{E}_{ab} & 0 \\ 0 & \eta_{\mu\nu} \end{pmatrix}, \quad \text{with} \quad \hat{E}_{ab} = \hat{G}_{ab} + \hat{B}_{ab}. \tag{7.7}$$

Exercise 1 *By using the action* (7.4), *prove that the canonical momentum* P_i *is given by*

$$2\pi P_i = G_{ij}\dot{X}^j + B_{ij}X'^j, \tag{7.8}$$

(*dot for* ∂_τ, *prime for* ∂_σ) *and that the Hamiltonian density* $\underline{H}$ *takes the form*

$$4\pi\,\underline{H} = (X', 2\pi P)\ \mathscr{H}(E)\ \begin{pmatrix} X' \\ 2\pi P \end{pmatrix}, \tag{7.9}$$

with the $2D \times 2D$ *matrix*

$$\mathscr{H}(E) = \begin{pmatrix} G - BG^{-1}B & BG^{-1} \\ -G^{-1}B & G^{-1} \end{pmatrix}. \tag{7.10}$$

The matrix $\mathscr{H}(E)$ is a $2D \times 2D$ symmetric matrix constructed out of G and B. It is called the 'generalized metric'. More precisely we will identify it with an object $\mathscr{H}^{MN}$ with $M, N = 1, \ldots, 2D$. It is convenient to write $\mathscr{H}$ and its inverse in product form as

$$\begin{aligned} \mathscr{H} &= \begin{pmatrix} G - BG^{-1}B & BG^{-1} \\ -G^{-1}B & G^{-1} \end{pmatrix} = \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & G^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -B & 1 \end{pmatrix}, \\ \mathscr{H}^{-1} &= \begin{pmatrix} G^{-1} & -G^{-1}B \\ BG^{-1} & G - BG^{-1}B \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ B & 1 \end{pmatrix} \begin{pmatrix} G^{-1} & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} 1 & -B \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{7.11}$$

$\mathscr{H}$ is non-degenerate because each of its factors is non-degenerate. It is useful to define another metric η with constant off-diagonal entries

$$\eta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{7.12}$$

With the metric η we are able to relate $\mathscr{H}$ and its inverse, so that, as you can check,

$$\eta \mathscr{H} \eta = \mathscr{H}^{-1}. \tag{7.13}$$

Such a constraint comes about because the generalized metric is a $2D \times 2D$ matrix symmetric matrix constructed from a single $D \times D$ matrix $E = G + B$. Thus is has to be constrained. We can view the parameterization of $\mathscr{H}$ in terms of G and B as a natural and general solution of the constraint.

Let us put indices on $\mathscr{H}$ like on a metric, so that we can identify

$$\begin{aligned} \mathscr{H} &\leftrightarrow \mathscr{H}^{MN}, \\ \mathscr{H}^{-1} &\leftrightarrow \mathscr{H}_{MN}. \end{aligned} \tag{7.14}$$

Then Eq. 7.13 becomes

$$\begin{aligned} \eta_{PM} \mathscr{H}^{MN} \eta_{NQ} &= \mathscr{H}_{PQ} \\ \mathscr{H}^{MN} \eta_{MP} \eta_{NQ} &= \mathscr{H}_{PQ}, \end{aligned} \tag{7.15}$$

so that lowering the indices of $\mathscr{H}$ with the metric η gives us the inverse $\mathscr{H}^{-1}$! The capitalized indices *M,N* run over $2D$ values, and will be called $O(D, D)$ indices.

7.2.2 Oscillator Expansions

The string coordinate $X^i = x^i + w^i\sigma + \tau G^{ij} p_j + \ldots$ has an expansion in terms of momenta, winding, and oscillators. The zero modes α_0 and $\tilde{\alpha}_0$ are given by

$$\begin{aligned} \alpha_0^i &= \frac{1}{\sqrt{2}} G^{ij}(p_j - E_{jk} w^k), \\ \tilde{\alpha}_0^i &= \frac{1}{\sqrt{2}} G^{ij}(p_j + E_{kj} w^k). \end{aligned} \tag{7.16}$$

Written with $p_i = \frac{1}{i}\frac{\partial}{\partial x^i}$ and $w^i = \frac{1}{i}\frac{\partial}{\partial \tilde{x}_i}$

$$\begin{aligned} \alpha_{0i} &= -\frac{i}{\sqrt{2}}\left(\frac{\partial}{\partial x^i} - E_{ik}\frac{\partial}{\partial \tilde{x}_k}\right) \equiv -\frac{i}{\sqrt{2}} D_i, \\ \tilde{\alpha}_{0i} &= -\frac{i}{\sqrt{2}}\left(\frac{\partial}{\partial x^i} + E_{ki}\frac{\partial}{\partial \tilde{x}_k}\right) \equiv -\frac{i}{\sqrt{2}} \bar{D}_i. \end{aligned} \tag{7.17}$$

We have thus defined derivatives that will play an important role later

$$\begin{aligned} D_i &= \partial_i - E_{ik}\tilde{\partial}^k, \qquad D^i \equiv G^{ij} D_j, \\ \bar{D}_i &= \partial_i + E_{ki}\tilde{\partial}^k, \qquad \bar{D}^i \equiv G^{ij} \bar{D}_j. \end{aligned} \tag{7.18}$$

It turns out that the Virasoro operators with zero mode number are given by

$$\begin{aligned} L_0 &= \frac{1}{2}\alpha_0^i G_{ij} \alpha_0^j + N - 1, \\ \bar{L}_0 &= \frac{1}{2}\bar{\alpha}_0^i G_{ij} \bar{\alpha}_0^j + \bar{N} - 1, \end{aligned} \tag{7.19}$$

where N and $\bar{N}$ are number operators counting the excitations. There is a constraint in closed string theory which matches the levels of the right and the left moving excitations in any state. It requires that $L_0 - \bar{L}_0 = 0$. Using the derivatives defined above we can express $L_0 - \bar{L}_0$ as:

$$L_0 - \bar{L}_0 = N - \bar{N} - \frac{1}{4}(D^i G_{ij} D^j - \bar{D}^i G_{ij} \bar{D}^j) = N - \bar{N} - \frac{1}{4}(D^i D_i - \bar{D}^i \bar{D}_i). \tag{7.20}$$

Exercise 2 *Show that*

$$\frac{1}{2}(D^i D_i - \bar{D}^i \bar{D}_i) = -2\partial_i \tilde{\partial}^i. \tag{7.21}$$

The constraint $L_0 - \bar{L}_0 = 0$ can now be expressed as a constraint on the number operators in the following way:

$$N - \bar{N} = -\partial_i \tilde{\partial}^i \equiv -\partial \cdot \tilde{\partial}. \tag{7.22}$$

The familiar massless fields with $N = \bar{N} = 0$ have the form

$$\sum_{p,w} e_{ij}(p,w)\alpha^i_{-1}\bar{\alpha}^j_{-1}c_1\bar{c}_1|p,w\rangle,$$
$$\sum_{p,w} d(p,w)(c_1c_{-1} - \bar{c}_1\bar{c}_{-1})|p,w\rangle, \tag{7.23}$$

with momentum space wavefunctions $e_{ij}(p,w)$ and $d(p,w)$. Here the matter and ghost oscillators act on a vacuum $|p,w\rangle$ with momentum p and winding w. On account of (7.22) we must require that the fields $e_{ij}(x,\tilde{x})$ and $d(x,\tilde{x})$ satisfy the constraint

$$\partial \cdot \tilde{\partial}\; e_{ij}(x,\tilde{x}) = \partial \cdot \tilde{\partial}\; d(x,\tilde{x}) = 0. \tag{7.24}$$

This constraint is a very important ingredient which any string theory and any double field theory has to satisfy.

