String Basics

J.L.F. Barbón [1]

Theory Division, CERN
CH-1211 Geneva 23, Switzerland
barbon@cern.ch

Abstract

This is a simple conceptual introduction to some basic facts about string theory.

[1]Lectures at "Taller de Altas Energías". April 2002, Peñíscola, Spain.

Prefacio

Las presentes notas se basan en el curso impartido con el título "String Basics" en el primer Taller de Altas Energías, celebrado en Peñíscola durante la primera mitad de Abril del año 2002. Aunque la versión preliminar existía con anterioridad al curso, esta versión para difusión pública contiene añadidos y mejoras puntuales como resultado de la interacción con los estudiantes durante el curso.

La presentación, un tanto atípica para un curso de supercuerdas, responde a la necesidad de hacer accesibles los resultados más importantes a una audiencia muy heterogénea, compuesta de doctorandos en temas teóricos y experimentales. Por esa razón he tratado de enfatizar los aspectos conceptuales sobre las aplicaciones, siempre sacrificando el formalismo a un estilo heurístico.

Los expertos saben bien que esta elección conlleva peligros considerables. Dado el carácter intrínsecamente técnico de la teoría de cuerdas, la línea que separa la pedagogía de la vulgarización no está siempre bien definida. Concretamente, uno de los resultados clave del curso; la deducción del espectro libre de cuerdas relativistas abiertas y cerradas (Lecture 2), se obtiene mediante un atajo que involucra la cuantización de cuerdas en la aproximación no relativista. Si bien trato de justificar el procedimiento mediante consideraciones de dualidad, el lector debe hacerse cargo de que una justificacón rigurosa de las fórmulas requiere un tratamiento completamente relativista e invariante gauge.

Dado que estos tratamientos son estándar y se pueden encontrar en numerosos libros de texto, he optado por sacrificar el rigor en aras de la brevedad. Al fin y al cabo, una cierta provisión de "trucos" es lo que caracteriza un curso oral como complemento al necesario aprendizaje autodidacta en un buen libro.

La organización de las charlas es la siguiente. En la primera se introduce un punto de vista heterodoxo sobre la teoría de perturbaciones covariante en teoría cuántica de campos. Se trata esencialmente del viejo punto de vista espacio-temporal de Feynman, clarificado mediante la representación de tiempo propio de Schwinger. Una buena parte de las peculiaridades técnicas de la teoría de cuerdas a nivel perturbativo se ven aquí en estado embrionario. Esto convierte en "natural" la generalización de teoría cuántica de campos a una teoría de objetos extensos. Cerramos la primera charla con el conjunto de "predicciones" genéricas de la teoría de cuerdas: gravitación, simetría gauge, dimensiones extra y supersimetría, cuyo desarrollo constituye el resto de las tres charlas, y una discusión de la principal motivación teórica de las teorías de cuerdas como un modelo de gravitación cuántica.

En la segunda charla obtenemos el espectro de cuerdas libres abiertas y cerradas. El objetivo principal es mostrar que las cuerdas requieren una dimensión espacio-temporal crítica superior a cuatro, que son genéricamente intestables en su versión más simple (bosónica) y que contienen bosones gauge (cuerda abierta) y gravitones (cuerda cerrada) de forma universal.

En la tercera y cuarta charlas introducimos un conjunto de resultados de carácter más concreto y que tienen que ver con propiedades dinámicas de las teorías de cuerdas. El mayor problema en la elaboración de los temas es la importancia del formalismo supersimétrico, un requisito no asumible en la cultura previa de los estudiantes para este curso. En este asunto hemos optado por introducir la idea básica de supersimetría para espacios de Fock libres, que es el andamiaje mínimo para construir el espectro libre de las teorías de supercuerdas, y hemos suprimido todos los detalles de las construcciones específicas de teorías de cuerdas de tipo II, tipo I y cuerdas heteróticas. A cambio, hay una discusión elemental del papel de supersimetría en modelos fenomenológicos a la escala del modelo estándar (problema de las jerarquías).

A continuación hay una discusión más bien descriptiva del impacto de supersimetría y simetrías de dualidad en la estructura no perturbativa de las teorías de cuerdas y una presentación de los principales prejuicios sobre la escala dinámica a la que aparecerían las cuerdas, basado en el comportamiento de los acoplamientos gauge (escenarios tipo GUT o tipo "dimensiones extra grandes").

Por último presentamos un argumento físico que explica la finitud de las teorías de cuerdas basado en el estado típico de una cuerda a muy alta energía. Utilizamos este resultado para introducir algunas ideas recientes sobre la relación entre teorías de cuerdas y agujeros negros cuánticos, así como el "principio holográfico".

Las notas se complementan con una lista de siete problemas resueltos que han sido seleccionados con el objetivo de ilustrar ciertos conceptos básicos, pero también como remedio improvisado a la incompletitud manifiesta de las notas en ciertos puntos. No se adjunta una colección de referencias, si bien recomendamos al lector interesado el estudio cuidadoso de los tratados clásicos: Green–Schwarz–Witten y Polchinski.

Agradecimientos

Vaya por delante mi agradecimiento a los estudiantes que han contribuido con sus preguntas a mejorar estas notas. Debo mencionar también el magnífico trabajo del tutor de problemas, Ernesto Lozano, así como las críticas al manuscrito de Luis Alvarez-Gaumé.

Desearía por último agradecer a los organizadores del taller: Joan Fuster, María José Herrero, José Ignacio Latorre y Javier Mas, su esfuerzo y visión al impulsar una iniciativa nueva en nuestro país: una escuela de doctorado que, a partir del formato ensayado con gran éxito en Santiago de Compostela durante una década de cursos teóricos, pone en contacto a experimentales y teóricos de altas energías. Se genera así un vínculo generacional que, en mi opinión, es fundamental para el futuro de nuestra disciplina. El éxito reconocido de esta primera edición debe mucho al entusiasmo de los organizadores. Sin embargo, la lección principal la proporcionan los propios estudiantes, que una vez más han demostrado ser la más rentable de las inversiones.

J.L.F. Barbón,

Ginebra, Abril 2002.

Lecture 1

From Point Particles to Strings

In the first lecture we recall the unconventional (original) formulation of covariant perturbation theory due to Feynman. The basic physical interpretation of relativistic Quantum Field Theory (QFT) is in terms of a many-particle system. In perturbation theory, the interactions really correspond to creation and annihilation of particles so that the number of particles is not conserved. In cases where a classical conservation law exists, it is realized quantum mechanically as the conserved number of particles *minus* the number of antiparticles. Although the existence of antimatter is the one *generic* prediction of QFT, it appears in the usual formalism in a rather technical way. The purpose of the next section is to introduce Feynman's picture based on the quantization of relativistic *single* particles. The inevitability of antimatter acquires then a nice geometric interpretation.

Feynman's point of view is the best technical way of introducing the string generalization of QFT. The result of this generalization is a remarkable list of *generic* predictions, much in parallel with the generic prediction of QFT (antimatter). The most surprising of these predictions is the dynamical emergence of gravity.

Feynman Diagrams and Particle Paths

In perturbation theory in powers of the coupling λ, the physics of a QFT with a simple Lagrangian such as

$$\mathcal{L}_\phi = -\frac{1}{2}\,(\partial\phi)^2 - \frac{\lambda}{6}\,\phi^3$$

reduces to the computation of Feynman diagrams.[1] Each diagram is a code for the mathematical expression

$$\sum_{\text{q. numbers}} \prod_{\text{vertices}} (-i\lambda) \prod_{\text{links}} G_{\text{link}},$$

where the sum over quantum numbers of intermediate particles includes momenta (or positions), spin, charge, etc. The hart of perturbation theory is the propagator function G associated to each link of the diagram. For the propagation between two points $x \to x'$ in the scalar model above we have

$$G(x',x) = \langle\, 0\,|\, \mathrm{T}\,[\,\phi(x')\,\phi(x)\,]\;|\,0\,\rangle, \tag{1}$$

i.e. the vacuum expectation value of the time-ordered product of two field operators. Upon calculation, this is

$$G(x',x) = \lim_{\epsilon\to 0} \int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot(x-x')}\, \frac{-i}{p^2 - i\epsilon}.$$

In a somewhat more symbolic form:

$$G(x',x) = \left\langle\, x' \,\right|\, -\frac{i}{p^2}\, \left| x \,\right\rangle,$$

where the pole at $p^2 = 0$ is resolved by giving p^2 a small negative imaginary part.

[1] We use the "gravitational" metric convention $\eta_{\mu\mu} = \text{diag}\,(-1,1,1,1)$, and $V^2 = V^\mu V_\mu = \eta_{\mu\nu}\,V^\mu V^\nu$.

Feynman's original derivation of his rules was in terms of path integrals for single-particle trajectories (rather than path integrals for *fields* like in modern textbooks). We can derive his starting point by working backwards from the expression of the propagator. Let us use the so-called Schwinger parametrization:

$$G(x',x) = \int_0^\infty ds\, \langle\, x' \,|\, e^{-i\,s\,p^2}\,|\, x\,\rangle$$

Defining the "Hamiltonian"

$$\mathcal{H} = p^2 = \frac{p^2}{2\cdot\frac{1}{2}},$$

the matrix element

$$\langle\, x' \,|\, e^{-i\,s\,\mathcal{H}}\,|\, x\,\rangle$$

looks like the propagation of a non-relativistic particle of mass 1/2 between points x and x' in time s. We can now write this in Lagrangian form by using Feynman's path integral formula:

$$\langle\, x' \,|\, e^{-i\,s\,\mathcal{H}}\,|\, x\,\rangle = \int_{x\to x'} Dx^\mu(\tau)\ \exp\left(i\int_0^s d\tau\ \mathcal{L}\left[x(\tau)\right]\right) \tag{2}$$

as a *formal* sum over paths $x^\mu(\tau)$ that connect x and x' in "time" s, i.e. $x(0) = x$, and $x(s) = x'$. The Lagrangian is that of a non-relativistic particle of mass 1/2, thus

$$\mathcal{L} = \dot{x}\cdot p - \mathcal{H} = \frac{1}{2}\,(\tfrac{1}{2})\,\dot{x}^2 = \frac{1}{4}\,\frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau} \tag{3}$$

The "time" variable τ parametrizes the quantum trajectory $x^\mu(\tau)$ of the relativistic particle. Notice that τ is a proper time along the trajectory and is fundamentally different from the external physical time $x^0 = t$.

Since τ simply parametrizes the curve $x^\mu(\tau)$, the particular choice of path parametrization should not be important. Under a reparametrization $\tau \to \tau'$ we have

$$d\tau = d\tau'\,\frac{d\tau}{d\tau'}, \qquad \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{d\tau'}\,\frac{d\tau'}{d\tau}$$

Therefore, in order to make the action invariant under reparametrizations we can introduce a "one-dimensional" metric $h_{\tau\tau} < 0$ with the property that

$$ds^2 = -h_{\tau\tau}\,d\tau\,d\tau$$

is invariant under reparametrizations. This means that

$$h_{\tau\tau} = h_{\tau'\tau'}\left(\frac{d\tau'}{d\tau}\right)^2,$$

so that the combination

$$h^{\tau\tau}\ \partial_\tau x^\mu\,\partial_\tau x_\mu$$

is invariant, where $h^{\tau\tau} \equiv 1/h_{\tau\tau}$. Also

$$d\tau\,\sqrt{-h_{\tau\tau}} = \text{invariant},$$

and s is interpreted as the "proper length" of the trajectory:

$$s = \int d\tau \sqrt{-h_{\tau\tau}}$$

As a result, we have that the action

$$S_{\rm P} = -\frac{1}{4} \int d\tau \sqrt{-h_{\tau\tau}}\, h^{\tau\tau}\, \partial_\tau x^\mu\, \partial_\tau x_\mu \tag{4}$$

is invariant under reparametrizations and reduces to (3) for a choice of "gauge" $h_{\tau\tau} = -1$. In this general view, the propagator is obtained as the path integral:

$$G(x', x) = \int \frac{Dh_{\tau\tau}}{\rm Diff}\, [Dx^\mu(\tau)]_{x\to x'}\; e^{iS_{\rm P}[h_{\tau\tau}, x^\mu(\tau)]},$$

i.e. "quantum gravity" in one (time) dimension. The integral over world-line metrics $h_{\tau\tau}$ modulo reparametrizations, or "diffeomorphisms", of the line is just a redundant way of writing the original path integral and should reduce to it upon fixing the gauge of the group of reparametrizations: Diff. In particular, since the metric $h_{\tau\tau}$ determines the proper length along the path, $s = \int d\tau\sqrt{-h_{\tau\tau}}$, and this is the only reparametrization-invariant property of the path, any functional $F[h_{\tau\tau}]$ of the metric should actually be a function of s only and we should have

$$\int \frac{Dh_{\tau\tau}}{\rm Diff}\, F\left[\int d\tau\sqrt{-h_{\tau\tau}}\right] \sim \int_0^\infty ds\; F[s],$$

so that the Schwinger-parameter integral is just the result of gauge-fixing the integral over one-dimensional metrics.

Back to QFT

We can also approach the problem with the method of canonical quantization. On quantizing (4) we may choose the standard gauge $h_{\tau\tau} = -1$ provided we impose its equation of motion as a constraint:

$$\frac{\delta S_{\rm P}}{\delta h^{\tau\tau}} \equiv -\frac{1}{2}\sqrt{-h_{\tau\tau}}\; T_{\tau\tau} = 0,$$

where

$$T_{\tau\tau} = \frac{1}{4}\, \partial_\tau x^\mu\, \partial_\tau x_\mu$$

In the quantum theory, we must impose this (analogous to Gauss's law in electrodynamics) by declaring that the quantum operator has vanishing matrix elements on physical states:

$$\langle \psi \,|\, T_{\tau\tau} \,|\, \phi \rangle = 0$$

Now, canonical quantization of the gauge-fixed action with $h_{\tau\tau} = -1$ gives

$$[\, p_\mu,\; x^\nu\,] = -i\, \delta_\mu{}^\nu,$$

with

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = \frac{1}{2}\, \dot{x}_\mu$$

Thus, we can realize the momentum as the quantum operator

$$p_\mu = -i\,\partial_\mu$$

and

$$T_{\tau\tau} = p^\mu\, p_\mu = \mathcal{H}$$

is nothing but the world-line Hamiltonian. Now the quantum constraint $\langle\psi|\,T_{\tau\tau}\,|\phi\rangle = 0$ imposes the vanishing of the matrix elements of $\mathcal{H}$ on physical states. The vanishing of the Hamiltonian is natural for a theory that is invariant under time reparametrizations. A suggestive presentation of this property is

$$\langle\, x\,|\, p^\mu p_\mu\,|\,\phi\,\rangle = -\partial^\mu\partial_\mu\,\langle\, x\,|\,\phi\,\rangle = -\partial^2\,\phi(x) = 0,$$

the Klein–Gordon equation!