7.2.3 $O(D, D)$ Transformations

It is important to understand the invariance of the physics under background transformations. In particular, $O(D,D)$ transformations play a prominent role in our case. In order to study them we start with the Hamiltonian, which can be constructed from the Hamiltonian density $\underline{H}$ in (7.9). One can show that

$$H = \int_0^{2\pi} d\sigma\, \underline{H} = \frac{1}{2} Z^t \mathscr{H}(E) Z + N + \bar{N} + \ldots \tag{7.25}$$

where the dots indicate terms irrelevant to the discussion and

$$Z = \begin{pmatrix} w^i \\ p_i \end{pmatrix},$$

is a $2D$ column vector consisting of integer winding and momentum quantum numbers. The $L_0 - \bar{L}_0 = 0$ condition (7.22) on the spectrum gives $N - \bar{N} = p_i w^i$, or equivalently,

$$N - \bar{N} = \frac{1}{2} Z^t \eta Z, \tag{7.26}$$

where η is the matrix defined in (7.12). Consider now a reshuffling of the quantum numbers

$$Z = h^t Z',$$

with some $2D \times 2D$ invertible matrix h with integer entries (h^{-1} should also have integer entries). Under such a transformation the physics should not change, and in particular the constraint (7.26) should be unchanged. For this it is then necessary that

$$Z'^t \eta Z' = Z^t \eta Z = Z'^t h \eta h^t Z', \tag{7.27}$$

which requires

$$h \eta h^t = \eta. \tag{7.28}$$

Exercise 3 *Show that* (7.28) *implies*

$$h^t \eta h = \eta, \tag{7.29}$$

The h matrices generate the $O(D, D)$ group. We write

$$h = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(D, D), \tag{7.30}$$

where a,b,c and d are $D \times D$-matrices. The conditions on a, b, c, and d following from (7.29) are

$$a^t c + c^t a = b^t d + d^t b = 0, \quad a^t d + c^t b = 1. \tag{7.31}$$

The conditions that follow from (7.28) are not independent but they are useful to have

$$ab^t + ba^t = cd^t + dc^t = 0, \quad ad^t + bc^t = 1. \tag{7.32}$$

Exercise 4 *Show that*

$$h^{-1} = \begin{pmatrix} d^t & b^t \\ c^t & a^t \end{pmatrix}. \tag{7.33}$$

More is still needed in order to ensure the invariance of the spectrum. The energy, or Hamiltonian must not change. This requires a change of the background field E: the shuffled quantum numbers are associated to a background field E'. From (7.25) we demand

$$Z^t \mathcal{H}(E) Z = Z'^t \mathcal{H}(E') Z'. \tag{7.34}$$

We thus have

$$Z'^t h \mathcal{H}(E) h^t Z' = Z'^t \mathcal{H}(E') Z'. \tag{7.35}$$

We therefore learn that

$$\mathcal{H}(E') = h \mathcal{H}(E) h^t. \tag{7.36}$$

Using the indices introduced in (7.14) we associate with h the transformation of coordinates

$$X'^M = h^M{}_N X^N, \quad X \equiv \begin{pmatrix} \tilde{x} \\ x \end{pmatrix}. \tag{7.37}$$

and then (7.36) becomes

$$\mathscr{H}^{MN}(E') = h^M{}_P h^N{}_Q \mathscr{H}^{PQ}(E). \tag{7.38}$$

Given that $\mathscr{H}$ is a rather complicated function of the metric G and the field B associated with $E = G + B$, it is not obvious that there is a transformation of E that induces the covariant transformation (7.36) (or (7.38)) of $\mathscr{H}$. The transformation of E in fact exists and is given by:

$$E' = h(E) = (aE + b)(cE + d)^{-1} \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} E. \tag{7.39}$$

This is actually a well known transformation which appears often in string theory. It looks like a modular transformation. The fields G and B in E have much more complicated transformation laws. This is an indication that E is a good variable to formulate our theories.

Exercise 5 *Show that*

$$E'^t = \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} E^t. \tag{7.40}$$

In order to show that (7.36) holds, we first consider the possibility that E is created from the identity background I by the action of h. Is it possible to create any such background from the identity? If so, then this would be a very convenient insight. Let us assume it is true for the moment and assign a transformation h_E to any E, so that

$$E = h_E(I). \tag{7.41}$$

To see that $h_E \in O(D, D)$ really does exist we re-write the field G in $E = G + B$. Since G is symmetric it can be written as $G = AA^t$, where A appears like a vielbein. Using now A and B in the explicit expression for h_E we find that

$$h_E = \begin{pmatrix} A & B(A^t)^{-1} \\ 0 & (A^t)^{-1} \end{pmatrix}. \tag{7.42}$$

It is easy to check that h_E is indeed an element of $O(D, D)$. In order to see that it satisfies (7.41) we compute

$$h_E(I) = (AI + B(A^t)^{-1})(0 \cdot I + (A^t)^{-1})^{-1} = (A + B(A^t)^{-1})A^t = AA^t + B = E. \tag{7.43}$$

This indeed shows that any background E can be created from the identity background by the transformation that we have explicitly constructed.

The transformation h_E is ambiguous since it is always possible to replace h_E by $h_E \cdot g$ where $g(I) = I$. These g are elements of $O(D, D)$, and in fact they form a subgroup.

Exercise 6 *Show that the elements g that satisfy* $g(I) = I$ *form an* $O(D) \times O(D)$ *subgroup of* $O(D, D)$ *and* $g^t g = g g^t = I$.

With these preparations we can now focus again on (7.36) and show that $\mathscr{H}$ transforms in the right way. For the construction of h_E we have split the metric G into a product of A and A^t, so that only A entered in h_E. In order to find a matrix with G it is natural to consider the product $h_E h_E^t$ which does not have the ambiguity of exercise 6. This can be calculated in a straightforward manner:

$$h_E h_E^t = \begin{pmatrix} A & B(A^t)^{-1} \\ 0 & (A^t)^{-1} \end{pmatrix} \begin{pmatrix} A^t & 0 \\ -A^{-1}B & A^{-1} \end{pmatrix} = \begin{pmatrix} G - BG^{-1}B & BG^{-1} \\ -G^{-1}B & G^{-1} \end{pmatrix} = \mathscr{H}(E). \tag{7.44}$$

Suppose now E' is a transformation of E by h, i.e. $E' = h(E) = hh_E(I)$. We also have $E' = h_{E'}(I)$. We thus see that $h_{E'} = hh_E g$, up to the ambiguous $O(D, D)$ subgroup formed by g. Now we can put all this together to compute

$$\mathscr{H}(E') = h_{E'} h_{E'}^t = hh_E g (hh_E g)^t = hh_E h_E^t h^t = h\mathscr{H}(E)h^t. \tag{7.45}$$

This proves (7.36).

Our aim is to show that it is natural to replace the standard notation in string theory based on G and B by E and $\mathscr{H}$, and in fact we will later re-write the Einstein action completely in terms of $\mathscr{H}$. In order to arrive there we still need a little more formalism.

First we need to understand how G and G' are related. This relation is not immediately visible. We claim that

$$(d + cE)^t G'(d + cE) = G. \tag{7.46}$$

This expression involves E but neither a nor b (from h) enter. It looks like a transformation law for tensors, but it is in fact a bit more complicated, since we have E-dependent matrices. In the end this will lead to a new kind of indices which are characterized by the fact that they transform like (7.46).

The metric G has the peculiar property that in addition it also satisfies

$$(d - cE^t)^t G'(d - cE^t) = G. \tag{7.47}$$

This has some deeper meaning, as we will see.

Exercise 7 *Prove that*

$$\begin{aligned} (d + cE)^t G'(d + cE) &= G, \\ (d - cE^t)^t G'(d - cE^t) &= G. \end{aligned} \tag{7.48}$$

Hint: Write $G' = \frac{1}{2}(E' + E'^t)$ *and use* (7.39) *for the first line. Write* $G' = \frac{1}{2}((E')^t + (E'^t)^t)$ *and use* (7.40) *for the second line.*

In order to sharpen notation let us introduce the matrices

$$\begin{aligned} M &\equiv (d - cE^t)^t, \\ \bar{M} &\equiv (d + cE)^t. \end{aligned} \tag{7.49}$$

With this abbreviation (7.48) becomes

$$\begin{aligned} G &= \bar{M} G' \bar{M}^t, \\ G &= M G' M^t. \end{aligned} \tag{7.50}$$

It is instructive to write these equations in index notation. These are in fact examples of $O(D, D)$ "tensors", which transform in the following way:

$$\begin{aligned} G_{\bar{i}\bar{j}} &= \bar{M}_{\bar{i}}{}^{\bar{p}} \bar{M}_{\bar{j}}{}^{\bar{q}} G'_{\bar{p}\bar{q}}, \\ G_{ij} &= M_i{}^p M_j{}^q G'_{pq}. \end{aligned} \tag{7.51}$$

Note that we have used two kinds of indices for the same object G. It is possible to describe G either with barred indices $G_{\bar{i}\bar{j}}$ or with unbarred indices G_{ij}. Each type of indices comes with a different transformation law, but still they describe the same transformation.

Previously we found indices M, N that are used for $O(D, D)$ tensors. Now we found other indices for which $O(D, D)$ transformations are generated by matrices M and $\bar{M}$. Thus we want to understand how these two index manipulations are related to each other. Consider some object with components

$$\Theta^M = \begin{pmatrix} \tilde{\theta}_i \\ \theta^i \end{pmatrix}.$$

We call such an object a "fundamental of $O(D, D)$" if $\Theta' = h\Theta$, or in components

$$\begin{pmatrix} \tilde{\theta}' \\ \theta' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \tilde{\theta} \\ \theta \end{pmatrix}, \tag{7.52}$$

and we say it transforms in the fundamental representation of $O(D, D)$. Now let us define two more objects

$$\begin{aligned} Y_i &\equiv -\tilde{\theta}_i + E_{ij}\theta^j, \\ \bar{Y}_i &\equiv \tilde{\theta}_i + E_{ji}\theta^j, \end{aligned} \tag{7.53}$$

using the θ's and the E. These objects will not transform just with h, since they now depend on E. Still, they have a simple transformation law, involving the M's:

$$\begin{aligned} Y_i &= M_i{}^j Y'_j, \\ \bar{Y}_{\bar{i}} &= \bar{M}_{\bar{i}}^{\bar{j}} \bar{Y}'_{\bar{j}}. \end{aligned} \tag{7.54}$$

Thus the above construction tells us how to move from an object Θ which transforms with h to an object Y which transforms with M.

Exercise 8 *Prove the first line of* (7.54). *For this use, and prove, the identity*

$$b^t - Ea^t = -ME'. \tag{7.55}$$

This has a useful application. Consider a fundamental object

$$X^M = \begin{pmatrix} \tilde{x}_i \\ x^i \end{pmatrix}.$$

The associated partial derivative is

$$\partial_M = \begin{pmatrix} \tilde{\partial}^i \\ \partial_i \end{pmatrix} \to \partial^M \equiv \eta^{MN}\partial_N = \begin{pmatrix} \partial_i \\ \tilde{\partial}^i \end{pmatrix}. \tag{7.56}$$

The derivative ∂^M is also in the fundamental representation. From this it is now possible to calculate

$$\partial^M \partial_M = 2\partial_i \tilde{\partial}^i = 0. \tag{7.57}$$

This is recognized as the constraint (7.22). In the same way as we have constructed the objects Y and $\bar{Y}$ in (7.53) above, we can construct derivatives transforming under the action of M. When we do that we find that the natural objects to write are

$$\begin{aligned} -\partial_i + E_{ij}\tilde{\partial}^j &= -D_i, \\ \partial_i + E_{ji}\tilde{\partial}^j &= \bar{D}_i, \end{aligned} \tag{7.58}$$

which are exactly the derivatives in (7.18). So we see, those derivatives we find in string theory are in fact $O(D, D)$-derivatives and transform covariantly under $O(D, D)$:

$$\begin{aligned} D_i &= M_i{}^j D'_j, \\ \bar{D}_i &= \bar{M}_i{}^j \bar{D}'_j. \end{aligned} \tag{7.59}$$

The last object whose transformation properties we have to understand better is that for the variation δE of the background field. We know already that $E' = h(E)$, which is a complicated expression when written out. While E does not transform as a tensor, its variation does. We find

$$\begin{aligned} E' + \delta E' &= h(E + \delta E) \\ &= (a(E + \delta E) + b)(c(E + \delta E) + d)^{-1} \\ &= (aE + b + a\delta E)(cE + d + c\delta E)^{-1} \\ &= E' + a\delta E(cE + d)^{-1} - E'c\delta E(cE + d)^{-1}, \end{aligned} \tag{7.60}$$

where we used $(A+\varepsilon)^{-1} = A^{-1} - A^{-1}\varepsilon A^{-1} + O(\varepsilon^2)$ in the last step. From this we get

$$\delta E' = (a - E'c)\delta E(cE+d)^{-1} = (a - E'c)\delta E(\bar{M}^t)^{-1}. \tag{7.61}$$

The last hurdle is a bit of manipulation.

Exercise 9 *Prove that $a - E'c = M^{-1}$. For this check that $M(a - E'c) = 1$ by explicit multiplication, using the results of Exercise* 8.

From (7.61) and the result of the above exercise one reads off the transformation law

$$\delta E = M\delta E'\bar{M}^t. \tag{7.62}$$

We see that E has one unbarred index and one barred index:

$$\delta E_{i\bar{j}} = M_i{}^p \bar{M}_{\bar{j}}{}^{\bar{q}} \delta E'_{p\bar{q}}. \tag{7.63}$$

We have set up a consistent formalism and have understood the transformation laws of the fundamental objects in our theory. We can use this in order to construct actions.

7.3 Double Field Theory Actions

For the construction of actions using the previously developed formalism we start with a background field $E_{i\bar{j}}$ and small fluctuations $e_{i\bar{j}}(x,\tilde{x})$. This should be thought of as a background configuration which contains a gravitational background as well as a background Kalb-Ramond field. In addition we include a dilaton $d(x,\tilde{x})$.

7.3.1 The Quadratic Action

First we focus on the quadratic part of the action, given by

$$S^{(2)} = \int dx d\tilde{x} \left[\frac{1}{4}e^{i\bar{j}}\Box e_{i\bar{j}} + \frac{1}{4}(\bar{D}^{\bar{j}}e_{i\bar{j}})^2 + \frac{1}{4}(D^i e_{i\bar{j}})^2 - 2dD^i\bar{D}^{\bar{j}}e_{i\bar{j}} - 4d\Box d\right], \tag{7.64}$$

where indices are raised by the background metric G^{ij} (or $G^{\bar{i}\bar{j}}$) and the box operator is given by $\Box = D^i D_i = \bar{D}^{\bar{i}}\bar{D}_{\bar{i}}$ with constraint $D^2 - \bar{D}^2 = 0$. This constraint is equivalent to $\partial\cdot\tilde{\partial} = 0$ and must be satisfied by all fields and gauge parameters.

Under $O(D, D)$-transformations the objects under the integral will transform with M or $\bar{M}$. Note that M and $\bar{M}$ depend on the background field E and not on the fluctuations $e_{i\bar{j}}$. This implies that derivatives will not act on M or $\bar{M}$. So this action is manifestly $O(D, D)$-invariant.

This action must have gauge symmetries, which must include those found in general relativity. In fact gauge symmetry fixes the relative values of the coefficients of the terms in the action.

Exercise 10 *Prove that the following are gauge invariances of* $S^{(2)}$:

$$\begin{aligned} \delta e_{i\bar{j}} &= \bar{D}_{\bar{j}}\lambda_i, \qquad \delta e_{i\bar{j}} = D_i\bar{\lambda}_{\bar{j}}, \\ \delta d &= -\frac{1}{4}D^i\lambda_i\ , \qquad \delta d = -\frac{1}{4}\bar{D}^{\bar{i}}\bar{\lambda}_{\bar{i}}. \end{aligned} \tag{7.65}$$

In order to get a better feeling for this action we write it out more explicitly, simplifying it by setting the background Kalb–Ramond field B_{ij} to zero, keeping only the fluctuations $e_{i\bar{j}} = h_{ij} + b_{ij}$ around the metric G_{ij}. The action becomes

$$\begin{aligned} S^{(2)} = \int dx d\tilde{x}\Bigg[&\frac{1}{4}h^{ij}\partial^2 h_{ij} + \frac{1}{2}(\partial^i h_{ij})^2 - 2d\partial^i\partial^j h_{ij} - 4d\partial^2 d \\ &+ \frac{1}{4}h^{ij}\tilde{\partial}^2 h_{ij} + \frac{1}{2}(\tilde{\partial}h_{ij})^2 + 2d\tilde{\partial}^i\tilde{\partial}^j h_{ij} - 4d\tilde{\partial}^2 d \\ &+ \frac{1}{4}b^{ij}\partial^2 b_{ij} + \frac{1}{2}(\partial^j b_{ij})^2 \\ &+ \frac{1}{4}b^{ij}\tilde{\partial}^2 b_{ij} + \frac{1}{2}(\tilde{\partial}^j b_{ij})^2 \\ &+ (\partial_k h^{ik})(\tilde{\partial}^j b_{ij}) + (\tilde{\partial}^k h_{ik})(\partial_j b^{ij}) - 4d\partial^i\tilde{\partial}^j b_{ij}\Bigg]. \end{aligned} \tag{7.66}$$

The first line contains the graviton and dilaton in the same way as one would get from the standard action. The second line is almost identical to the first line but contains dual derivatives. This is to be expected since the whole action should be $O(D, D)$-invariant. The third line contains the contributions of the Kalb–Ramond field strength, while the fourth line again complements it with terms involving dual derivatives. The last line contains terms with mixed derivatives. These terms have no counterpart in conventional field theory actions.