This is rather remarkable. By quantization of a *single* free *relativistic* scalar particle we have obtained the free *field* equation. It remains to see how the multiparticle interpretation of QFT comes out. It turns out that this fact can be traced to simple geometrical facts in Feynman's formalism, even at the level of the free particle's propagation.

To see this, notice that in computing the Feynman propagator out of the path integral (2) one must specify boundary conditions. In other words, what kind of paths are to be included in the path integral?

A first guess would be that, in order to respect causality, one should restrict to the quantum trajectories that lie inside the relative future light-cone of $x' - x$, that is we would require that the trajectories are "causal" in that the tangent vector to the curves is never spacelike:

$$\eta_{\mu\nu}\,\dot{x}^\mu(\tau)\,\dot{x}^\nu(\tau) \leq 0$$

However, if we do this we do not obtain the right answer. The correct propagator in QFT is (1), where the instruction of time ordering is essential. This is related to Feynman's $i\epsilon$ prescription which adds a small negative imaginary part to p^2. The poles are shifted as

$$p^0 = \pm i\epsilon \mp |\vec{p}\,|,$$

so that one may as well compute the propagator by analytic continuation through the Wick rotation $p^0 \to ip^0$, or $x^0 \to -ix^0$. This sends us to Euclidean space where there is no light cone and no notion of causality, so that the natural class of paths to be included has no constraints on the tangent vectors. After analytic continuation back to Minkowski space on finds that the class of paths that must be considered do not lie inside the relative light-cone of the initial and final point, i.e. the quantum particle wanders off to space-like regions with

$$\eta_{\mu\nu}\,\dot{x}^\mu(\tau)\,\dot{x}^\nu(\tau) > 0$$

In particular, the quantum trajectory may cut back and forth the surface of constant time x^0. This is interpreted as pair creation of particle-antiparticle pairs, an idea that Feynman credits to a night phone call by Wheeler!

In fact, antimatter is the one generic prediction when one marries quantum mechanics and special relativity or, what is the same, QFT.

The String Generalization

This construction of *perturbative* QFT makes the transition to string theory rather natural. If we replace the point-particle trajectories by extended string trajectories (world-lines by world-tubes or world-sheets) most of the previous construction generalizes by simply adding the dependence on the extension of the string.

The most evident improvement is in the description of the interactions. The splitting of a string into two or more becomes a smooth process, topological in nature. The entire amplitude can be smoothly represented by a sum over all possible surfaces with a given topology, weighted by the same action that applies to the free propagation.

The coupling constant of three closed strings, g_s, or three open strings, g_o, determines the coupling weight of any diagram because it must be a topological invariant of the surface. Let us consider for simplicity oriented strings which only produce oriented two-dimensional surfaces in their motion.

For a surface with the topology of a sphere with n handles and m holes, the handles represent closed-string loops, whereas the holes represent external string states. The Euler number

$$\chi = 2 - 2n - m$$

is a topological invariant that can be computed by use of the Gauss–Bonnet theorem:

$$\chi = \frac{1}{4\pi}\int_{\Sigma}\mathcal{R}^{(2)} + \frac{1}{2\pi}\int_{\partial\Sigma}\mathcal{K}^{(1)},$$

where Σ is the world-sheet, $\mathcal{R}^{(2)}$ is the two-dimensional curvature scalar and $\mathcal{K}^{(1)}$ is the extrinsic curvature on the boundary of the world-sheet.

Since each handle is a closed-string loop, it must be weighted by a factor of g_s^2. Thus, the complete surface with n handles and m holes has a weight proportional to

$$(g_s)^{-\chi}$$

This implies that, if external states are normalized with one power of g_s, the classical normalization of the effective action generating all connected tree-level diagrams is

$$\Gamma(\text{closed})_{\text{tree}} \propto \frac{1}{g_s^2}$$

For open oriented strings, vacuum diagrams with no external states are necessarily equal to some closed-string diagrams, i.e. spheres with handles and holes. Each hole is one extra open string loop, thus it is suppressed by one power of g_o^2. This means that

$$g_o^2 = g_s$$

The lowest-order open-string diagram has topology of a disk, or a sphere with one boundary and is weighted by $g_o^{-2} = g_s^{-1}$. This is the normalization of the generator functional of tree open-string diagrams:

$$\Gamma(\text{open})_{\text{tree}} \propto \frac{1}{g_o^2} = \frac{1}{g_s}$$

Hence, canonical normalization of the propagator of open string states, $g_o^0 = 1$, requires that external open strings be weighted by a power of

$$g_o = \sqrt{g_s}$$

The most general amplitude with N_o external open strings, N_c external closed strings, L_o open-string loops and L_c closed string loops is a sphere with L_c handles and $L_o + 1 + N_c$ holes. The coupling dependence is

$$(g_s)^{-1+L_c+L_o+N_c+\frac{1}{2}N_o}$$

In this formula, the case with no open strings is interpreted as $L_o = -1$. We learn that there is only one fundamental coupling, the closed string coupling g_s. This gives string interactions a universality that is not shared by point-particle theories. In the space-time construction of point-particle perturbation theory the splitting of particles at interaction points must be specified in a somewhat arbitrary fashion. For strings, on the other hand, the path integral for a given perturbative contribution is a sum over smooth surfaces Σ of a given topology with an appropriate generalization of (4).

We can say that, in string theory, knowing the free propagation (free spectrum) of the string determines univocally all perturbative interactions.

The spectrum is determined by the world-sheet action. The covariant and reparametrization-invariant action (4) generalizes to the so-called Polyakov action

$$S_{\rm P} = -\frac{1}{4\pi\alpha'}\int d^2\sigma\,\sqrt{-h}\,h^{ab}\,\partial_a X^\mu\,\partial_b X^\nu\,\eta_{\mu\nu}, \tag{5}$$

where $a, b = 0, 1$ runs over the world-sheet coordinates and

$$h \equiv \det(h_{ab}).$$

As in the point-particle case, the metric h_{ab} is also redundant and can be gauge-fixed to $h_{ab} = \eta_{ab}$ in conventional cartesian coordinates for the world-sheet (τ, σ), for which the Polyakov action reads,

$$S_{\rm P}[h_{ab} = \eta_{ab}] = \frac{1}{4\pi\alpha'}\int d\tau\,d\sigma\,\left(\partial_\tau X^\mu\,\partial_\tau X_\mu - \partial_\sigma X^\mu\,\partial_\sigma X_\mu\right). \tag{6}$$

This redundancy of the two-dimensional metric is crucial to the theory and requires the extra symmetry of (5) of local Weyl rescalings (see Problem 1):

$$h_{ab}(\sigma) \to e^{2\omega(\sigma)}\,h_{ab}(\sigma).$$

The constant

$$\alpha' \equiv \ell_s^2 \equiv \frac{1}{m_s^2}$$

is called the Regge slope and has been introduced for dimensional reasons. Its physical interpretation is that

$$\frac{1}{2\pi\alpha'}$$

measures the tension (mass per unit length) of the string. To see this, we use the equation of motion of the metric

$$-4\pi\sqrt{-h}\,\frac{\delta S_{\rm P}}{\delta h_{ab}} \equiv T_{ab} = 0$$

It is proven in Problem 1 that the Polyakov action (5) evaluated on the solutions of $T_{ab} = 0$ gives the so-called Nambu–Goto action

$$S_{\rm NG} = -\frac{1}{2\pi\alpha'} \int d^2\sigma \sqrt{-\det\left(\partial_a X^\mu\, \partial_b X_\mu\right)} \tag{7}$$

The integral in this action is the area of the world-sheet embedded in space-time by the string coordinates $X^\mu(\sigma^a)$. To see this, consider a straight string at rest along the coordinate axis X^1. Choosing the world-sheet coordinates as $\tau = X^0, \sigma = X^1$ we find

$$S_{\rm NG} = -\frac{1}{2\pi\alpha'} \int d\tau\, d\sigma = -\frac{L}{2\pi\alpha'} \int d\tau,$$

where L is the length of the string. Therefore, we recover the interpretation of α' as an inverse tension parameter.

The quantization program for the Polyakov action follows step by step the analogous program for the path integral over particle paths. One can either compute the path integral over surfaces by appropriately fixing the gauge (introducing ghosts for the reparametrization group) or quantize the action (6) in the canonical formalism. Since (6) is the action of d free scalar fiels in two dimensions, one can quantize the string coordinates directly in terms of d chains of harmonic oscillators. The analogue of the constraint equation

$$\langle\, \psi \,|\, T_{\tau\tau} \,|\, \phi \,\rangle = 0$$

in the particle case is now

$$\langle\, \psi \,|\, T_{ab} \,|\, \phi \,\rangle = 0,$$

where T_{ab} is the two-dimensional energy-momentum tensor:

$$T_{ab} = \frac{1}{\alpha'}\, \partial_a X^\mu\, \partial_b X_\mu - \tfrac{1}{2}\, \eta_{ab}\, (\text{trace}).$$

The zero-mode of this constraint equation still corresponds to the vanishing of the world-sheet Hamiltonian and yields the free spectrum of the theory:

$$p^2 + M^2 = 0$$

with M^2 given symbolically by

$$M^2 \propto \frac{1}{\alpha'}\, (\text{Osc} + C) \tag{8}$$

in terms of the total excitation number of the oscillators and a constant contribution C coming from zero-point Casimir fluctuations. The scale of the mass spectrum is set by the Regge slope α' and the rest of the constraint equations decouple negative-norm states much like in the Gupta–Bleuler formalism for gauge theory. The result is that we can just consider transverse oscillations after gauge-fixing.

The derivation of (8) is the subject of many textbook treatments. In these notes we will obtain the basic formula (8) from a non-rigorous short-cut.

The result of carrying out this program is rather impressive. If this line of reasoning gave for QFT its one generic prediction, namely the *inevitability* of antimatter, for strings we have, at the same level of *inevitability*, the following list:

- Closed strings always contain gravity. This is the single most remarkable property of string theory. It always contains a graviton excitation whose perturbative interactions are consistent and finite.
- Open strings always contain gauge fields. This is similar to the former but regarding the origin of gauge symmetry.
- For consistency, always more than four dimensions are required. This means that relation to the real world must proceed by choosing ground states where $d-4$ spacetime dimensions are *small*. This is the old Kaluza–Klein idea, that takes a new look in string theory. In particular, this gives another universal mechanism to generate gauge symmetries.
- Since the coupling of open and closed strings is related by $g_s = g_o^2$ and closed strings generate gauge symmetry out of the (stringy) Kaluza–Klein mechanism, it turs out that gauge interactions are naturally unified at the string mass scale and, at the same time, are also unified with gravity (there is a natural connection with the GUT idea).
- Generic string models based on bosonic strings are *unstable*. The only known stabilization mechanism is supersymmetry. Thus, supersymmetry at stringy energy scales is another (perhaps more technical) generic prediction of string theory.

The fact that this list includes most of the popular ideas about what may lie beyond the Standard Model has turned string theory into a major fashion and, in a sense, a general scenario for theoretical thinking about the relation between gauge symmetry, gravity and supersymmetry.

Since the most remarkable item is the first one in the list (from a strictly theoretical point of view), we pause to review what is the most immediate problem posed by the quantization of gravity.

Interlude: The Problem of Gravity

Einstein gravity in d space-time dimensions is based on the so-called Einstein–Hilbert Lagrangian

$$S_{\rm EH} = \frac{1}{16\pi G_{\rm N}} \int d^d x \sqrt{-g}\, \mathcal{R}, \tag{9}$$

where $\mathcal{R} = g^{\mu\nu} \mathcal{R}_{\mu\nu}$ and $\mathcal{R}_{\mu\nu}$ is the Ricci tensor containing up to two derivatives of the metric $g_{\mu\nu}$. That (9) is non-renormalizable in perturbation theory around flat space can be understood by simple considerations of power-counting in Feynman diagrams. Writing

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa\, h_{\mu\nu}$$

with

$$\kappa = \sqrt{8\pi G_{\rm N}} = \ell_P^{\frac{d-2}{2}} \sqrt{8\pi}$$

we have canonical normalization for $h_{\mu\nu}$, i.e. a kinetic term of the symbolic form

$$S_{\rm FP} = -\frac{1}{2} \int d^d x \; h\, \mathcal{D}_{\rm FP}\, h,$$

with $\mathcal{D}_{\rm FP}$ a second-order differential operator and $h_{\mu\nu}$ of mass dimension $(d-2)/2$. This so-called Fierz–Pauli Lagrangian describes free gravitons and will be studied in Problem 2. Then the three-point couplings are proportional to κ and thus must have two derivatives by elementary dimensional analysis:

$$\int V_3 \sim \kappa \int h\, \partial h\, \partial h$$

In fact, since the Ricci scalar itself contains only up to second derivatives of the metric, all vertices of the theory in perturbation theory have two derivatives.

A diagram with N_e external lines would formally come from a contact term in the Lagrangian of the form

$$\mathcal{L}_{\text{contact}} \sim \eta\, (\partial)^{N_\partial}\, (h)^{N_e}$$

Therefore, by counting dimensions, the mass dimensionality of an amputated amplitude with N_e external lines is

$$\dim(\eta) = d - N_\partial - N_e \frac{d-2}{2}$$

At the same time, if the diagram contains N_v three-point vertices, its *generic* behaviour as the external momenta $|p|$ vanish in comparision with *all* internal momenta is

$$\kappa^{N_v}\, |p|^{N_\partial}\, \Lambda^{\text{div}},$$

where Λ is an ultraviolet cutoff. Combining the two previous equations we obtain

$$\text{div} = d - N_e \frac{d-2}{2} - N_v \dim(\kappa)$$

and we see that any amplitude can become divergent at a sufficiently large order of perturbation theory (large N_v) provided the coupling has negative mass dimension. This is the standard criterion for renormalizability. Symmetries can reduce the order of divergence of various diagrams, but a finite symmetry (such as gauge symmetry, general covariance, supersymmetry, etc) is not expected to be able to cancel all possible divergent diagrams that arise with an *arbitrarily large* number of external legs.

Since the gravitational coupling has mass dimension

$$\dim(\kappa) = \frac{2-d}{2},$$

gravity is non-renormalizable above $d = 2$. In particular it is non-renormalizable in any $d \geq 4$. This means that to get finite physical amplitudes we have to add counterterms with an arbitarily large number of graviton fields. Since the couplings of these operators are not predicted, the theory is strictly-speaking, not predictive at all.

As any non-renormalizable theory, it is still useful provided we use it as an effective low-energy action at energies $E \ll M_P$. If we expand any physical quantity in powers of E/M_P, at a given order n in this expansion, we just need $O(n)$ counterterms. Of course, this means that predictability is lost as the typical energy approaches the Planck scale. In fact, perturbation theory simply breaks down at these energies.

Consider the tree-level scattering of gravitons in four dimensions. It is of order

$$(\kappa p^2) \cdot \frac{1}{p^2} \cdot (\kappa p^2) \sim \kappa^2 p^2 \sim \alpha_G(p)$$

So, the effective dimensionless expansion parameter is just

$$\alpha_G(E) = \left(\frac{E}{M_P}\right)^2$$

This is of $O(1)$ at $E \sim M_P$ and thus perturbation theory breaks down there.

There are, generally speaking, two schools of thought about what this might mean. One of them assumes that the basic Einstein Lagrangian and metric degrees of freedom are to be kept at $E > M_P$, but they are simply strongly coupled and the theory must be studied with non-perturbative methods. In particular, one may imagine that the non-renormalizability is an artifact of perturbation theory. This is a logical possibility, albeit one that has led to little progress in practice, mostly because it is difficult to do any calculation if one gives up the possibility of having some weakly-coupled dynamics.

The second school of thought is that of string theory, that interprets the threshold at M_P as a signature of the breaking of the metric as a good description of the degrees of freedom. The idea is that one may find a weakly coupled description based on entirely new degrees of freedom that become apparent at the Planck scale.

This is exactly parallel to the situation in Fermi's theory of weak interactions. The analogue of Newton's constant is Fermi's constant $G_{\rm F} \sim 10^{-5}$ GeV^{-2}. Therefore, the dimensionless expansion parameter $\alpha_F(E) \sim G_{\rm F}\, E^2 = O(1)$ at about $E \sim G_{\rm F}^{-1/2} \sim 300$ GeV, which is exactly analogous to the Planck mass in this context. However, before that, at $E \sim 100$ GeV, we know that the four-fermion vertex responsible for muon decay gets smeared out by the exchange of W-bosons with mass $M_W \sim 80$ GeV. In fact, we have

$$G_{\rm F} \sim \frac{g_W^2}{M_W^2},$$

where g_W is the appropriate gauge coupling constant of the Standard Model. Therefore, at $E > 100$ GeV, the effective dimensionless coupling is

$$\alpha_W = \frac{g_W^2}{4\pi},$$

which is only renormalized logarithmically, and in particular decreases towards the ultraviolet. This means that we solve the problem of the non-renormalizability of Fermi's theory by postulating new degrees of freedom at a scale below (but close) to $G_{\rm F}^{-1/2}$ and we can make sense of the dynamics at all energies without ever having to give up the weakly coupled description.

String theory does exactly this for the case of gravity, and it is the only known way of doing so. The local graviton-graviton vertex is smeared out by the exchange of strings, i.e. extended objects, with characteristic mass scale $m_s < M_P$. The analogue of the W boson is the string, the analogue of M_W is m_s, the string mass scale, and the analogue of g_W is the dimensionless coupling of three strings that we denote by g_s. Thus we have, just on dimensional grounds:

$$G_{\rm N} \sim \kappa^2 \sim \frac{g_s^2}{m_s^2} = g_s^2\, \ell_s^2,$$

where $\ell_s = 1/m_s$ is the string length. In higher dimensions this generalizes to

$$\kappa_d^2 \sim g_s^2\, \ell_s^{d-2}.$$

Hence, to the extent that $g_s < 1$, we would have achieved a fully perturbative regularization of gravity that exactly mimics that of the electroweak model.

The most notorious property of strings is that they are not simply ultraviolet gadgets that are used in a more-or-less *ad hoc* manner to regularize quantum gravity. It turns out that the strings have massless excitations (in spite of the characteristic mass m_s being close the Planck scale) and in particular gravitons themselves are one of these excitations that appear *universally* for all closed-string theories.

Lecture 2

Quantization

In a rigorous treatment one would quantize the covariant Polyakov action (5). In canonical quantization, one deals with the gauge-fixed action (6):

$$S_{\rm P} = \frac{1}{4\pi\alpha'}\int d\tau\, d\sigma\; (\partial_\tau X^\mu\, \partial_\tau X_\mu - \partial_\sigma X^\mu\, \partial_\sigma X_\mu\,)$$

It is convenient to define light-cone coordinates:

$$\sigma^\pm = \tau \pm \sigma.$$

In these coordinates we have equations of motion

$$\partial_+\, \partial_-\, X^\mu = 0, \tag{10}$$

and constraints

$$\langle\, \phi\,|\; \partial_\pm\, X^\mu\, \partial_\pm\, X_\mu\; |\, \psi\,\rangle = 0 = \langle\, \phi\,|\; T_{\pm\pm}\; |\, \psi\,\rangle. \tag{11}$$

The constraint $\langle\, \phi\,|\, T_{+-}\,|\, \psi\,\rangle = 0$ is automatic provided conformal invariance holds at the quantum level.

This procedure is reviewed in many textbooks. Here we will follow a heuristic short-cut that has some interest in itself.

Physically, the vibrational modes of the string should be classified by representations of the Poincare group, i.e. they must come with a given mass and spin. Thus, our first objective is the calculation of the spectrum of the operator M^2 in the relativistic dispersion relation

$$p^2 + M^2 = 0.$$

A hint at this spectrum can be obtained by considering the infinite-momentum frame. Let us look at the string from a highly boosted frame, say in the $X^{||}$ direction. Then the energy takes the form

$$p^0 = \sqrt{p_{||}^2 + p_\perp^2 + M^2} = |p_{||}| + \frac{\mathcal{H}_\perp}{2|p_{||}|} + \dots \tag{12}$$

as $|p_{||}| \to \infty$. $\mathcal{H}_\perp$ gives the contribution to the energy squared from transverse degrees of freedom:

$$\mathcal{H}_\perp = p_\perp^2 + M^2. \tag{13}$$

We obtain a non-relativistic dispersion relation with $|p_{||}|$ playing the role of the rest mass. On the other hand, the infinite boost causes an infinite Lorentz contraction of the string configurations in the $X^{||}$ direction. Hence, we can neglect longitudinal fluctuations of the string in this frame and concentrate on the purely transverse ones.

T-Duality and the Nonrelativistic Dispersion Relation

The general solution (before imposing boundary conditions) of the field equation is

$$X^\mu(\tau,\sigma) = X^\mu_L(\sigma^+) + X^\mu_R(\sigma^-),$$

where X_L represents left-moving traveling waves and X_R gives the right-moving ones. Both the equations of motion (10) and the constraints (11) are symmetric under a change of sign of the right-movers: $\partial_+ X^\mu \to \partial_+ X^\mu$, but $\partial_- X^\mu \to -\partial_- X^\mu$. In terms of the original coordinates:

$$\partial_\tau X^\mu \longleftrightarrow \partial_\sigma X^\mu \tag{14}$$

This innocent-looking field redefinition acquires a deep significance when we compactify the space where the string moves. Suppose we compactify the $X^{||}$ direction on a circle of radius R. Then the momentum becomes discrete with values

$$p_{||} = \frac{n}{R},$$

and the infinite boost limit corresponds either to $n \to \infty$ or $R \to 0$. Let us separate the oscillator part and the zero-mode part of the solutions of the equation of motion (10) and write

$$\partial_\tau X^\mu = a^\mu + \text{Oscillations}, \qquad \partial_\sigma X^\mu = b^\mu + \text{Oscillations},$$

where a^μ, b^μ are constant vectors. Hence the symmetry (14) acts on the zero modes as $a^\mu \leftrightarrow b^\mu$.

The interpretation of a^μ is the following; the momentum of the centre of mass is given as the standard conjugate variable to $\partial_\tau X^\mu$, that is

$$p_\mu = \int_0^\ell d\sigma \, \frac{\delta S_{\rm P}}{\delta \partial_\tau X^\mu} = \frac{1}{2\pi\alpha'} \int_0^\ell d\sigma \, \partial_\tau X_\mu = \frac{\ell}{2\pi\alpha'} \, a_\mu,$$

so that

$$a^\mu = \frac{2\pi\alpha'}{\ell} \, p^\mu,$$

where ℓ is a conventional normalization for the σ coordinate. On the other hand, for a *static* string that is stretched and winds w times in the $X^{||}$ direction we have

$$\partial_\sigma X^{||}_{\rm stretch} = b^{||} = 2\pi R \, \frac{w}{\ell}$$

Hence, the symmetry (14) acts on the quantized momentum and winding numbers as the combined operation

$$n \longleftrightarrow w, \qquad R \longleftrightarrow \frac{\alpha'}{R}$$

This symmetry is called T-duality and it is of the utmost importance. It simply means that the physics of quantized momentum modes on a circle of radius R is equivalent to the physics of quantized winding modes on a circle of radius α'/R.

The T-duality symmetry is deeply rooted in the extended character of the string and we will return to it later on. For now we will use it to point out that a large boost in a compact direction is equivalent to a large stretching in the dual circle. In the infinite boost limit, $R \to 0$, the dual system is that of an infinitely long string. In this case we can perform a non-relativistic quantization of the oscillation modes.

According to T-duality, the rest mass of a long winding string equals the large compact momentum of a boosted string in the dual circle. Hence, we shall quantize the relativistic string by computing the non-relativistic spectrum of transverse fluctuations for a heavy macroscopic string of rest mass M_0:

$$E = M_0 + \frac{\mathcal{H}_\perp}{2M_0} + \dots$$

and afterwards reinterpreting $M_0 \to |p_{||}|$ and $\mathcal{H}_\perp = p_\perp^2 + M^2$. We insist on the heuristic character of this procedure. Strictly speaking, it is no more than a useful mnemonic to obtain the right formulas. We have decided to derive the spectrum in this form because the non-relativistic quantization is interesing in itself and illustrates most of the physics issues in the general relativistic case.

Quantizing Transverse Fluctuations

Let us consider a macroscopic string of length L and rest mass

$$M_0 = \frac{L}{2\pi\alpha'}$$

It is easy to quantize the small oscillations about this equilibrium configuration. We can choose the world-sheet parameters so that

$$X^0 = \tau, \qquad X^{||} = \sigma.$$

This is the so-called static gauge. It is just incorporating the fact that longitudinal oscillations (in the parallel direction $X^{||}$) can be absorbed in a reparametrization, and that we can choose the world-sheet time equal to the physical time X^0. The induced metric is

$$(h_{ab}) = \begin{pmatrix} -1 + (\partial_\tau X_\perp)^2 & \partial_\tau X_\perp \cdot \partial_\sigma X_\perp \\ \partial_\tau X_\perp \cdot \partial_\sigma X_\perp & 1 + (\partial_\sigma X_\perp)^2 \end{pmatrix}.$$

Thus, expanding the determinant to quadratic order in derivatives gives

$$\sqrt{-\det(h_{ab})} = \sqrt{1 + (\partial_\sigma X_\perp)^2 - (\partial_\tau X_\perp)^2 + \dots} = 1 + \frac{1}{2}\left[-(\partial_\tau X_\perp)^2 + (\partial_\sigma X_\perp)^2) + \dots\right].$$

The Nambu–Goto action in the same approximation is

$$S_{\rm NG} \approx -\frac{1}{2\pi\alpha'} - \frac{1}{4\pi\alpha'}\int d\tau\, d\sigma\, (\partial X_\perp)^2 = -\frac{L}{2\pi\alpha'}\, X^0 - \frac{1}{2}\int (\partial\phi_\perp)^2,$$

where $\phi_\perp \equiv X_\perp/\sqrt{2\pi\alpha'}$. Therefore, we find that the small transverse oscillations are given by $d-2$ massless fields in $1+1$ dimensions. In momentum space, this is just a set of free oscillators with Hamiltonian

$$H \approx M_0 + \frac{1}{2}\int_0^L d\sigma \left[(\partial_\tau\phi_\perp)^2 + (\partial_\sigma\phi_\perp)^2\right] = \frac{L}{2\pi\alpha'} + \sum_\omega \omega\left(n_\omega + \frac{1}{2}\right).$$

In this formula, n_ω is the occupation number of each independent oscillator whereas

$$\frac{1}{2}\sum_\omega \omega$$

is the energy of zero-point fluctuations. This is infinite as usual and must be renormalized. However, the string has a finite length L, so that a part of this energy is physical and must be kept, i.e. the Casimir energy:

$$C = \frac{1}{2}\sum \omega(L) - \frac{1}{2}\sum \omega(L=\infty)$$

We use the normal prescription of subtracting the infinite Casimir energy at infinite volume. In this way, we renormalize away only the short-distance part of the divergence.

The $\phi_\perp$ are massless fields in $1+1$ dimensions. Thus, their normal excitations are naturally divided in right-movers with dispersion relation:

$$\omega_R = +k,$$

and left-movers:

$$\omega_L = -k,$$

whith k the momentum carried by the oscillation in the $X^{||}$ direction. In order to find the allowed values of the momentum we consider the case of open and closed strings separately.

Open Strings $\longrightarrow$ Gauge Fields

For open strings, right- and left-movers are reflected into one another at the end-points. Thus, the stationary states are standing waves on the open string. The allowed momenta are $|k| = 2\pi/\lambda$, with λ the allowed values of the wave-length. The nodes of the standing waves are separated by half a wave-length. Hence, this distance must be an integral fraction of the total length:

$$\frac{1}{2}\lambda = \frac{n}{L}, \qquad n = \text{positive} \quad \text{integer},$$

and we have found that

$$\omega = \frac{\pi n}{L}$$

for each of the $d-2$ transverse directions.