The symmetries of this action are conveniently described in terms of redefined gauge parameters

$$\varepsilon_i = \frac{1}{2}(\lambda_i + \bar{\lambda}_i) \qquad \tilde{\varepsilon}_i = \frac{1}{2}(\lambda_i - \bar{\lambda}_i). \tag{7.67}$$

Using these the gauge symmetries (7.65) are found to be

$$\begin{aligned} \delta h_{ij} &= \partial_i \varepsilon_j + \partial_j \varepsilon_i & \tilde{\delta} h_{ij} &= \tilde{\partial}_i \tilde{\varepsilon}_j + \tilde{\partial}_j \tilde{\varepsilon}_i \\ \delta b_{ij} &= -(\tilde{\partial}_i \varepsilon_j - \tilde{\partial}_j \varepsilon_i) & \tilde{\delta} b_{ij} &= -(\partial_i \tilde{\varepsilon}_j - \partial_j \tilde{\varepsilon}_i) \\ \delta d &= -\frac{1}{2} \partial \cdot \varepsilon & \tilde{\delta} d &= \frac{1}{2} \tilde{\partial} \cdot \tilde{\varepsilon}. \end{aligned} \tag{7.68}$$

To appreciate the novel aspects of the above, consider the familiar linearized gauge symmetries of the low energy (non-double) action (7.2):

$$\begin{aligned} \delta h_{ij} &= \partial_i \varepsilon_j + \partial_j \varepsilon_i & \tilde{\delta} h_{ij} &= 0, \\ \delta b_{ij} &= 0, & \tilde{\delta} b_{ij} &= -(\partial_i \tilde{\varepsilon}_j - \partial_j \tilde{\varepsilon}_i) \\ \delta d &= -\frac{1}{2} \partial \cdot \varepsilon & \tilde{\delta} d &= 0. \end{aligned} \tag{7.69}$$

In (7.69) we have two columns. The left one corresponds to the symmetry of diffeomorphisms, with gauge parameter ε_i. The gravity fluctuation transforms, the b field does not, and the dilaton d transforms as a scalar density. The conventional scalar dilaton Φ is given by $\Phi \equiv d + \frac{1}{4} h^i{}_i$ and is gauge invariant. In the double field theory case (7.68) the b field transforms using the tilde derivatives to form the required antisymmetric right-hand side.

In the second column of (7.69) the gauge parameter $\tilde{\varepsilon}_i$ generates the b field transformations. No other field transforms under it. But in the corresponding column of (7.68) we see h transforming under what we could call dual diffeomorphisms and d transfoming as a dual density. The combination $\tilde{\Phi} \equiv d - \frac{1}{4} h^i{}_i$ is invariant under the $\tilde{\varepsilon}$ symmetry. Since $\tilde{\Phi}$ is not invariant under the ε transformation nor is Φ invariant under $\tilde{\varepsilon}$ transformations there is no dilaton that is a scalar under *both* diffeomorphisms and dual diffeomorphisms.

7.3.2 The Cubic Action

For going beyond the free theory cubic terms should be added to the action. Indeed there are cubic terms which are $O(D, D)$-invariant and can be added consistently to the action. This results in fact in a nonlinear extension of the gauge invariance.

For simplicity we focus only on a few possible terms in the cubic part $S^{(3)}$ of the action and refer to the literature for complete details:

$$\begin{aligned} S^{(3)} = \int dx d\tilde{x} \frac{1}{4} e_{ij} \Big((D^i e_{kl})(\bar{D}^j e^{kl}) - D^i e_{kl} \bar{D}^l e^{kj} - D^k e^{il} \bar{D}^j e_{kl} \Big) \\ + de^2 \text{ terms} + d^2 e \text{ terms} + d^3 \text{ terms}. \end{aligned} \tag{7.70}$$

The nonlinear extension of the gauge symmetry can be seen from the variation of e, which is given by

$$\delta_\lambda e_{i\bar{j}} = \bar{D}_{\bar{j}} \lambda_i + \frac{1}{2} \Big[(D_i \lambda^k) e_{k\bar{j}} - (D^k \lambda_i) e_{k\bar{l}} + \lambda_k D^k e_{i\bar{j}} \Big]. \tag{7.71}$$

While the construction up to cubic order has been completed, higher orders may be very nontrivial. It may even happen that higher orders do not exist as long as one restricts oneself to a formulation involving only the massless fields e_{ij} and d.

We have stressed that all fields and gauge parameters must satisfy the constraint that they are annihilated by $\partial \cdot \tilde{\partial}$. This was enough for the quadratic action and in fact for the cubic action. But even for the gauge transformations (7.71) there is an important subtlety. It is not true that $\partial \cdot \tilde{\partial}$ annihilates a product of two fields, even if each field is annihilated individually. Thus the terms in brackets in (7.71) do not satisfy the contraint; they should since they represent a variation of the constrained field $e_{i\bar{j}}$. Thus one must include for those terms in brackets a projector to the space of functions that satisfy the constraint. Such projectors are not needed in the cubic action (the integration does the projection automatically) but they complicate matters considerably when trying to construct the quartic terms of the action.

To be able to proceed more simply we impose a stronger constraint. We simply demand that the operator $\partial \cdot \tilde{\partial}$ annihilates all fields *and* all products of fields.

Let $A_i(x, \tilde{x})$ be fields or gauge parameters which are annihilated by $\partial_M \partial^M$. When we require now that all products $A_i A_j$ be also killed by $\partial_M \partial^M$ this leads to the condition

$$\partial_M A_i \partial^M A_j = 0, \quad \forall i, j. \tag{7.72}$$

We may call this the "strong" $O(D, D)$ constraint.

In fact this is a *very* strong constraint, and while it makes the calculations easier it makes us lose much physics. It turns out that this strong constraint makes the theory independent of the dual coordinates in the following sense:

Theorem 1 *For a set of fields $A_i(x, \tilde{x})$ that satisfies (7.72) there is a duality frame $(\tilde{x}'_i, x'^i)$ in which the fields do not depend on $\tilde{x}'_i$.*

Even if it is always possible to find such a frame, we need not specify it explicitly, i.e. we need not break $O(D, D)$ invariance. The constraint (7.72) is indeed $O(D, D)$ invariant. Hence we are in a situation where we can formulate a theory using dual coordinates in the action, keeping the full $O(D, D)$ invariance, while physically only half of the coordinates matter.

7.4 Courant Brackets

In a theory with a metric $g_{ij}(x)$ and a Kalb–Ramond field $b_{ij}(x)$ the diffeomorphisms are generated by vector fields $V^i(x)$ and Kalb–Ramond gauge transformations are generated by one-forms $\xi_i(x)$. These are formally added and thus written as $V + \xi \in T(M) \oplus T^*(M)$, where $V \in T(M)$ and $\xi \in T^*(M)$ are elements of the tangent bundle and the cotangent bundle, respectively. We can formulate the gauge transformations in a geometric language

$$\begin{aligned}\delta_{V+\xi} g &= \mathscr{L}_V g, \\ \delta_{V+\xi} b &= \mathscr{L}_V b + d\xi,\end{aligned} \tag{7.73}$$

where $\mathscr{L}_V$ is the Lie derivative along the vector field V. Recall that when acting on forms the Lie derivative is

$$\mathscr{L}_V = \iota_V d + d\,\iota_V, \tag{7.74}$$

where ι_V is contraction with V. It follows that the Lie derivative and the exterior derivative commute,

$$\mathscr{L}_V d = d\mathscr{L}_V. \tag{7.75}$$

Acting on the metric the Lie derivatives gives

$$(\mathscr{L}_V g)_{ij} = (\partial_i V^k) g_{kj} + \partial_j V^k g_{ik} + V^k \partial_k g_{ij}. \tag{7.76}$$

Lie derivatives satisfy interesting algebraic relations:

$$\begin{aligned}[][\mathscr{L}_X, \mathscr{L}_Y] &= \mathscr{L}_{[X,Y]}, \\ [\mathscr{L}_X, \iota_Y] &= \iota_{[X,Y]}.\end{aligned} \tag{7.77}$$

The left hand sides are commutators of operators and on the right-hand side we find brackets of vector fields, defined as $[V_1, V_2]^k = V_1^p \partial_p V_2^k - (1 \leftrightarrow 2)$.