¿From here we may calculate the Casimir energy of the $d-2$ massless scalar fields with the following regularization:

$$C = \frac{d-2}{2}\cdot\frac{\pi}{L}\sum_{1}^{\infty} n\, e^{-\epsilon n/\sqrt{L}} - [\text{same with} \quad L=\infty] = -\frac{\pi(d-2)}{24L}$$

In this calculation, we have used the identity

$$\sum_n n\, e^{-\epsilon n/\sqrt{L}} = -\sqrt{L}\,\frac{d}{d\epsilon}\sum_n e^{-\epsilon n/\sqrt{L}} = -\sqrt{L}\,\frac{d}{d\epsilon}\left(\frac{1}{1-e^{-\epsilon/\sqrt{L}}}\right).$$

This regularization has been chosen so that the divergent term is independent of L and scales with ϵ^{-2}.

Adding a boost of momentum $p_\perp$ in the transverse directions and putting all pieces together we find

$$H = M_0 + \frac{\mathcal{H}_\perp}{2M_0} = M_0 + \frac{p_\perp^2}{2M_0} + \frac{\pi}{L}\left(\sum_{n>0} nN_n - \frac{d-2}{24}\right),$$

where

$$N_n = \sum_{i \in \perp} N_{n_i}$$

is the occupation number of the i-th transverse mode with frequency $\omega_{n_i} = \pi n_i / L$. Using now $M_0 = L/2\pi\alpha'$ we finally obtain the mass spectrum of a relativistic open string.

$$M^2 = \frac{1}{\alpha'} \left(N - \frac{d-2}{24} \right), \tag{15}$$

where we denote the total oscillator level by

$$N = \sum_n n N_n$$

Then, at $N = 0$ we have states labelled by the momentum $|p\rangle$ with no transverse oscillations, i.e. scalars with respect to the transverse $SO(d-2)$ rotation group. These states have mass squared

$$M^2 \, |p\rangle = -\frac{d-2}{24\alpha'} \, |p\rangle$$

They are tachyonic for $d > 2$. Consider now the first excitation level of the transverse oscillators, $N = 1$, i.e. states of the form

$$|p, i\rangle = (a_1^i)^\dagger \, |p\rangle.$$

They transform as a vector of $SO(d-2)$ with $d-2$ degrees of freedom. If the spectrum is to be Lorentz-invariant, they must fit in representations of the full Lorentz group $SO(1, d-1)$. This would be the case if the states $|p, i\rangle$ were the physical polarizations of a massless vector field.

To see this, recall that the number of physical polarizations of a massive particle in d spacetime dimensions corresponds to a representation of the *Little Group*, i.e. the symmetry that remains when we look at a particle at rest. This is the $SO(d-1)$ group of rotations in *spatial* directions. Hence, a massive particle is in a representation of $SO(d-1)$. For massless particles however, we cannot sit on top of them to have them at rest. The best we can do is to choose a special frame with all transverse momenta vanishing:

$$(p^\mu) = (1, 1, 0_\perp)$$

In this frame, we still have the $SO(d-2)$ group of *transverse* rotations. Thus, it is natural to classify massless particles by representations of this transverse group. Since $SO(d-2)$ is smaller than $SO(d-1)$, massless particles usually have less degrees of freedom than massive particles. For example, a massive vector in four spacetime dimensions has 3 polarizations from the $SO(3)$ group. However, a massless vector (a photon) has only 2 polarizations (of course, related by CPT).

Since the mass from (15) is

$$M^2 \, |p, i\rangle = \frac{26-d}{24\alpha'} \, |p, i\rangle,$$

we find that Lorentz invariance of the spectrum requires a critical dimension

$$d = 26$$

in which the states $|p, i\rangle$ represent the physical polarizations of a massless *photon* field.

The great significance of this result is that open strings generate gauge symmetry *dynamically*, even if such a symmetry was not postulated in the first place. At a very fundamental level, interacting theories of *massless vectors* always include a gauge symmetry. This is a consequence of the tension between Lorentz covariance and the small number of polarizations of massless particles. From this point of view, gauge symmetries are just useful redundancies that constraint the possible interactions of massless vectors.

This is easily seen by noticing that a vector field in a fully Lorentz-invariant description must involve a vector potential A_μ with d components, while we know that there are only $d-2$ physical polarizations of the photon. Hence, we need a prescription to project out the physical polarizations in a Lorentz-invariant way and conserve this projection through interactions. The standard solution of this problem is to impose a gauge symmetry, i.e. an equivalence of vector potentials under $A_\mu \to A_\mu + \partial_\mu \lambda$. In momentum space:

$$A_\mu(p) \to A_\mu(p) + \lambda\, p_\mu$$

We can fix the gauge symmetry by imposing the Lorentz gauge $\partial_\mu A^\mu = 0$, which means that the polarization must be transverse:

$$p^\mu\, A_\mu(p) = 0$$

This condition removes one polarization degree of freedom. Notice however that the transversality condition does not fix shifts by longitudinal polarizations of the form $\delta A_\mu \propto p_\mu$, provided the particle is massless and *on shell*: $p^2 = 0$.

Hence, one can reduce the physical components of massless vectors to the two polarizations orthogonal to the direction of motion. Since the gauge condition and the gauge transformation are Lorentz-covariant, we can be sure that unphysical polarizations are not generated in interactions provided the Lagrangian is gauge-invariant at the quantum level and the gauge-fixing procedure is Lorentz-invariant.

It is interesting in this respect that Maxwell's action can be "deconstructed" from the perhaps more obvious Klein–Gordon action if one is careful about gauge symmetry. An immediate guess for the action of a vector potential describing particles with massless dispersion relation $\omega = |\vec{p}\,|$ is the Klein–Gordon action for each component:

$$S_{\text{naive}} = \frac{1}{2} \int A_\mu\, \partial^2\, A^\mu .$$

This action is Lorentz-invariant and and has massless dispersion relations. However, since $\eta_{00} = -1$, the action of the A_0 component has an overall minus sign which leads to negative-norm states in the Hilbert space. We can circunvent this problem by enforcing the transversality conditions above. In order to maintain Lorentz-covariance we can work in terms of projected vectors:

$$A^\mu_P \equiv P^\mu{}_\nu\, A^\nu$$

where the projector is given by

$$P^\mu_\nu = \delta^\mu_\nu - \frac{\partial^\mu\, \partial_\nu}{\partial^2}$$

We can deal with this non-local projector by working in momentum space. The important property of the projector is that A^μ_P is gauge-invariant and transverse. It is then a simple exercise to show that the projected Klein–Gordon action is nothing but the ordinary Maxwell action:

$$S_{\text{Maxwell}} = \frac{1}{2} \int A_\mu^P \, \partial^2 \, A_P^\mu = -\frac{1}{4} \int F_{\mu\nu} \, F^{\mu\nu}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

In summary, we have found that open strings naturally lead to massless gauge fields in extra ($d = 4 + 22$) dimensions. In addition, they tend to have tachyonic ground states.

Boundary Conditions and D-branes

In our derivation of the open-string spectrum we have concentrated on the oscillation degrees of freedom and have been careless about the boundary conditions at the endpoints of the open string. The simplest assumption about the physical properties of the endpoints is that they have no mechanical degrees of freedom on their own (for example no rest mass). This means that we have to choose boundary conditions at $\sigma = 0, \ell$ that are compatible with stationarity of the action for arbitrary variations of the field X^μ. From the basic action we have

$$\delta \, S_{\text{P}} = \frac{1}{2\pi\alpha'} \int d\tau \, d\sigma \, (\partial_\tau X^\mu \, \partial_\tau(\delta X_\mu) - \partial_\sigma X^\mu \, \partial_\sigma(\delta X_\mu)) \, .$$

Integrating by parts and dropping total time derivatives we obtain

$$\delta \, S_{\text{P}} = -\frac{1}{2\pi\alpha'} \int d\tau \, d\sigma \, \delta X^\mu \, (\partial_\tau^2 - \partial_\sigma^2) \, X_\mu - \frac{1}{2\pi\alpha'} \int d\tau \, [\delta X^\mu \, \partial_\sigma X_\mu]_{\sigma=0}^{\sigma=\ell}$$

The first term yields the usual equation of motion (10) and the second term yields the so-called Neumann boundary conditions:

$$\partial_\sigma \, X^\mu \, |_{\text{endpoints}} = 0$$

However, there is another natural boundary condition that also makes the action stationary. We can declare that the endpoints, rather than moving freely, are *confined* to a fixed position (in a given direction). This corresponds to

$$\delta X \, |_{\text{endpoints}} = 0,$$

i.e. X(endpoints) is constant in time (Dirichlet boundary condition). We can in fact combine both and consider some directions with Neumann boundary conditions and the others with Dirichlet boundary conditions:

$$\partial_\sigma \, X^N \, |_{\text{endpoints}} = 0, \qquad \partial_\tau \, X^D \, |_{\text{endpoints}} = 0. \tag{16}$$

The result is a defect in space-time, a hyperplane defined by X^D = constant on which open strings have confined endpoints and otherwise move freely in the Neumann directions. If we have p spatial Neumann directions, we say that we have a Dp-brane.

The oscillation spectrum (15) is insensitive to the boundary conditions on the endpoints. On the other hand, the distinguished hyperplane breaks the full Lorentz symmetry $SO(1, d-1)$ to the Lorentz group in the Neumann directions $SO(1, p)$ times the transverse Lorentz group $SO(d - 1 - p)$. Therefore, the d-dimensional vector field A_μ splits into a $(p + 1)$-dimensional vector A_N, plus a set of $d - 1 - p$ scalars $A_D = \Phi_D$. All fields are built from open-strings and therefore are confined to the D-brane. This means that they are functions of x^N only and the d-dimensional field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ decomposes as

$$F_{NN'} = \partial_N A_{N'} - \partial_{N'} A_N, \qquad F_{ND} = \partial_N \Phi_D, \qquad F_{DD'} = 0\,,$$

and the Maxwell action:

$$\int d^d x\, F_{\mu\nu}\, F^{\mu\nu} \longrightarrow \int d^{p+1}x \left(F_{ab}\, F^{ab} - 2 \sum_D (\partial_a \Phi_D)^2 \right), \tag{17}$$

where $a, b = 0, 1 \ldots, p$. The presence of massless scalars could have been foreseen from the fact that the D-brane breaks translational invariance along all the transverse directions, so that we expect the corresponding Goldstone bosons. The gauge coupling of the gauge fields on a D-brane $g_{\mathrm{D}p}$ comes from open-string diagrams, so that $g_{\mathrm{D}p}^2 \sim g_o^2 \sim g_s$, and must have mass dimension $4 - (p + 1) = 3 - p$. Hence, on dimensional grounds

$$g_{\mathrm{D}p}^2 \sim g_s\, \ell_s^{p-3}$$

D-branes give a novel mechanism to generate non-abelian gauge symmetry. If we consider N parallel Dp-branes on top of each other one can have strings that strecth between different D-branes as well as strings with both endpoints on a given D-brane, for a total of N^2 species. Therefore we have N^2 massless vectors. Since these vectors interact nontrivially through higher order string diagrams, they must furnish a non-abelian gauge theory. The simplest setting generates $U(N)$ gauge symmetry, although $SO(N)$ and $Sp(N)$ are not difficult to generate as well.

In this case, the non-abelian Yang–Mills action on the world-volume descends from the d-dimensional Yang–Mills action in terms of $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ via

$$F_{ND} = D_N\, \Phi_D, \qquad F_{DD'} = [\Phi_D, \Phi_{D'}],$$

with $D_N = \partial_N + [A_N, \quad]$ the covariant derivative in the adjoint representation. We end up with a Yang–Mills theory in $p + 1$ space-time dimensions with $d - p - 1$ scalar fields in the adjoint representation with a potential proportional to $[\Phi, \Phi]^2$.

D-brane Mass and Size

A further important property of D-branes is that they have a finite mass. If we consider the gravitational interaction between two identical heavy branes at some transverse distance of each other, the gravitational force is proportional to $G_{\mathrm{N}}\, M^2$. On the other hand, the leading gravitational interaction comes from the exchange of a single graviton, i.e. a single closed string. This is a diagram with the topology of a cylinder whose boundaries are mapped to the D-branes. The power of the string coupling is thus $g_s^0 = 1$. Since $G_{\mathrm{N}} \sim g_s^2$ we learn that the mass per unit volume, or tension of the Dp-brane scales like

$$T_{\mathrm{D}p} \sim \frac{m_s^{p+1}}{g_s}$$

Incidentaly, notice that the region of spacetime where the gravitational field of such an object is significative is the scale of the Schwarzschild radius

$$R_{\mathrm{grav}} \sim (G_{\mathrm{N}}\, T_{\mathrm{D}p})^{d-3-p} \sim g_s\, \ell_s$$

Therefore, in the $g_s \to 0$ limit, the perturbative realm, the gravitational radius goes to zero in string units. This is the reason why a very heavy object such a D-brane manages to appear as a sharp defect of zero thickness in string perturbation theory.

The D-brane tension decouples in the perturbative limit $g_s \to 0$. This implies that they are nonperturbative states in the Hilbert space of the closed string theory. Unlike other solitonic objects in QFT, such as 't Hooft–Polyakov monopoles whose mass scales like g^{-2}, D-branes are comparatively lighter at weak coupling. If D-branes were to mediate some semiclassical tunneling process, the WKB amplitude would be proportional to $\exp(-C/g_s)$ for some constant C. These are stronger than typical nonperturbative effects in QFT. In Problem 6 we argue that this is related to a stronger growth of perturbative amplitudes at high orders.

D-branes and T-duality

Finally, we comment on the interplay between D-branes and T-duality. Since D-branes are defined in terms of boundary conditions for open strings, it is easy to track the action of T-duality for a compactification on a circle of radius R.

If the compact direction is of Neumann type, i.e. the D-brane is wrapped on the circle, T-duality converts momentum modes into winding modes of the dual circle of radius α'/R. In order for an open string to have winding modes its endpoints must be fixed. Therefore we find that T-duality maps Neumann directions into Dirichlet directions and viceversa. In view of (16), this is also obvious from the action of T-duality on the world-sheet fields (14).

Hence, a Dp-brane wrapped on a circle of radius R is T-dual to a D$(p-1)$-brane localized on the dual circle of radius α'/R. Since we have realized T-duality as a symmetry of the D-brane spectrum, it must be true that the tension of the localized D$(p-1)$-brane is equal to the tension of the wrapped Dp-brane times the wrapping length $2\pi R$. This gives a relation

$$2\pi R \, \frac{m_s^{p+1}}{g_s} = \frac{\tilde{m}_s^p}{\tilde{g}_s},$$

where the tilded mass and coupling correspond to the theory after T-duality. Under T-duality $\alpha' = 1/m_s^2$ is invariant. Therefore, we find that the string coupling constant *does* transform under T-duality as

$$g_s \to g_s \, \ell_s / R$$

T-duality can also be used to obtain a considerable improvement over the low-energy Maxwell Lagrangian for the $U(1)$ gauge field (17). The minimal coupling of a gauge field is described by the covariant derivative or $-i\partial_\mu + A_\mu$. In momentum space

$$p_\mu \to p_\mu + A_\mu$$

Therefore, a constant vector potential induces a shift of the momentum given by

$$p = \frac{n}{R} \to \frac{n}{R} + A$$

Under T-duality,

$$\frac{n}{R} \longrightarrow \frac{w\,R}{\alpha'} = \frac{1}{2\pi\alpha'}\, 2\pi R\, w$$

Therefore, $2\pi\alpha' A$ is a fractional shift of the total winding length $2\pi R w$. This corresponds to the endpoints of the open string being separated a distance $2\pi\alpha' A$ on the circle.

The natural interpretation of this is that the zero-mode of the gauge field A maps under T-duality to the transverse position of the D-brane. Hence, the scalars that result from the d-dimensional gauge field are

$$\Phi_D = \frac{\mathrm{X}^D}{2\pi\alpha'}$$

Consider now a heavy D-brane that moves slowly. The relativistic action of such heavy object contains the usual term:

$$S = -T_{\mathrm{D}p} \int dt \sqrt{1 - (\partial_t \mathrm{X}^D)^2},$$

where $\partial_t X^D$ is nothing but the transverse velocity in the Dirichlet directions. Now, under T-duality the tensions of the different Dp-branes map into each other and $\mathrm{X}^D = 2\pi\alpha'\,\Phi_D \to 2\pi\alpha' A_N$. In particular this means that the velocity $\partial_t X^D \to 2\pi\alpha'\,\partial_t A_N$ and is interpreted as the N-th component of the electric field on the D$(p+1)$-brane.

Hence, from the non-linear action above we can find the non-linear completion of the Maxwell action that is compatible with T-duality and Lorentz invariance. It is the so-called Dirac–Born–Infeld action:

$$S_{\mathrm{DBI}} = -T_{\mathrm{D}p} \int d^{p+1}x \sqrt{-\det\left(\eta_{ab} + 2\pi\alpha'(F_{ab} + \partial_a \Phi^D \partial_b \Phi_D\right))}$$

This action is good provided we neglect the space-time variation of the field strength (see Problem 7).

Closed Strings $\longrightarrow$ Gravity

After doing the hard work for the open-string case, closed strings are an easier matter. The closed string is like an open string for which the end-points have been joined together. This means that vibrational modes are no longer reflected and travel independently. There are however two subtleties that have to be dealt with.

First, the allowed frequencies are still

$$\omega = |k|,$$

with $|k|$ the allowed momenta. But these are determined now by the condition that the oscillations are periodic on the closed string, i.e. on the size of the closed string L we must fit an integer number of wavelengths and we have

$$\omega_n = \frac{2\pi n}{L}, \qquad n = \text{ positive integer}$$

This gives an overall factor of 2 in the spectrum with respect to the open-string case.

There is also a constraint on the total oscillation number of left- versus right-movers, coming from the fact that the soldering point is arbitrary. This means that the states must be invariant under a translation of the origin of the string coordinate

$$\sigma \to \sigma + \text{constant}$$

Notice that we require *invariance* rather than *covariance* of the states under σ-translations, because this is just a reparametrization of the world-sheet, and these were *gauge* symmetries in the Polyakov action.

The Noether charge for σ-translations is the two-dimensional momentum in the σ direction:

$$P_\sigma = \sum_k k\, a_k^\dagger \cdot a_k = \sum_{n>0} \frac{2\pi n}{L} \left(N_n^L - N_n^R\right) = \frac{2\pi}{L} \left(N^L - N^R\right),$$

where we have only summed the contribution of oscillations. We assume that the centre of mass of the string has no momentum in the longitudinal direction, since this can be added at the end by a Lorentz boost.

On physical states we thus require that the total oscillation number from right- and left-movers is equilibrated:

$$(N^L - N^R)\,|\Psi\rangle = 0.$$

This is called the *level matching constraint.* Therefore, the Hilbert space of a closed string is just the direct product of two separate open-string Hilbert-spaces divided by the level matching constraint:

$$\mathcal{H}_{\text{closed}} = \frac{\mathcal{H}^L_{\text{open}} \otimes \mathcal{H}^R_{\text{open}}}{\text{Level Matching}}$$

The formula for the spectrum is:

$$M^2 = \frac{2}{\alpha'} \left(N^L + N^R - \frac{d-2}{12}\right), \tag{18}$$

supplemented by the level matching constraint. Hence, we can also write it as

$$M^2 = \frac{4}{\alpha'} \left(N - \frac{d-2}{24}\right),$$

with N the total *left-moving* oscillator level.

We now study the low-lying levels of this spectrum. The ground state with no oscillation is again tachyonic for $d > 2$:

$$M^2\,|p\rangle = -\frac{d-2}{6\alpha'}\,|p\rangle$$

The first oscillation level corresponding to states of the form

$$|p,t\rangle \equiv t^{ij}\,|p,ij\rangle = t^{ij}\;\; (a_1^i)_L^\dagger\,(a_1^j)_R^\dagger\,|p\rangle,$$

are characterized by an arbitrary rank-two polarization tensor t_{ij}. They are massless at the critical dimension $d = 26$:

$$M^2\,|p,t\,\rangle = \frac{26-d}{6\alpha'}\,|p,t\,\rangle$$

Under the transverse rotation group $SO(d-2)$, the general rank-two tensor decomposes into a symmetric-traceless part, an antisymmetric part, and a pure trace describing a scalar degree of freedom:

$$t^{ij} = \frac{1}{2}\left(t^{ij} + t^{ji} - \frac{2}{d-2} t^i_i\right) + \frac{1}{2}\left(t^{ij} - t^{ji}\right) + \frac{1}{d-2}\,t^i_i.$$

We call the symmetric-traceless part h_{ij}, the antisymmetric part b_{ij} and the scalar ϕ. Thus we have symbolically:

$$[t_{ij}] = [h_{ij}] \oplus [b_{ij}] \oplus [\phi]$$

This is the universal massless spectrum of closed-string theories. The scalar is called the *dilaton* and plays an special role in the quantum expansion of string theory. The antisymmetric tensor field descends from a full antisymmetric tensor field in d dimensions $b_{\mu\nu}$ with gauge invariance:

$$b_{\mu\nu} \to b_{\mu\nu} + \partial_\mu \lambda_\nu - \partial_\nu \lambda_\mu$$

This gauge symmetry can be partially fixed by a transversality condition:

$$\partial^\mu b_{\mu\nu} = 0.$$

One can still shift the polarization on-shell (at $p^2 = 0$) so that only the transverse components b_{ij} remain. This is entirely analogous to the discussion of the massless vector in the open-string sector.

The symmetric traceless tensor h_{ij} descends from a symmetric tensor in d dimensions $h_{\mu\nu}$. Now the gauge invariance is

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \lambda_\nu + \partial_\nu \lambda_\nu$$

There is a subtlety in that the dilaton mixes with the trace of the symmetric tensor field. This is studied in Problem 2.

The result is that $h_{\mu\nu}$ represents quantized perturbations of the space-time metric (quantized gravitational waves) and thus corresponds to the graviton. Just as we pointed out before in the case of open strings, this fact is rather remarkable. The string is quantized as an elementary object with an intrinsic mass scale, and one obtains, among other things, an interacting theory of gravitons (general relativity) modified at short distances in a consistent way.

To be precise, one does not get exactly general relativity, but a close cousin of it. In Problem 2 we argue that the tree-level low-energy effective Lagrangian of massless closed-string modes is proportional to $\exp(-2\phi)$ where ϕ is the dilaton:

$$S_{\text{eff}} = \int e^{-2\phi} \left(\mathcal{R} + 4(\partial\phi)^2 + \dots \right) \tag{19}$$

Since we already knew that the effective action was proportional to $1/g_s^2$, we learn that the expectation value of the dilaton controls the string coupling constant:

$$g_s = \exp(\langle \phi \rangle).$$

This is a very important result. It implies that there are no adjustable parameters in string theory. The strength of the interactions is determined by the dynamics via the vacuum value of a scalar field in the string spectrum.

Notice that (19) gives a low-energy modification of Einstein's gravitational theory with an effective Newton's constant depending on the expectation value of a massless field (this is one of the Brans–Dicke type of theories). Such modification of gravity is severely constrained experimentally. Hence, the dilaton should become massive by quantum corrections of appropriate size. In supersymmetric string theories the dilaton is usually massless to all orders in perturbation theory, but can obtain a potential at a nonperturbative level. In fact, the generation of a potential for the dilaton is a generic consequence of supersymmetry breaking.

Since the dilaton determines Newton's constant together with all other gauge couplings of the low-energy physics, the calculation of the effective dilaton potential is of the utmost importance. Unfortunately, no systematic method exists to compute this effective potential after supersymmetry breaking.

Even if the dilaton should obtain a mass at low energy, the Lagrangian (19) does describe the dilaton-graviton sector of supergravity models that serve as low-energy limits of string theories with unbroken supersymmetry. As such, it has applications to very early universe cosmology or to the study of nonperturbative duality symmetries.

Lectures 3 & 4

Dynamical Aspects of String Theory

In the remainder of these notes we discuss a number of separate topics related to dynamical questions. Most of the issues we touch upon are related to fundamental open problems in string theory.

We start by introducing supersymmetry on the basis of the consistency of string models at tree level. We remain very qualitative at this point because the machinery of supersymmetry was not part of the knowledge assumed for these lectures. We finish by drawing some general patterns on the energy scales that are expected to be relevant in the structure of gauge interactions in superstring models and we introduce the connection between strings and black holes on the basis of the high-energy behaviour of perturbative string states.

The Problem of the Tachyon

One generic feature of the open and closed string spectra found before is the negative mass-squared of the ground states. This means that the lowest mass field, $T(x)$, is a tachyon in either string spectrum. Its low-energy effective Lagrangian has the form

$$\mathcal{L}_{\text{Tachyon}} = -\frac{1}{2}\,(\partial T)^2 - V(T) + \dots$$

with

$$\frac{\partial^2 V}{\partial T^2}(T=0) = M^2 < 0.$$

That is, the perturbative classical vacuum state with vanishing expectation value $\langle T \rangle = 0$ corresponds to a local maximum of the potential and therefore it is *unstable*. While we may suspect that $V(T)$ has a stable local minimum at some $\langle T \rangle \neq 0$, our ability to calculate non-linear terms to $V(T)$ is limited by our perturbative methods. In general, the dynamical relaxation of tachyonic instabilities is an open problem in string theory.

Another possible strategy is to look for *stable* ground states of string theory by adding extra degrees of freedom to the string world-sheet beyond the minimal bosonic complement. In this context, an interesting question is the following: what is the minimal modification of the bosonic string that achieves stability with the same "good" features of the massless spectrum (namely gravity and gauge fields)?

The standard answer to this question is based on making strings supersymmetric. We must emphasize that a satisfactory solution of this problem is not guaranteed to exist *a priori*. Recall that both gauge fields and gravitons arise from the first oscillation level of the open and closed strings respectively. Therefore, these excitations have *positive* oscillation energy. What makes them massless is the cancellation of this oscillation energy against the ground-state Casimir energy. This means that the negative-definite Casimir energy is responsible for *both* the good and the bad properties of the bosonic string.

Unfortunately, the construction of supersymmetric string theories is still a rather technical matter. Here we will content ourselves with the introduction of the main ideas and their implications.

Susy in One Page

The essential idea of supersymmetry is that there is a conserved charge Q, i.e. it commutes with the Hamiltonian:

$$[\,Q,\,H\,] = [\,Q^\dagger,\,H\,] = 0,$$

such that its action, if non-vanishing, converts bosonic states into fermionic states:

$$Q\,|\,\mathrm{Bose}\,\rangle = |\,\mathrm{Fermi}\,\rangle,$$

and viceversa for $Q^\dagger$. For *free* field theories, such as the ones that arise in the free approximation to string theory, the implementation of supersymmetry is rather trivial. Since a free field theory is an assembly of free harmonic oscillators with frequency $\omega = \sqrt{\vec{p}^{\,2} + m^2}$, the supersymmetric generalization simply involves adding one fermionic oscillator for each bosonic oscillator. In terms of two pairs of creation and annihilation operators:

$$[a, a^\dagger] = 1, \qquad [a, a] = [a^\dagger, a^\dagger] = 0,$$

for the bosons and

$$\{b, b^\dagger\} = 1, \qquad \{b, b\} = \{b^\dagger, b^\dagger\} = 0$$

for the fermions, the Hamiltonian is

$$H = \frac{1}{2}\,\omega_B\,(a^\dagger\,a + a\,a^\dagger) + \frac{1}{2}\,\omega_F\,(b^\dagger\,b + b\,b^\dagger)$$

Therefore, the spectrum is given by

$$E = \omega_B\,\left(N_B + \frac{1}{2}\right) + \omega_F\,\left(N_F - \frac{1}{2}\right)$$

for N_B, N_F the occupation numbers for bosons and fermions respectively. Clearly, in order to have a symmetry between bosons and fermions we need

$$\omega_B = \omega_F = \omega,$$

which leads to a cancellation of zero-point fluctuations of the oscillators and

$$E = \omega\,(N_B + N_F)$$

The supersymmetry charge operators are

$$Q = \sqrt{\omega}\,a^\dagger\,b, \qquad Q^\dagger = \sqrt{\omega}\,b^\dagger\,a,$$

and their algebra is given by

$$\{ Q, Q^\dagger \} = H. \tag{20}$$

This anticommutator is the essence of supersymmetry. It means that supersymmetry transformations square to time translations. In a Lorentz-invariant theory there must be other supercharges so that the anticommutator is Lorentz-covariant. Also, because of the spin-statistics connection the supercharges must have half-integer spin. This leads to the usual algebras

$$\{ Q_\alpha, Q^\dagger_\beta \} = \gamma^\mu_{\alpha\beta}\, p_\mu + \dots ,$$

with Q_α in a spinor representation of the Lorentz group. The constants $\gamma^\mu_{\alpha\beta}$ appear here as Clebsch–Gordan coefficients for the decomposition of a product of two spinorial representations. They are of course the Dirac matrices. In four dimensions, since the minimal spinor (Weyl or Majorana) has four real components, there are at least four independent supercharges in a Lorentz-invariant four-dimensional model. This is the so-called $\mathcal{N} = 1$ or minimal supersymmetry.