7.4.1 Motivating the Courant Bracket

Suppose one has a theory of a metric and an antisymmetric tensor field and one has derived the transformation laws (7.73), how can one determine the gauge algebra? First we compute the algebra of gauge transformations on the metric g by evaluating the bracket

$$[\delta_{V_2+\xi_2}, \delta_{V_1+\xi_1}] g = \mathscr{L}_{V_1} \mathscr{L}_{V_2} g - (1 \leftrightarrow 2) = \mathscr{L}_{[V_1, V_2]} g. \tag{7.78}$$

On the Kalb–Ramond field b the computation is a little less trivial. We find

$$[\delta_{V_2+\xi_2}, \delta_{V_1+\xi_1}] b = \mathscr{L}_{V_1}(\mathscr{L}_{V_2} b + d\xi_2) - (1 \leftrightarrow 2) = \mathscr{L}_{[V_1, V_2]} b + d(\mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1). \tag{7.79}$$

When we compare this with (7.73) we conclude that acting on the fields

$$[\delta_{V_2+\xi_2}, \delta_{V_1+\xi_1}] = \delta_{[V_1, V_2] + \mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1}. \tag{7.80}$$

This last expression defines a "bracket" on $T(M) \oplus T^*(M)$:

$$[V_1 + \xi_1, V_2 + \xi_2] = [V_1, V_2] + \mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1. \tag{7.81}$$

The first term on the right-hand side is a vector field, the last two give a one-form. One may ask now if this bracket is a Lie bracket. It is because it is antisymmetric and the Jacobi identity is satisfied (as a calculation shows).

There is, however, an ambiguity in the one-form because this one-form appears in the gauge transformation acted by the exterior derivative. Indeed,

$$\delta_{V+\xi} b = \mathscr{L}_V b + d\xi = \mathscr{L}_{V+(\xi+d\sigma)} b.$$

Thus the one-form ξ is ambiguous up to an exact term $d\sigma$. This ambiguity also is present in (7.81). To see this we calculate the exterior derivative of the form on the right-hand side

$$d(\mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1) = d(\underline{d\,\iota_{V_1}\xi_2} + \iota_{V_1} d\xi_2 - (1 \leftrightarrow 2)) \tag{7.82}$$

The underlined term is killed by the action of d, so without loss of generality we may change the coefficient in front of it. We will do so by replacing it with $1 - \frac{\beta}{2}$:

$$d(\mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1) = d\left(\mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1 - \frac{1}{2}\beta d(\iota_{V_1}\xi_2 - \iota_{V_2}\xi_1)\right). \tag{7.83}$$

This ambiguity should be reflected in our definition of the bracket. So we replace (7.81) by

$$[V_1 + \xi_1, V_2 + \xi_2]_\beta = [V_1, V_2] + \mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1 - \frac{1}{2}\beta d(\iota_{V_1}\xi_2 - \iota_{V_2}\xi_1). \tag{7.84}$$

One complication with this bracket is that is does not satisfy a Jacobi identity as long β does not vanish. Does it make sense to consider brackets with $\beta \neq 0$ at all ? Yes it does! One can show that, with $Z_i = V_i + \xi_i,\ i = 1, 2, 3$, the "Jacobiator" takes the form

$$[Z_1, [Z_2, Z_3]] + \text{cyclic} = dN(Z_1, Z_2, Z_3). \tag{7.85}$$

The right hand side is not zero but an exact 1-form. Since exact one-forms do not generate gauge transformations, the failure of the Jacobi identity does not cause inconsistency.

This bracket is not a new invention, but it has been considered before by T. Courant in 1990. He had reasons to fix $\beta = 1$ and therefore defined a bracket called the *Courant bracket* as

$$[V_1 + \xi_1, V_2 + \xi_2]_{\beta=1} = [V_1, V_2] + \mathscr{L}_{V_1}\xi_2 - \mathscr{L}_{V_2}\xi_1 - \frac{1}{2}d(\iota_{V_1}\xi_2 - \iota_{V_2}\xi_1). \tag{7.86}$$

In fact for $\beta = 1$ there is an extra automorphism of the bracket, called B-transformation. This is what makes it interesting from a mathematical point of view. Given a closed 2-form B with $dB = 0$, the B-transformation acts on a pair (X, ξ) of gauge parameters as follows,

$$B\text{-transformation: } X + \xi \mapsto X + (\xi + \iota_X B). \tag{7.87}$$

So this map has the effect that it changes the 1-form. If B-transformations are an automorphism of the bracket one must have:

$$[X + \xi + \iota_X B, Y + \eta + \iota_Y B] = [X + \xi, Y + \eta] + \iota_{[X,Y]} B. \tag{7.88}$$

Exercise 11 *Show that the existence of this automorphism selects* $\beta = 1$ *in* (7.84), *thus giving* (7.86).

The reason why automorphisms like the B-transformation are interesting for us is that they tell us something about the symmetries of a theory. Consider a manifold with some metric g. We say that some vector field V is an isometry (and therefore generates a symmetry of the metric) if the Lie derivative $\mathscr{L}_V g$ vanishes. If we have an anti-symmetric field b on a manifold, one is tempted to demand that symmetries correspond to vector fields for which the Lie derivative of b vanishes. In fact this is too restrictive. Instead it is reasonable to demand that $\mathscr{L}_V b$ vanishes up to some exact form, since any such change of b can be undone by a b-field gauge transformation. Therefore, $V + \xi \in TM \oplus T^*M$ is a symmetry of b if

$$\mathscr{L}_V b = d\xi. \tag{7.89}$$

Consider a 2-form B with $dB = 0$. Imagine changing b by adding B to it. What are the symmetries of the new $b + B$ field? We claim that the B-transform of $V + \xi$ is a symmetry of $b + B$,

$$\mathscr{L}_V (b + B) = d(\xi + \iota_V B). \tag{7.90}$$

It is straightforward to verify this by explicit calculation. From this we see that B-transformations of b do not change the symmetries of the theory. Thus it is reasonable to promote B-transformations to automorphisms of the bracket, thus selecting the Courant-bracket.

7.4.2 *Algebra of Gauge Transformations: From Courant Brackets to C Brackets*

In order to determine the algebra of gauge transformations we switch to a more uniform notation in which we mark all one-forms by tildes while vectors stay undecorated. Hence we consider objects

$$\xi^M = \begin{pmatrix} \tilde{\xi}_i \\ \xi^i \end{pmatrix},$$

denoting gauge parameters in the sum of tangent and cotangent space of the manifold. In an abuse of notation we sometimes write this as

$$\xi^M = (\xi + \tilde{\xi})^M.$$

The gauge algebra is governed by a C-bracket $[\cdot, \cdot]_C$, which is closely related to the Courant-bracket but applies to doubled fields! The Courant bracket does not, of course. Consider the Mth component of such a bracket:

$$\begin{aligned}([\xi_1, \xi_2]_C)^M &= \xi^P_{[1} \partial_P \xi^M_{2]} - \frac{1}{2}\eta^{MN}\eta_{PQ}\xi^P_{[1}\partial_N \xi^Q_{2]} \\ &= \xi_{[1} \cdot \partial \xi^M_{2]} - \frac{1}{2}\xi_{P[1}\partial^M \xi^P_{2]},\end{aligned} \tag{7.91}$$

where the brackets on indices indicate anti-symmetrization. Because of the consistent use of our capitalized indices $M, N, \ldots,$ this bracket is $O(D, D)$ covariant. Note that the second term on the right-hand side involves a contraction of indices and therefore contains the metric η. In a conventional theory it would be unthinkable to include a metric-dependent term in a bracket. In our case the use of the constant metric η causes no complications.