One essential property of the basic commutator (20) is that its expectation value on any state is necessarily positive if the Hilbert space only admits positive-norm states:

$$0 \leq \langle \psi \,|\{Q, Q^\dagger\}|\, \psi \rangle = |Q\,|\psi\rangle|^2 + |Q^\dagger\,|\psi\rangle|^2 = \langle \psi \,|\, H \,|\, \psi\rangle \tag{21}$$

Therefore, we find that in a supersymmetric system of the type constructed here (in terms of free Fock spaces) the energy is positive definite and it vanishes if and only if the state is *supersymmetric*, i.e. it is invariant under the action of Q. For the oscillator system this is the oscillator vacuum. Notice that given a generic state of positive energy $|\omega\rangle$, it is paired with $Q|\omega\rangle$ at the same energy, this is called a *supermultiplet*. This is true for all states except for the ground states: since $Q|\omega = 0\rangle = 0$, vacua do not need to be paired since they are supersymmetric *singlets*.

Equation (21) also shows that supersymmetry is spontaneously broken if and only if the ground state (vacuum) has strictly positive energy.

Susy and Particle Phenomenology

Supersymmetry is attractive for particle phenomenology because it reconciles the use of weakly-coupled scalars (Higgs) in the mechanism of electroweak symmetry breaking, with the generic quadratic renormalization of these scalars. This is the so-called *hierachy problem*.

For the fermions of the Standard Model, masses are only logarithmically renormalized, so that the contribution to the physical mass coming from vacuum fluctuations is of order

$$\delta\, m_f \sim \alpha\, m_f\, \log(M_X/m_f),$$

where $\alpha = g^2/4\pi$ is a typical fine-structure constant and M_X is some high-energy scale at which the point-like description of the fermion breaks down. The logarithmic renormalization is a consequence of the chiral symmetry at $m_f = 0$ that forbids any perturbative mass renormalization at the massless limit. Given the actual masses of quarks and leptons and the value of the gauge couplings, we have $\delta\, m_f \sim m_f$ with M_X in the vicinity of the Planck scale. If we come up with some way of breaking chiral symmetry weakly, with $m_f/M_X \ll 1$, then we have a very satisfactory resolution of the XIX century crisis of the self-energy of the electron!

In the Standard Model, the electroweak gauge symmetry is chiral, so that the mass of the fermions is naturally associated to the scale of gauge symmetry breaking. Our comment here is that the vacuum fluctuations do not upset this picture and we can actually think of the fermion masses in a rather physical way.

On the other hand, if we have point-like scalars such as Higgs fields, the typical contribution of vacuum fluctuations to their mass depends quadratically on the cutoff:

$$\delta\, m_h^2 \sim \alpha\,(M_X^2 - m_h^2).$$

If we want the Higgs fields to effect the electroweak symmetry breaking *within weak coupling* then m_h cannot be much larger than M_W and we have $\delta\, m_h^2 \sim \alpha\, M_X^2 \gg m_h^2$. Hence, the vacuum-fluctuations contribution to the Higgs mass is huge and must be cancelled with extreme precission by the bare mass at the high scale.

Supersymmetry solves this problem by restoring the logarithmic running for the scalar mass. The quadratic contribution is cancelled by bosons and fermions running in the virtual loops. Since supersymmetry must be broken, with splittings of $O(M_{ss})$ between superpartners, the contribution to the mass of the Higgs after supersymmetry breaking is of order

$$\delta\, m_h^2 \sim \alpha\, M_{ss}^2$$

Hence, if we do not want the fine-tunning problem back, we need M_{ss} to be not too much larger than m_h, and we expect that if supersymmetry has anything to do with the ratio $M_W/M_P \sim 10^{-16}$, then superpartners should appear below $O(\text{TeV})$ energies.

Besides stabilizing the hierarchy M_W/M_P under quantum corrections, supersymmetry is also attractive in generating the hierarchy dynamically. The most natural method known to generate mass hierarchies is based on the logarithmic running of four-dimensional gauge couplings. With logarithmic accuracy, the relation between fine-structure constants at different energy scales is

$$\frac{1}{\alpha(\mu)} = \frac{1}{\alpha(M_X)} + \frac{\beta_0}{2\pi}\,\log\left(\frac{\mu}{M_X}\right), \tag{22}$$

where β_0 is a numerical coefficient that depends on the gauge group and the matter content. If $\mu < M_X$, models with $\beta_0 > 0$ are asymptotically free, i.e. the effective coupling grows towards low energies. If the theory develops strong coupling at $\mu \sim M_{\text{strong}} < M_X$ we have

$$M_{\text{strong}} \sim M_X \;\; \exp\left(-\frac{2\pi}{\beta_0\,\alpha(M_X)}\right).$$

This formula is just a rough estimate obtained by putting $\alpha(M_{\text{strong}}) = \infty$ in (22). It shows that the low-energy scale M_{strong} depends on the high-energy data M_X and $\alpha(M_X)$ non-perturbatively on the coupling. This means that a moderately small high-energy coupling can induce huge hierarchies of masses.

In fact, this is the most natural explanation of the ratio between the proton mass and the Planck mass. The 1 GeV mass of the proton is mostly due to the highly relativistic binding energy of the quarks and gluons, rather than to the bare masses. Hence the mass of the proton is tied to the non-perturbative effects associated to strong SU(3) gauge coupling (confinement and chiral symmetry breaking).

This leads to the natural idea that M_W should be associated to some strong-coupling phenomena, i.e. the mass of Standard Model particles would be zero in perturbation theory, and the small ratio M_W/M_P would be a consequence of these masses being generated by small non-perturbative effects. While it is technically easy to have massless vectors and fermions (via gauge and chiral symmetry constraints, such as those arising in the Standard Model), it is much more difficult to have naturally massless charged scalars, such as Higgs bosons. By relating scalars to chiral fermions, supersymmetry explains the lightness of charged scalars, either at tree level or under perturbative corrections. Hence, in the context of low-energy supersymmetry, one tries to relate M_W/M_P to the dynamics of supersymmetry breaking, something that involves nonperturbative physics at some (perhaps intermediate) stage.

The major dilemma that will be (hopefully) resolved by LHC is whether the breaking of $SU(2) \times U(1)$ in the Standard Model is associated to a weakly-coupled scalar (a Higgs field), a situation that would "naturally" call for supersymmetry, or there are no light scalars at all and the symmetry breaking is entirely due to strong-coupling phenomena. This "minimalistic" alternative is very attractive theoretically, although no compelling models could be constructed to date (the technicolor saga).

This means that of all the phenomena that are associated with strings, supersymmetry is the one that could be tested experimentally in the near future. In fact, *low-energy* supersymmetry is independent of whether supersymmetry plays a role at very high energies in the stringy domain. But the story as presented here is certainly suggestive.

This scenario has important theoretical problems though. They are all related to the fact that, while supersymmetry is attractive in generating the small M_W/M_P ratio, we already know that the supersymmetry-breaking dynamics cannot be completely generic. This is true even when we explicitly arrange for exotic processes such as proton decay to be suppressed. The reason is that the Standard Model has other unnaturally small parameters (not explained by renormalization-group arguments) whose values can be upset by a generic supersymmetry breaking dynamics at the TeV scale. There are unnaturally small ratios of masses of quarks and leptons. Consider for example m_e/m_t or even better m_ν/m_t (this is the most naive formulation of the so-called flavour problem). There is also a very small contribution of strong interactions to CP violation (the so-called strong CP problem $|\theta| < 10^{-9}$). Even more striking is the mysteriously small value of the cosmological constant.

A consequence of (21) is that generation of positive vacuum energy is always associated to spontaneous supersymmetry breaking. Since the vacuum energy density (cosmological constant) is observed to be almost zero:

$$\rho_{\text{vac}} \approx \left(10^{-15}\right)^4 \quad (\text{TeV})^4,$$

while the scale of supersymmetry breaking should be about 1 TeV, this looks like a huge problem. Fortunately, in the presence of gravity, the left hand side of (21) gets corrections of $O(\kappa^2)$. These corrections are *negative* definite, so that one can cancel both effects and be left with a small energy density after supersymmetry breaking. The problem remains, though, that this cancellation is another fine-tunning for which we have no explanation. Supersymmetry does not help because this is happening after supersymmetry breaking.

In fact, in a typical model of supersymmetry breaking, the *ad hoc* adjustement needed to fit ρ_{vac} will also affect the predictions for the spectrum of massess and mixings at the supersymmetry breaking scale M_{ss}, as well as the strength of CP violation from different sectors.

There is no single model of low-energy supersymmetry breaking that meets all phenomenolgical constraints in an "elegant" way, especially after large portions of the parameter space of the simplest models are already excluded experimentally.

This means that if supersymmetry is found at the TeV scale, the study of supersymmetry-breaking dynamics will be a fascinating journey in which experiment will be considerably ahead of theory.

Superstrings

Coming back to the string scale, at the level of the free approximation ($g_s = 0$) in flat Minkowski space, a supersymmetric string spectrum must have H real and positive, and negative mass-squared states are excluded. Also, notice that the origin of the tachyon in bosonic string theories was the negative Casimir energy of the finite-size string. In a supersymmetric model this Casimir energy vanishes by cancellation of bosonic and fermionic zero-point energies. Therefore, the spectrum reads

$$M^2 = \frac{c}{\alpha'} \, (N_B + N_F) \geq 0,$$

where N_B and N_F are the total transverse oscillator level

$$N = \sum_\omega n_\omega$$

in bosonic and fermionic oscillators of the open string or the total oscillator level of *left-moving* oscillators of the closed string. Then $c = 1$ for open strings and $c = 4$ for closed strings.

The massless spectrum corresponds now to vanishing oscillator levels $N_B = N_F = 0$. If we want the massless spectrum to be supersymmetric and contain the massless vector in the open-string case and the massless graviton in the closed-string case, we need to build the states from the zero-modes of the fermionic oscillators.

Consider the open-string case for definiteness. The ordinary fermionic oscillators of non-zero frequency have the property

$$b_\omega^\dagger = b_{-\omega}$$

where b_ω comes from a Fourier mode of negative frequency of the fermionic world-sheet fields. Canonical anticommutation relations then yield the usual relation

$$\{ \, b_\omega, \, b_{-\omega} \, \} = 1$$

The zero-frequency oscillators are hermitian $b_0 = b_0^\dagger$ and the same canonical commutation relations imply:

$$\{ \, b_0^\alpha, \, b_0^\beta \, \} = \delta^{\alpha\beta} \tag{23}$$

This is a Clifford algebra, and therefore the b_0 can be represented as Dirac matrices. Supersymmetry demands that the index α must take $d-2$ values, because these oscillators must be in one-to-one correspondence with the $d-2$ bosonic oscillators that describe transverse vibrations of the string. Hence, the dimension of the representation is

$$\dim \, (\text{Dirac}) = 2^{\left[\frac{d-2}{2}\right]},$$

where the brackets denote the integer part. These states must contain the $d-2$ polarizations of the massless vector field plus the same number of fermionic superpartners. So we have

$$d - 2 = \frac{1}{2} \, 2^{[(d-2)/2]},$$

which is solved for

$$d = 10$$

This is the critical dimension of supersymmetric string theories.

Another argument for $d = 10$ can be given based on the spin-statistics connection. The fermionic oscillators create fermionic states and these must be in spinor representations of the transverse group of rotations $SO(d-2)$. This implies that the index α in (23) is a spinor index. On the other hand the same fermionic oscillators are related by supersymmetry to the $d-2$ bosonic oscillators a_i with a *vector* index. This means that, in order to have supersymmetry acting consistently both on the space-time and on the world-sheet, we need some special symmetry between spinor and vector representations of $SO(d-2)$. The "magical" case is $SO(8)$ that has 3 fundamental representations related by a "triality" symmetry (a symmetry of the Dynkin diagram that permutes the representations). These representations are the vector and the two chiral spinor representations, all of dimension eight.

The massless spectrum of the open-string theory contains a massless vector A_μ and a spinor with the same number of on-shell degrees of freedom, namely 8. A Majorana spinor in $d = 10$ has 32 real components. In ten dimensions it is possible to enforce a Weyl projection on top of the Majorana reality condition. This yields 16 components. Furthermore, the Dirac equation relates half of them to the other half and this finally gives 8 components. So we have a massless vector and a Majorana–Weyl spinor. The number of supercharges is the number of components of the Majorana–Weyl spinor, i.e. 16, or $\mathcal{N} = 4$ in four-dimensional terms.

For closed strings we have the same structure with the product of left- and right-movers. The result are the multiplets of ten-dimensional supergravity.

There are 32 supercharges from

$$Q_L \oplus Q_R,$$

where Q_L and Q_R are the respective supercharges carried by left- and right-moving excitations. Depending on whether these supercharges are chosen with opposite or the same Weyl handedness we have the two ten-dimensional closed superstring theories with maximal supersymmetry, the so-called type IIA and type IIB theories.

Ten-dimensional string theories with 16 supercharges are considerably more complicated, partially as a result of various anomalies that require special constructions, such as multi-D-branes and non-orientable (orientifolds) open strings and closed-strings where right- and left-movers live in different spaces (heterotic strings). These special mechanisms induce large non-abelian gauge groups with chiral representations, a fact that makes them ideal starting points for phenomenological models.

D-branes as Partially Supersymmetric States

Remember that open strings were equivalent to D-branes, which in turn were to be interpreted as topological defects in the theory of closed strings. From this point of view the existence of half as many supersymmetries in the theory of open strings as there are in the theories of closed strings can be interpreted by saying that the D-brane ground states are annihilated by half of the supercharges of the closed string theory.

The unbroken supercharges that act on open-string excitations are a linear combination of those carried by left- and right-movers

$$Q_{\text{open}} = Q_L + \Gamma_{\text{D}p}\, Q_R,$$

where $\Gamma_{\mathrm{D}p}$ is an appropriate matrix with ± 1 eigenvalues that depends on the dimension of the Dp-brane. They annihilate the open-string vacuum:

$$Q_{\mathrm{open}} \, |\mathrm{vac}\,\rangle_{\mathrm{open}} = 0$$

The orthogonal linear combination are the so-called broken supercharges that generate a multiplet of

$$2^{16/2} = 2^8 = 256$$

D-brane states. This is the size of the massless multiplet of $d = 10$ supergravity. The fact that a multiplet of massive states has the same size as a massless multiplet is a consequence of the preservation of supersymmetries by the massive state. This phenomenon is called BPS saturation and it is responsible for most of what is known about dualities in string theory.

The phenomenon is very general. Suppose that we have two supercharges Q_+, Q_- that close a supersymmetry algebra with the Hamiltonian H and some $U(1)$ charge Z, given by

$$\{Q_\pm^\dagger, Q_\pm\} = H \pm Z$$

and all other (anti-) commutators trivial. Then, positivity of the anticommutator of supercharges implies the so-called BPS bound

$$E \geq |Z|,$$

with saturation iff some supercharge annihilates a state of charge $\pm|Z|$. Since one linear combination of the supercharges acts trivially on the BPS-saturated state, the supermultiplet has only two states, as opposed to the generic supermultiplet for $E > |Z|$ that has four states.

D-branes are characteristic BPS-saturated states in string theory. In that case the $U(1)$ charge Z is provided by the so-called Ramond–Ramond charge of the D-brane.

The discrete difference between saturated and generic states should be preserved by continuous variations of parameters, such as the coupling. Therefore, properties of BPS states such as their mass formulas and degeneracies can be followed into strong coupling and they serve as tests of the various duality conjectures.

Duality

As stated above, systems with sufficient supersymmetry are amenable at extrapolations from weak coupling to strong coupling. In string models, the massless spectrum is normally protected by gauge symmetries and supersymmetry and therefore remains invariant under a variation of the coupling. Massive modes with mass of $O(1)$ in string units typically increase their mass and nonperturbative states such as D-branes decrease their mass. In the limit of $g_s \to \infty$ we can have D-branes becoming light and fundamental string states becoming non-perturbative states. Normally, this is an indication of a *duality*, i.e. a symmetry that relates strong and weak-coupling.

The template for all these dualities is the electric-magnetic duality of Maxwell's electromagnetism. Under this duality, the electromagnetic tensor transforms into its dual:

$$F_{\mu\nu} \to \widetilde{F}_{\mu\nu} = \frac{1}{2}\, \epsilon_{\mu\nu\rho\sigma}\, F^{\rho\sigma}$$

Therefore, the source-free Maxwell equation and the Bianchi identity are exchanged:

$$\partial^\mu F_{\mu\nu} = 0, \quad \leftrightarrow \quad \partial_{[\mu} F_{\nu\rho]} = 0.$$

Electric and magnetic fields are also exchanged. Hence, if we wish to generalize it to the theory with charged sources, we just consider both an electric current j_e^μ and a magnetic current j_m^μ. As shown by Dirac, the quantum mechanical consistency of such spectrum requires that the fundamental quanta of electric and magnetic charge satisfy:

$$\frac{q_e\, q_m}{2\pi} = \text{integer}$$

To see this, consider the Dirac–Aharonov–Bohm effect for an electron of charge q_e in the background of a monopole field of charge q_m. Under a closed loop γ of the electron, the wave function picks up the Dirac–Aharonov–Bohm phase

$$\exp\left(i\, q_e \oint_\gamma \vec{A}\cdot d\vec{x}\right)$$

We want to write the circulation of the vector potential in terms of the magnetic flux using Stokes's theorem:

$$\oint_\gamma \vec{A}\cdot d\vec{x} = \int_{D_\gamma} \vec{B}\cdot d\vec{S}$$

We have two choices for the surface that is bounded by γ, D_γ and D'_γ, depending on whether we enclose the monopole or not. The difference between the two choices gives the total flux of the magnetic field on $D_\gamma \cup D'_\gamma$, which is simply a two-sphere around the monople. Hence it measures the total magnetic flux of the monopole:

$$\int_{\mathbf{S}^2} \vec{B}\cdot d\vec{S} = q_m$$

On the other hand, this ambiguity should not have physical effects, so that the net phase is trivial:

$$e^{2\pi i n} = 1 = \exp\left(i\, q_e \int_{\mathbf{S}^2} \vec{B}\cdot d\vec{S}\right) = e^{i\, q_e\, q_m},$$

which proves the Dirac quantization condition. This means that the natural couplings of electrons and monopoles are reciprocal. Therefore the duality is nonperturbative, inverting the coupling constant $g \to 1/g$. This is called an S-duality.

It is one thing to prove that a duality must map states in a certain fashion, but it is quite a different matter to actually prove that such states exist in the model under consideration. It has been possible to argue in favour of various versions of S-duality in theories with enough supersymmetry, including string theory. In this case, D-branes play an instrumental role in the discussion.

Another example of duality that can be discussed with more detail, being perturbative in the coupling constant, is T-duality. We have argued that this is a duality between momentum modes and winding modes of strings compactified on a circle. The perturbative spectrum must be invariant. To find the spectrum, consider a closed-string winding w times a circle of radius R. Its rest mass is given by the tension times the length

$$M_{\text{rest}} = \frac{2\pi R w}{2\pi\alpha'} = \frac{R w}{\alpha'}.$$

Boosting to a finite momentum in the compact dimension $p = n/R$ we obtain an energy

$$E = \sqrt{\frac{n^2}{R^2} + \frac{R^2 w^2}{\alpha'^2}}$$

Finally, adding the oscillator mass (18) and boosting in an arbitrary non-compact direction we obtain the energy spectrum

$$E = \sqrt{\vec{p}^{\,2} + \frac{n^2}{R^2} + \frac{R^2 w^2}{\alpha'^2} + M_{\rm osc}^2}$$

Hence, upon compactification on a circle, the spectrum of closed-string states on the non-compact 25-dimensional space-time gets an effective mass squared:

$$M_{\rm eff}^2 = \left(\frac{n}{R}\right)^2 + \left(\frac{w\,R}{\alpha'}\right)^2 + \frac{2}{\alpha'}\,(N_L + N_R - 2) \tag{24}$$

The level-matching condition gets modified in an interesting way. Recall that level-matching was a result of projecting on states with vanishing momentum in the longitudinal σ-direction. If the string winds once around the longitudinal direction, and we add a contribution to P_σ from the momentum of the centre of mass, we have

$$P_\sigma = \frac{2\pi n}{L} + \frac{2\pi}{L}\,(N_L - N_R)\,,$$

so that the constraint is modified to

$$N_R - N_L = n$$

with n the units of centre-of-mass momentum in the compact direction. If the string has w units of winding number in the longitudinal direction, the frequency of oscillations of the string is fractionalized by a factor of w, since the string is actually w times longer. In this case the level matching condition becomes

$$N_R - N_L = n\,w.$$

Considering permutations of the axis and Lorentz-covariance we have the general constraint

$$N_R - N_L = \vec{n}\cdot\vec{w} \tag{25}$$

for general quantized momenta and windings in different directions.

We see that the spectrum is invariant, so that we can simply consider radii $R > \sqrt{\alpha'}$, the rest being obtained by the action of T-duality. This is an indication of string theory possessing a minimum length!

Physically, what happens is that at large $R \gg \ell_s$ the light spectrum is given by momentum modes. At $R \sim \ell_s$ momentum modes and winding modes are of the same mass and they can convert into each other. As $R\ell_s < 1$ the winding modes become the lightest states and any momentum mode will tend to decay into them. Thus at $R \to 0$ we obtain a theory dominated by winding modes. This is actually the same theory as before but written in different variables. In fact, there is a practical difference. Because of the transformation of the string coupling

$$g_s \to g_s\,\ell_s/R,$$

the $R \to 0$ limit at fixed ℓ_s and fixed g_s takes us to a dual theory with $\tilde{R} \to \infty$ but also strongly coupled $\tilde{g}_s \to \infty$. Therefore, a further S-duality will be needed in order to understand such theory.

Although bosonic closed strings as presented here are self-dual under T-duality, this is not a general rule. For example, for supersymmetric strings we have that IIA goes to IIB under

T-duality. Hence T-duality reverses the relative chirality of the supercharges. This is easily understood because T-duality acts with a relative sign between right-movers and left-movers.

Using T-dualities and S-dualities that are suggested by considerations of BPS-saturation, one can draw an impressive web of relationships between string theories in various vacua with different couplings. The whole picture indicates that string theory is unique (so-called M-theory) and that the different string perturbation theories that we construct *ab initio* are useful expansions around particular vacua of the unique M-theory.

The simplest, and in many ways the most fundamental *subweb* of dualities is that of the models with 32 supercharges and $SO(1,9)$ Lorentz symmetry. This includes the two closed superstrings IIA and IIB in $d = 10$, that are related by T-duality, as well as a compactification on a circle of an eleven-dimensional vacuum of M-theory. This vacuum is rather mysterious. It is characterized by a dimensionful Planck length ℓ_p and no adjustable moduli (no analogue of the dilaton). It's low-energy dynamics is described by eleven-dimensional supergravity (a rather unique theory) and has solitonic two-branes and five-branes that are BPS states. A first view of this duality and its power is offered through Problem 6.

In situations where there is a duality one can answer an interesting question of principle. We motivated strings in the first lecture by analogy with the electroweak solution to the Fermi's theory problems. Namely, even if gravity becomes strong at the Planck mass m_p, the threshold of *weakly-coupled* strings lies at parametrically lower energies (in ten dimensions):

$$m_s = \left(g_s^2\, m_p^4\right)^{1/4} < m_p$$

What happens for strong coupling, $g_s > 1$? In all examples of nonperturbative dualities between two different weakly-coupled string theories, it turns out that the Planck mass is invariant, whereas $g_s \to 1/g_s$ and m_s transforms accordingly. Hence, after the duality the string mass of the dual string theory

$$\tilde{m}_s = \left(\tilde{g}_s^2\, m_p^4\right)^{1/4}$$

still remains below the Planck scale. This means that in such systems $m_s < m_p$ always, with $m_s \sim m_p$ at the self-dual point $g_s \sim 1$.

Considerations of non-perturbative dualities and D-branes have enlarged considerably the types of models that one can use for phenomenological applications. This gives a richer set of possibilities for the connection between string theories and the low-energy world, at the price of losing some of the previously considered qualitative "model-independent" predictions of the theory. A prime example of this is explained in the next sections.

The Origin of Gauge Symmetry in String Theory

String theory has natural mechanisms to produce gauge interactions. Given that this is the main dynamical feature of the Standard Model, this is an interesting subject to discuss.

We have seen that a simple and very fundamental origin for gauge fields arises from the quantization of open strings or, what is the same, D-branes. In particular, non-abelian gauge theories arise as low-energy limits of systems with coincident D-branes. In these systems, the gauge coupling in $p + 1$ dimensions is given by

$$g_{\rm YM}^2 \sim g_s\, \ell_s^{p-3}$$

The other generic gauge symmetries that arise are ten-dimensional gravity, with coupling

$$\kappa_{10}^2 = g_s^2\, \ell_s^8,$$

and a $U(1)$ symmetry of the antisymmetric tensor field with coupling of gravitational strength. There are also higher-rank antisymmetric tensors supporting $U(1)$ symmetries with couplings that given just by the string scale (independent of g_s) that we will not discuss further.

In closed-string theories (the gravitational sector) one can generate gauge symmetries using the old mechanism of Kaluza–Klein (see Problem 4). One assumes that spacetime is compactified as $\mathbf{R}^4 \times K_6$ where K_6 is some compact six-manifold. If K_6 has some rigid isometries (a global symmetry group G), then at energy scales $E \ll 1/R_K$ with R_K the typical size of K_6, we cannot detect the dependence of the fields on the details of K_6. Hence we have a symmetry under G-transformations of K_6 at *any* point on $\mathbf{R}^4$, i.e. we have a gauge symmetry in $\mathbf{R}^4$ with gauge group G. The components of the metric tensor $g_{\mu\nu}$ with one index in K_6 and one index in $\mathbf{R}^4$ give the required gauge fields.

Unfortunately, together with the gauge fields, it is often the case that other massless scalar fields appear, associated to the fact that spaces of the form $\mathbf{R}^4 \times K_6$ that satisfy the Einstein's equations come in continuous families, depending on the size and shape of K_6. These continuous parameters translate into massless scalar fields (analogous to Goldstone bosons) in the low-energy effective theory. These extra scalar fields (so-called "radions" or "moduli" in general) are always problematic for phenomenology and we will return to them later.

There is a general rule to derive the matching of coupling constants under the Kaluza–Klein reduction. Suppose we have a general Lagrangian defined on a space of the form $\mathbf{R}^4 \times K$:

$$S[\,\Phi\,] = \frac{1}{g^2} \int_{\mathbf{R}^4\times K} \mathcal{L}\,[\Phi(x,y)],$$

where $x \in \mathbf{R}^4$ and $y \in K$ are coordinates in the macroscopic Minkowski space and the compact manifold respectively. The fields Φ stand for matter, gauge or gravitational fields on the higher-dimensional theory and g is a gauge or gravitational coupling. Now, on energy-momentum scales $E \ll 1/R_K$ we can approximate $\Phi(x,y)$ by the averaged fields

$$\overline{\Phi}(x) = \frac{1}{\mathrm{Vol}\,(K)} \int_K dy\; \Phi(x,y),$$

since we cannot distinguish the details of the variations of the fields on the compact manifold. In a first approximation at large distances we can just substitue the fields by their averages over K and write an effective Lagrangian of the form

$$S[\,\overline{\Phi}\,]_{\rm eff} = \frac{1}{g_{\rm eff}^2} \int_{\mathbf{R}^4} \mathcal{L}\,[\,\overline{\Phi}(x)\,] + \dots,$$

where the dots stand for derivative terms that come from integrating-out the higher modes and

$$\frac{1}{g_{\rm eff}^2} = \frac{\mathrm{Vol}\,(K)}{g^2} \tag{26}$$

The volume factor arises because the Lagrangian density is independent of the y coordinate once it is evaluated on the averaged fields. This formula is the basic rule of the Kaluza–Klein reduction of couplings in the tree level approximation. If the coupling g^2 has length dimension a, so that the effective expansion parameter is

$$\alpha_g(E) = g^2\, E^a\,,$$

then the effective coupling after compactification has length dimension $a - d_K$, with $d_K = \dim\,(K)$. Then the effective expansion parameter at $E \ll 1/R_K$ is

$$\alpha_{g_{\rm eff}}(E) = g_{\rm eff}^2\, E^{a-d_K} = \frac{\alpha_g(E)}{(E\,R_K)^{d_K}}\,.$$

Therefore, both dimensionless couplings match at $E \sim 1/R_K$. Notice that the higher-dimensional theory has always a stronger ultraviolet behaviour. This means that Kaluza–Klein thresholds are always close to nonperturbative physics, unless the low-energy QFT is substituted by some ultraviolet completion, such as string theory.