Evaluating this bracket between $\xi_1 + \tilde{\xi}_1$ and $\xi_2 + \tilde{\xi}_2$ displays the appearance of some unusual terms:

$$\begin{aligned}\left[\xi_1 + \tilde{\xi}_1, \xi_2 + \tilde{\xi}_2\right]_C &= [\xi_1, \xi_2] + \mathscr{L}_{\tilde{\xi}_1}\xi_2 - \mathscr{L}_{\tilde{\xi}_2}\xi_1 - \frac{1}{2}\tilde{d}(\tilde{\iota}_{\tilde{\xi}_1}\xi_2 - \tilde{\iota}_{\tilde{\xi}_2}\xi_1) \\ &\quad + [\tilde{\xi}_1, \tilde{\xi}_2] + \mathscr{L}_{\xi_1}\tilde{\xi}_2 - \mathscr{L}_{\xi_2}\tilde{\xi}_1 - \frac{1}{2}d(\iota_{\xi_1}\tilde{\xi}_2 - \iota_{\xi_2}\tilde{\xi}_1),\end{aligned} \tag{7.92}$$

where the dual exterior derivatives acting on functions give objects with a vector (upper) index: $(\tilde{d}f)^i \equiv \tilde{\partial}^i f$. It is unusual to see $\mathscr{L}_{\tilde{\xi}_2}\xi_1$, since Lie derivatives are taken with respect to vector fields and not one-forms. In our case this alternative is allowed since we have (dual) derivatives with upper indices, so that a contraction with a one-form is possible. In the same way it is no surprise to see a bracket of one-forms giving a one-form (an object with a lower index): $[\tilde{\xi}_1, \tilde{\xi}_2]_j \equiv \tilde{\xi}_{[1_i}\tilde{\partial}^i\tilde{\xi}_{2]_j}$.

If we drop the $\tilde{x}$-dependence of the C-bracket this will set $\mathscr{L}_{\tilde{\xi}} \to 0$, $\tilde{d} \to 0$ and $[\tilde{\xi}, \tilde{\xi}] \to 0$. The C bracket reduces to

$$[\xi_1 + \tilde{\xi}_1, \xi_2 + \tilde{\xi}_2]_C\big|_{\tilde{x}\equiv 0} = [\xi_1, \xi_2] + \mathscr{L}_{\xi_1}\tilde{\xi}_2 - \mathscr{L}_{\xi_2}\tilde{\xi}_1 - \frac{1}{2}d(\iota_{\xi_1}\tilde{\xi}_2 - \iota_{\xi_2}\tilde{\xi}_1). \tag{7.93}$$

We recognize the right-hand side as the Courant-bracket (7.86). Therefore we can view the C-bracket as $O(D, D)$ covariant, double field theory generalization of the Courant-bracket. It can be shown that the β-parameter cannot be incorporated into the C bracket while preserving $O(D, D)$ covariance.

7.4.3 *B*-Transformations

Having identified the algebraic basis of our theory, we now want to understand what are the B-transformations in our setup. Take an element of $O(D, D)$,

$$h = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \tag{7.94}$$

where b is antisymmetric and constant. Acting with this map on E it is easy to compute the transformation

$$E \mapsto E' = h(E) = (E + b)(1)^{-1} = E + b. \tag{7.95}$$

From this one can read off that the transformation h has the effect of leaving G untouched while B is mapped to $B + b$. So indeed h is a B-transformation. Now it is straightforward to see the action of this map on the gauge parameters ξ^M. Explicit evaluation shows that

$$\begin{pmatrix} \tilde{\xi} \\ \xi \end{pmatrix} \mapsto \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\xi} \\ \xi \end{pmatrix} = \begin{pmatrix} \tilde{\xi} + b\xi \\ \xi \end{pmatrix}, \tag{7.96}$$

so that in components the B-transformation is given by

$$\xi^i \mapsto \xi^i, \qquad \tilde{\xi}_i \mapsto \tilde{\xi}_i + b_{ij}\xi^j, \qquad \tilde{\partial}^i \mapsto \tilde{\partial}^i. \tag{7.97}$$

Note that invariance of the dual derivatives implies that B-transformations leave the constraint $\frac{\partial}{\partial \tilde{x}}\phi = 0$ appropriate for the Courant bracket unchanged. One sees that (7.97) implies

$$\xi + \tilde{\xi} \to \xi + \tilde{\xi} + \iota_\xi b,$$

which is exactly the expected result.

We see now how nicely the parts fit together to form a larger picture: from the physics point of view we have arrived at this formulation because we took T-duality seriously and considered it as basic component of our field theory. From the mathematics point of view the B-transformations play a fundamental role as automorphisms of the Courant-bracket, and in fact now we see that they are just the counterpart of certain T-duality transformations that must be incorporated in an $O(D, D)$ invariant formulation.

7.5 Background Independent Action

We now want to put the various parts together and come to a formulation of a doubled action. We have written down before the perturbative action for a double field theory in terms of a background E_{ij} and fields $e_{ij}(x, \tilde{x})$, depending on both the usual coordinates x and their duals $\tilde{x}$. We made an explicit distinction between the background field and its fluctuation, very similar to the splitting $g_{ij} = \eta_{ij} + h_{ij}$ in linearized gravity. In the end, however, one is looking for a manifest background independent version of the action which does not rely on this distinction.

7.5.1 Background Independent Formulation

To stress the point of background independence we introduce the field

$$\mathcal{E}_{ij}(X) = E_{ij} + e_{ij}(x, \tilde{x}) + O(e^2), \tag{7.98}$$

which at the linearized level is the sum of E and e. We have seen how E and e behave under T-duality, and there is also a natural way to transform $\mathcal{E}$. Since $X' = hX$ (recall (7.37)) we expect that $\mathcal{E}$ transforms like

$$\mathcal{E}'(X') = (a\mathcal{E}(X) + b)\,(c\mathcal{E}(X) + d)^{-1}\,. \tag{7.99}$$

The dilaton d is expected to be $O(D, D)$ invariant, so its transformation law should be

$$d'(X') = d(X). \tag{7.100}$$

This is the analogue of the scalar field Lorentz transformation in conventional field theory.

All the identities and constructions presented in previous sections above did not make use of any X-independence of E. Therefore they can be immediately generalized by replacing E with $\mathcal{E}$, keeping the formal expressions unchanged. For example the derivatives D_i in (7.18) can be generalized to curly $\mathcal{D}_i$, and similarly for the $\bar{\mathcal{D}}_i$'s; but now they are defined with the full metric $\mathcal{E}$,

$$\begin{aligned} D_i &= \partial_i - E_{ik}\tilde{\partial}^k \quad \longrightarrow \quad \mathcal{D} \equiv \partial_i - \mathcal{E}_{ik}(X)\tilde{\partial}^k, \\ \bar{D}_i &= \partial_i + E_{ki}\tilde{\partial}^k \quad \longrightarrow \quad \bar{\mathcal{D}} \equiv \partial_i + \mathcal{E}_{ki}(X)\tilde{\partial}^k. \end{aligned} \tag{7.101}$$

These derivatives will now transform with generalized M matrices, that now depend on $\mathcal{E}(X)$ as

$$\begin{aligned} M &= \left(d - cE^t\right)^t \quad \longrightarrow \quad M(X) = \left(d - c\mathcal{E}^t\right)^t, \\ \bar{M} &= \left(d + cE^t\right)^t \quad \longrightarrow \quad \bar{M}(X) = \left(d + c\mathcal{E}^t\right)^t. \end{aligned} \tag{7.102}$$

Indeed, any object will now transform correctly with $M(X)$ and $\bar{M}(X)$ exactly in the way as they transformed with M and $\bar{M}$ before. This is because the transformations come from (7.99), and one needs no derivatives to derive them. We will also write

$\mathscr{E} = g + b$ without any reference to a background field, and the generalized metric in (7.10) becomes

$$\mathscr{H}(\mathscr{E}) = \begin{pmatrix} g - bg^{-1}b & bg^{-1} \\ -g^{-1}b & g^{-1} \end{pmatrix}. \tag{7.103}$$

In particular, the metric $g(X)$ itself is an $O(D, D)$ tensor, so from (7.50) we have

$$\begin{aligned} g(X) &= \bar{M}(X)g'(X')\bar{M}^t(X), \\ g(X) &= M(X)g'(X')M^t(X). \end{aligned} \tag{7.104}$$

Moreover, the transformation of the Hamiltonian in Eq. (7.36) becomes

$$\mathscr{H}(\mathscr{E}'(X')) = h\mathscr{H}(\mathscr{E}(X))h^t. \tag{7.105}$$

We can repeat all the steps that gave the transformation law (7.62) for the variation of $\mathscr{E}$, this time finding

$$\delta\mathscr{E}(X) = M(X)\delta E'(X')\bar{M}^t(X). \tag{7.106}$$

This relation applies to any derivative of $\mathscr{E}$, thus, for example,

$$\partial_i\mathscr{E} = M(X)\partial_i\mathscr{E}'\bar{M}^t(X), \quad \tilde{\partial}^i\mathscr{E} = M(X)\tilde{\partial}^i\mathscr{E}'\bar{M}^t(X). \tag{7.107}$$

This also means that the same transformations apply to the calligraphic derivatives of $\mathscr{E}$:

$$\mathscr{D}_i\mathscr{E} = M(X)\mathscr{D}_i\mathscr{E}'\bar{M}^t(X), \quad \bar{\mathscr{D}}_j\mathscr{E} = M(X)\bar{\mathscr{D}}_i\mathscr{E}'\bar{M}^t(X). \tag{7.108}$$

The derivatives above can also be transformed, if desired (see (7.111) below). Finally, the transformation of the dilaton under gauge transformation is given by

$$\delta d = -\frac{1}{2}\partial_M\xi^M + \xi^M\partial_M d. \tag{7.109}$$

This implies that

$$\delta e^{-2d} = \partial_M\left[\xi^M e^{-2d}\right], \tag{7.110}$$

which tells us that e^{-2d} is a density. Therefore it is identified as $\sqrt{-g}e^{-2\phi} = e^{-2d}$.

There is one small complication which appears when one takes multiple derivatives. To understand this, we observe that the derivatives (7.101) transform covariantly

$$\begin{aligned} \mathscr{D}_i &= M_i^{\,j}(X)\mathscr{D}'_j, \\ \bar{\mathscr{D}}_i &= \bar{M}_i^{\,j}(X)\bar{\mathscr{D}}'_j. \end{aligned} \tag{7.111}$$

Since M is not a constant anymore, multiple derivatives would not transform correctly. We handle this problem simply by not using higher derivatives in the formulation of our action. We can define $O(D, D)$ covariant derivatives, but they will not be needed here.

7.5.2 The $O(D, D)$ Action

After these preparations we can now present the full background independent $O(D, D)$ action for the fields $\mathcal{E}$ and d. The action is given by

$$S_{\mathcal{E},d} = \int dx d\tilde{x} e^{-2d} \Big[-\frac{1}{4} g^{ik} g^{j\ell} \mathcal{D}^p \mathcal{E}_{k\ell} \mathcal{D}_p \mathcal{E}_{ij} + \frac{1}{4} g^{k\ell} \left(\mathcal{D}^j \mathcal{E}_{ik} \mathcal{D}^i \mathcal{E}_{j\ell} + \bar{\mathcal{D}}^j \mathcal{E}_{ki} \bar{\mathcal{D}}^i \mathcal{E}_{\ell j} \right) + \left(\mathcal{D}^i d \bar{\mathcal{D}}^j \mathcal{E}_{ij} + \bar{\mathcal{D}}^i d \mathcal{D}^j \mathcal{E}_{ji} \right) + 4 \mathcal{D}^i d \mathcal{D}_i d \Big]. \quad (7.112)$$

Each term is independently $O(D, D)$ invariant, and so is the whole action. This also means, though, that the action is not completely determined by $O(D, D)$ invariance, since the numerical factor in front of each term is arbitrary. What finally fixes the action is diffeomorphism and Kalb-Ramond gauge invariance. There is a particular combination of the coefficients, so that the theory is consistent and exhibits these expected gauge invariances. Also, one can expand this action and recover to quadratic and cubic part of the action exactly as in (7.64) and (7.70). Moreover, taking $\tilde{\partial} = 0$, $S_{\mathcal{E},d}$ reduces to an action that is identical to the standard Einstein action plus antisymmetric field plus dilaton, when $\sqrt{-g} e^{-2\phi} = e^{-2d}$. So all this is consistent and fixes the action uniquely.

This action is invariant under the following gauge transformations

$$\delta_\xi \mathcal{E}_{ij} = \partial_i \tilde{\xi}_j - \partial_j \tilde{\xi}_i + \mathcal{L}_\xi \mathcal{E}_{ij} + \mathcal{L}_{\tilde{\xi}} \mathcal{E}_{ij} - \mathcal{E}_{ik} \left(\tilde{\partial}^k \xi^\ell - \tilde{\partial}^\ell \xi^k \right) \mathcal{E}_{\ell j}. \quad (7.113)$$

This is in fact quite a natural expression. The first three terms are the standard terms including the Kalb-Ramond gauge transformation and the usual Lie derivative. The last three terms are zero in a situation where the theory does not depend on the dual coordinate $\tilde{x}$. They are the counterparts to the first three terms which make the transformation compatible with $O(D, D)$. The field $\mathcal{E}$ appears additionally in the last terms in order to get the right index structure. Hence, all the terms that appear here are expected and natural. However, proving the gauge invariance directly is hard.

7.5.3 Formulation Using the Generalized Metric

As next step we want to arrive at an even better formulation of the action without explicit reference to the metric g. Ideally we want to express everything in terms of the generalized metric only, in a form that resembles the Einstein-Hilbert action as far as possible.

For example, for the dilaton we previously found the $O(D, D)$ invariant term

$$4 \mathcal{D}^i d \mathcal{D}_i d.$$

This is actually a complicated term since the $\mathscr{E}$ is contained in the derivatives $\mathscr{D}$. We can also try to formulate a dilaton term with usual partial derivatives only, but then we must be careful how to contract the indices. Certainly a contraction with η is not reasonable, since then the constraint $\partial^M A \partial_M B = 0$ would kill this term. The only other possibility is to contract the indices with $\mathscr{H}$, yielding a term

$$4\mathscr{H}^{MN}\partial_M d \partial_N d.$$

It takes only little calculation to see that this terms is identical to the dilaton term used above. The advantage of this formulation is that we got rid of the explicit appearance of $\mathscr{E}$ and introduced $\mathscr{H}$ instead.

This does not only work for the dilaton term, but also all other terms in this action can be rephrased in this way. Doing so one finds the action

$$S_{\mathscr{H}} = \int dx d\tilde{x} e^{-2d}\Big(\frac{1}{8}\mathscr{H}^{MN}\partial_M \mathscr{H}^{KL}\partial_N \mathscr{H}_{KL} - \frac{1}{2}\mathscr{H}^{MN}\partial_N \mathscr{H}^{KL}\partial_L \mathscr{H}_{MK} - 2\partial_M d \partial_N \mathscr{H}^{MN} + 4\mathscr{H}^{MN}\partial_M d \partial_N d\Big). \tag{7.114}$$

This action is $O(D, D)$-invariant since all indices are correctly contracted. This action is identical to the action in (7.112) although this takes some computation to verify. Finally, by dropping the $\tilde{x}$-dependence it reduces to the expected low-energy action (7.2).

7.5.4 Generalized Lie Derivative

The action (7.114) also comes with a gauge symmetry, and this is quite surprising and rather elegant. In a conventional setting the Lie derivatives appearing in such a theory are

$$\mathscr{L}_\xi A_M = \xi^P \partial_P A_M + \partial_M \xi^P A_P, \tag{7.115}$$

$$\mathscr{L}_\xi B^N = \xi^P \partial_P B^N - \partial_P \xi^N B^P. \tag{7.116}$$

In our setting here we cannot use these; there is a very basic reason why the normal Lie derivative is not applicable. Since we include the Kalb-Ramond field in our theory, there are redundant gauge transformations where the one-form gauge parameter is d-exact. In double field theory the vector field gauge parameter can also be trivial. Indeed, consider the gauge parameter ξ^M to be the derivative of some χ, in components

$$\xi^M = \begin{pmatrix} \tilde{\xi}_i \\ \xi^i \end{pmatrix} = \begin{pmatrix} \partial_i \chi \\ \tilde{\partial}^i \chi \end{pmatrix} = \partial^M \chi. \tag{7.117}$$

The one-form $\tilde{\xi}_i$ is trivial because it is a derivative and so is the vector ξ^i being a dual derivative. Hence ξ^M is a trivial gauge parameter and it should generate no Lie derivative. We see, however, that

$$\mathscr{L}_{\xi=\partial\chi} A_M = \partial^P \chi \partial_P A_M + \partial_M \left(\partial^P \chi\right) A_P \neq 0. \tag{7.118}$$

The first term is zero because of the constraint, but the second term is not zero. Since the Lie derivative does not vanish we should modify its definition. In fact there is a natural way to do so. Using the metric η^{MN} it is possible to define a generalized Lie derivative by

$$\begin{aligned} \widehat{\mathscr{L}}_\xi A_M &\equiv \xi^P \partial_P A_M + \left(\partial_M \xi^P - \underline{\partial^P \xi_M}\right) A_P, \\ \widehat{\mathscr{L}}_\xi B^N &\equiv \xi^P \partial_P B^N - \left(\partial_P \xi^N - \underline{\partial^N \xi_P}\right) B^P. \end{aligned} \tag{7.119}$$

The underlined terms are new and writing them uses the metric twice: once to raise the derivative index and once to lower the gauge parameter index. The conventional Lie derivative distinguishes very much between covariant and contravariant indices. The generalized Lie derivative is more democratic and treats covariant and contravariant indices in a more symmetric way. It is now easy to verify that the generalized Lie derivative along a trivial field vanishes:

$$\widehat{\mathscr{L}}_\xi = \partial \chi A_M = \partial^P \chi \partial_P A_M + \left(\partial_M \partial^P \chi - \partial^P \partial_M \chi\right) A_P = 0. \tag{7.120}$$

$\widehat{\mathscr{L}}$ is the correct Lie derivative to use in our theory. Generalized tensors are objects with $O(D, D)$ indices $M, N, \cdots$, up or down, for which the (generalized) Lie derivative takes the form implied by (7.119).

With the new generalized Lie derivative at hand we can now write the gauge transformations. The gauge transformations of the generalized metric are given by

$$\delta \mathscr{H}^{MN} = \widehat{\mathscr{L}}_\xi \mathscr{H}^{MN}. \tag{7.121}$$

For the dilaton we have

$$\delta e^{-2d} = \partial_M \left[\xi^M e^{-2d}\right]. \tag{7.122}$$

Both transformations vanish for $\xi^M = \partial^M \chi$.

The commutator of two generalized Lie derivatives gives a very elegant expression

$$\left[\widehat{\mathscr{L}}_{\xi_1}, \widehat{\mathscr{L}}_{\xi_2}\right] = -\widehat{\mathscr{L}}_{[\xi_1,\xi_2]_C}. \tag{7.123}$$

The commutator is itself a generalized Lie derivative with parameter obtained by the C-bracket. This shows that the C-bracket determines the algebra of symmetries of this theory.

Exercise 12 *Use* (7.119) *to prove that* (7.123) *holds when acting on* A_M.

7.5.5 Generalized Einstein–Hilbert Action

We have constructed two Lagrangians $\mathcal{L}_{\mathcal{E},d}$ and $\mathcal{L}_{\mathcal{H}}$ which look very different since they are formulated in different variables, but are in fact equal. Both are T-duality invariant, and they use field variables that reflect the doubling of coordinates. The second one, $\mathcal{L}_{\mathcal{H}}$, is perhaps most novel because it completely relies on the use of the generalized metric, which is some kind of metric for a space with doubled coordinates.

Although the Lagrangian $\mathcal{L}_{\mathcal{H}}$ is already written in a reasonably nice form, one can try to take this construction even further. One may ask if there is such a thing as a generalized Ricci curvature or a generalized scalar curvature. In fact, the answer is positive and both objects can be constructed out of the generalized metric *and* the dilaton. Curiously, it seems that there is no "generalized" Riemann curvature, although this has not been established for certain. We do not need the Riemann curvature for writing down a generalized Einstein–Hilbert action, so we will leave this question aside.

The generalized scalar curvature $\mathcal{R}$ is given by the expression

$$\begin{aligned}\mathcal{R} = 4\mathcal{H}^{MN}\partial_M\partial_N d - \partial_M\partial_N\mathcal{H}^{MN} - 4\mathcal{H}^{MN}\partial_M d\partial_N d + 4\partial_M\mathcal{H}^{MN}\partial_N d \\ + \frac{1}{8}\mathcal{H}^{MN}\partial_M\mathcal{H}^{KL}\partial_N\mathcal{H}_{KL} - \frac{1}{2}\mathcal{H}^{MN}\partial_M\mathcal{H}^{KL}\partial_K\mathcal{H}_{NL}.\end{aligned} \tag{7.124}$$

It does contain second derivatives, which is indeed expected since just like in gravity one cannot construct a scalar curvature with just one derivative. Note that the derivatives appearing here are ∂ and not $\mathcal{D}$, so this imposes no problem since they transform with constant h. Each term in (7.124) is $O(D, D)$ invariant, but only the full combination of terms is a generalized scalar.

A simple rearrangement of total derivatives in $S_{\mathcal{H}}$ shows that

$$S_{\mathcal{H}} = \int dx d\tilde{x} e^{-2d}\mathcal{R}(\mathcal{H}, d). \tag{7.125}$$

We see that the action takes a very simple form in terms of the generalized scalar curvature. It looks rather analogous to the conventional Einstein–Hilbert action.

In order to prove the gauge invariance of $S_{\mathcal{H}}$ we can calculate $\delta_\xi\mathcal{R}$ using $\delta_\xi\mathcal{H}$ and $\delta_\xi d$. A substantial calculation confirms that $\mathcal{R}$ is a generalized scalar:

$$\delta_\xi\mathcal{R} = \xi^M\partial_M\mathcal{R}. \tag{7.126}$$

Since $\mathcal{R}$ is a generalized scalar and e^{-2d} is a generalized density, the action is gauge invariant. When the dependence on $\tilde{x}$ is ignored (that is, setting $\tilde{\partial} = 0$) the generalized scalar curvature reduces to

$$\mathcal{R}|_{\tilde{\partial}=0} = R + 4\left(\Box\phi - (\partial\phi)^2\right) - \frac{1}{12}H^2, \tag{7.127}$$

with $H = dB$ and R being the conventional Ricci scalar. This shows that scalars in general relativity do not necessarily correspond to generalized scalars in the double field theory. In general relativity all three terms on the right-hand side of (7.127) are scalars but are not separately $O(D, D)$ invariant. In $\mathscr{R}$ all terms are $O(D, D)$ invariant, but separately are not generalized scalars.

In these lectures we have given a self-contained introduction to double field theory. We have constructed Lagrangians that implement T-duality more explicitly than before. We have seen the natural emergence of the Courant-bracket and how the generalized metric provides a natural variable for the formulation of the theory. One can view the Lagrangians built here as rewritings of the familiar theory that make $O(D, D)$ symmetry manifest. To obtain such Lagrangians we had to impose the "strong" constraint, and it is not yet clear if this constraint may be relaxed. This also means that the power of double field theory has not yet been fully unleashed.

Acknowledgments I would like to thank the organizers of the International School for their invitation to lecture and for their hospitality. I am also grateful to Marco Baumgartl and Nicolas Moeller who prepared an excellent version of the lecture notes that was easy to edit and finalize. Finally, I thank Olaf Hohm for comments and suggestions on this draft.

References

1. Hull, C., Zwiebach, B.: Double field theory. JHEP **0909**, 099 (2009). [arXiv:0904.4664 [hep-th]]
2. Hull, C., Zwiebach, B.: The gauge algebra of double field theory and courant-brackets. JHEP **0909**, 090 (2009). [arXiv:0908.1792 [hep-th]]
3. Hohm, O., Hull, C., Zwiebach, B.: Background independent action for double field theory. JHEP **1007**, 016 (2010). [arXiv:1003.5027 [hep-th]]
4. Hohm, O., Hull, C., Zwiebach, B.: Generalized metric formulation of double field theory. JHEP **1008**, 008 (2010). [arXiv:1006.4823 [hep-th]]
5. Siegel, W.: Superspace duality in low-energy superstrings. Phys. Rev. **D48**, 2826–2837 (1993). [hep-th/9305073]
6. Siegel, W.: Two vierbein formalism for string inspired axionic gravity. Phys. Rev. **D47**, 5453–5459 (1993). [hep-th/9302036]
7. Tseytlin, A.A,: Duality symmetric closed string theory and interacting chiral scalars, Nucl. Phys. **B350**, 395–440 (1991)