In string theory, once the compactification manifold is of stringy size $R_K \sim \ell_s$, the Kaluza–Klein (KK) reduction gets modified by subtle stringy effects. First, the compact geometry is not describable in classical terms (it becomes "fuzzy"). Second, there are new long-distance effects. For example, winding modes can become massless and enhance the gauge symmetry from $U(1)$ to $SU(2)$ or larger groups. We can illustrate the basic phenomenon from the formulas (24) and (25) for the effective mass after compactification on a circle in the X^1 direction. At a generic value of the radius R, the low energy spectrum has two gauge fields. One coming from the KK reduction of the metric $A_\mu \sim g_{\mu 1}$, and the other coming from the KK reduction of the antisymmetric tensor $A'_\mu \sim b_{\mu 1}$. Thus the generic gauge group is $U(1) \times U(1)'$.

In the basic construction of the graviton multiplet in terms of oscillators, $g_{\mu\nu}$ and $b_{\mu\nu}$ differ by the relative sign of left- versus right-moving oscillations. This means that they are exchanged by T-duality. Since momentum modes are charged with respect to the first $U(1)$ we conclude that winding modes are charged with respect the second $U(1)'$. Now, the mass formula at the self-dual radius under T-duality, $R = \sqrt{\alpha'}$, becomes:

$$M_{\rm eff}^2 = \frac{(n-w)^2}{\alpha'} + \frac{4}{\alpha'}\,(N_L - 1) = \frac{(n+w)^2}{\alpha'} + \frac{4}{\alpha'}\,(N_R - 1)\,.$$

Hence, with one oscillator mode excited in either left- or right-movers, by adjusting at the same time $n, w = \pm 1$ we have four extra massless vectors with non-vanishing values of winding and/or momentum. This means that each $U(1)$ gets extended to an $SU(2)$ and the total gauge group at the self-dual radius is $SU(2) \times SU(2)'$.

The extreme example of this is the heterotic string. In this model the left-movers are like in a standard superstring in ten dimensions, but the right-movers are purely bosonic and are compactified from $d = 26$ to $d = 10$ in a 16-dimensional torus of stringy size (in fact self-dual under T-duality just like in the previous example). The result is a huge enhancement of gauge symmetry to a group $SO(32)$ or $E_8 \times E_8$ that is present in the ten-dimensional theory. This models are particularly nice for phenomenology.

Not surprisingly, the two general methods to generate gauge symmetry (ordinary or stringy Kaluza–Klein and D-branes) are related by nonperturbative dualities.

In general, gauge groups that exist in $d = 10$ and come from the closed-string sector (the heterotic construction) will have a coupling of same order as the gravitational coupling, in string units:

$$g_{\rm YM}^2 \sim g_s^2\, \ell_s^6$$

Couplings and Scales in String Theory

The previous relations can be used to obtain a generic feature of string compactifications. If the gauge symmetry originates from the closed-string sector in ten dimensions then the ratio between the gauge coupling and the gravitational coupling stays fixed under compactification. This means that the ratio of the four-dimensional gauge couplings to the gravitational coupling is of order one in string units, so that the unification of gauge and gravitational couplings should occur at around the string scale. We have

$$G_{\rm N} \sim \kappa_4^2 \sim \frac{g_s^2 \ell_s^8}{R_K^6}, \qquad \alpha = \frac{g^2}{4\pi} \sim \frac{g_s^2 \ell_s^6}{R_K^6}$$

Then, we can eliminate the internal KK volume and obtain

$$G_{\rm N} = \ell_P^2 \sim \alpha\, \ell_s^2$$

This means that, since $\alpha = \alpha_{\rm GUT} \sim 1/25$ in supersymmetric unification, the Planck length and the string length are not too far away. On the other hand, if the string is weakly coupled, so that we can believe the calculation of the spectrum, then $g_s < 1$ implies

$$R_K^6 < \ell_s^6 / \alpha$$

and the compactification radius is also in the order of magnitude of the string scale. Putting $R_K^{-1} \sim M_{\rm GUT} \sim 10^{16}$ GeV as suggested by LEP fits to supersymmetric models, one obtains (keeping the factors of 2 and π):

$$G_{\rm N} > \frac{\alpha_{\rm GUT}^{4/3}}{M_{\rm GUT}^2},$$

which is too large by a factor of $O(400)$.

Hence, the hierarchy between the GUT unification scale, the string scale and the Planck scale comes out roughly right in order of magnitude but not quite exactly so. Since the GUT extrapolation is so bold, this could be interpreted as a big success, and one could imagine many complications, such as threshold effects or extra low-energy matter, that would bring a better agreement.

A more radical alternative to evade this problem is to detach the evolution of the gravitational coupling under compactification from that of the gauge coupling. This is easily achieved through the use of D-branes. In that case, the gauge and gravitational couplings scale with different powers of the string coupling and we can adjust independently all parameters to fit $G_{\rm N}$ to the experimental value, while keeping $M_{\rm GUT} \sim 10^{16}$ GeV and $\alpha_{\rm GUT} \sim 1/25$.

Since D-branes can have various dimensionalities, the power of R_K in the effective four-dimensional couplings is also a matter of model-choice. For example, on a D3-brane, gauge couplings are given by $\alpha \sim g_s$ with no leading dependence on R_K or ℓ_s. This idea of using D-branes to localize gauge fields can be pushed to the extreme, to actually lower the Planck scale to the TeV range. This sounds impossible, but in fact it isn't. What is lowered is the fundamental Planck scale in the string theory, i.e. the ten-dimensional Planck scale. The apparent Planck scale is given by $G_{\rm N}$ which is

$$\frac{1}{G_{\rm N}} = \frac{R_K^6}{g_s^2 \ell_s^8}$$

as before. We can set $m_s = O(\text{TeV})$ with $g_s < 1$ by taking R_K sufficiently large, say R_K^{-1} of the order of a fraction of a Fermi! This looks crazy because the extra dimensions are so large that they should have been detected at LEP. However, if we confine the whole of the non-gravitational Standard Model to a D3-brane, its threshold for new physics on the brane is set by $m_s \sim$ TeV and not by R_K^{-1}. The situation is even more strange in models (the Randall–Sundrum type) where the extra dimensions are *non-compact* and remain unseen because of strong curvature effects, i.e. the role of R_K is played by the curvature radius in the direction of the large extra dimensions.

Direct detection of the large extra dimensions must proceed by purely gravitational effects, such as tests of Newton's gravitational law. The experimental bounds on this are rather weak, since they tolerate R_K in the submilimiter range! Effects from the light KK states will however appear in virtual processes and this can be used to put (not very stringent) bounds on R_K and m_s.

Hence, we see that the scale at which strings appear could be anywhere between a few TeV and 10^{19} GeV. If strings or whatever physics is associated to quantum gravity shows up at low energies, then we will be able to explore quantum gravity in the laboratory, opening a revolutionary period of the same depth as the elaboration of quantum mechanics out of atomic physics. If, on the other hand, the string scale is close to the effective four-dimensional Planck scale, then we will need to rely on indirect evidence and hard calculation to test our ideas about quantum gravity.

Strings and Predictability

The idea that the string scale could be within experimental reach around the TeV range is psychologically irresistible, even if it comes at the high price of rendering irrelevant all the hints at a gauge unification around $M_{\rm GUT} \sim 10^{16}$ GeV. A different matter is whether a given model with low string scale is viable when it comes to the details (recall that the "details" are the killer for technicolor-type models). We will not enter this discussion here, partly because low-scale models are comparatively less surveyed.

However, it is striking that a theory without adjustable free parameters gives no clue as to where the scale of fundamental physics lies. In string theory, this is equivalent to the problem of fixing the values of the string coupling g_s and the size and shape of the compactification space K_6. These parameters are associated to scalar fields in four dimensions whose effective potential is generically flat before supersymmetry breaking. The general problem of fixing the expectation values of dilaton and "radions" is called the vacuum stabilization problem.

After supersymmetry breaking, a potential $V(\phi)$ will be generated for the dilaton. This occurs via some nonperturbative dynamics and hence it depends on the string coupling as $V \sim \exp(-C/g_s^a)$ with $a \geq 1$. Typically $a = 2$ or perhaps $a = 1$ (see Problem 5). This means that the leading scaling with the dilaton is

$$V(\phi) \sim \exp\left(-C\, e^{-a\,\phi}\right), \quad \text{as} \quad \phi \to -\infty,$$

and we learn that in the weak-coupling regime $\phi \to -\infty$ the potential slopes to zero with restoration of supersymmetry at $\phi = -\infty$, which corresponds to the free, $g_s = 0$, theory. Hence, there are no natural vacuum states that can be studied with weak-coupling methods. This is a generic problem that occurs for most of the scalar fields that are massless in the supersymmetric approximation. This difficulty of fixing their expectation values is called the *moduli problem* (or Dine–Seiberg problem for the specific case of the dilaton).

This problem is the main reason why string models of low-energy physics remain *non-predictive*, i.e. they do not actually determine the low-energy spectrum, much less the numerical values of the parameters of the Standard Model Lagrangian. One must rely on special mechanisms with some degree of fine-tunning to achieve a "technical" vacuum at weak coupling, or assume that a vacuum at strong coupling $g_s = O(1)$ exists with the right properties. In this case one must explain why the gauge couplings are relatively small with a strong fundamental string coupling. However, the main problem of this alternative is that one loses computability in practice (notice that dualities do not help here, since de region $g_s = O(1)$ is self-dual).

To the extent that the problem of moduli stabilization is related to supersymmetry breaking, it is afflicted by all the low-energy constraints discussed before, most notably by the flavour, CP, and cosmological constant problems. There are other problems that are specific and have to do with the presence of scalars with couplings of gravitational strength. While these fields might be useful for various cosmological duties, such as inflation, they also tend to destroy the succesful predictions of nucleosyntesis.

The pattern of constraints to be satisfied is so intrincate that being *numerically predictive* in string theory might require a rather complete control of the nonperturbative dynamics at the *non-supersymmetric* level. Despite the revolutionary progress of the last years by mapping-out the duality web, much of what has been learned concerns vacua with lots of supersymmetry.

The fact that *consistent* vacua exist in which supersymmetry is *exact* brings the question of the vacuum selection to a rather unconfortable neighbourhood from the epistemological point of view. Since the theory contains solutions that are consistent and are manifestly different from our universe, it seems that there is a certain degree of "historical contingency", in the cosmological sense. This is usually referred to as the "anthropic principle". It simply means that perhaps some parameters of the low energy Lagrangian are cosmological accidents and therefore cannot be predicted from fundamental physics. In this respect, the fact that the cosmological constant can be nailed to within one order of magnitude on the basis of very weak anthropic arguments is rather remarkable!

The problem with this is that, once the cosmological constant is fixed anthropically, it is possible that other constants are anthropic too, especially within a typical scenario of low-energy supersymmetry breaking. Of course, this would be a rather unhappy state of affairs because it would put some limits to our ability of inferring short-distance physics from low-energy information (the parameters of the Standard Model). This limitation would be there independently of the amount of computational power. One possibility is that non-supersymmetric vacua of string theory with the right properties turn out to be scarce and easily identifiable. Of course, we could also be very lucky and find strings at low energy, so that experient will guide us in the unraveling of the quantum gravity realm!

There is also the exciting possibility that our poor understanding of vacuum-selection issues is a symptom of a new theoretical crisis, i.e. the need for completely new ideas. In the next sections we will describe how string theory hints at such a change of paradigm, regarding non-perturbative questions in quantum gravity.

High Energy Behaviour

After the long discussion of vacuum properties of string theory we now return to the basic question of why strings manage to quantize gravity in a consistent way.

Strings succesfully smear the gravitational interaction on length scales of the order of the string length ℓ_s. From the point of view of the free spectrum, we found that a string theory looks like an infinite tower of massive fields, plus a rather universal massless spectrum. Much of what we like about strings has to do with the massless spectrum, namely the presence of gravity and gauge symmetry as well as supersymmetry.

If string theory was simply a tower of massive fields it is not obvious how this is enough to regularize consistently the gravitational interaction, that has resisted all previous attemps at a quantum treatment based on QFT ideas.

Although more sophisticated answers exist, there is a simple physical argument that explains the success of strings on this account. The idea is to look at the typical state of string theory at very high energy. The spectrum at very high mass has the distinctive property of having exponential degeneracy of states. From Problem 3 we know that the density of closed-string mass levels in d dimensions grows as

$$\rho(M) \sim \ell_s \, \frac{e^{\beta_s M}}{(\ell_s M)^d},$$

where $\beta_s \sim \ell_s$. From here, using the normal dispersion relation

$$E = \sqrt{\vec{p}^{\,2} + M^2}$$

one can compute the complete density of single-string states:

$$\omega(E) \sim m_s^{d-2} \, V \, \frac{e^{\beta_s E}}{(\ell_s \, E)^{\frac{d+1}{2}}}$$

The physical interpretation of these formulas is that for typical energies $E\ell_s \gg 1$ it is entropically favourable for the string to distribute the total energy E on oscillation degrees of freedom rather than on the boost of the centre of mass.

A model that represents the typical oscillation state at very high energy is a random-walk representation of the string. Let us build a string configuration from n steps of a random walk in d dimensions, each of string-length stretch. The total energy is then

$$\ell_s \, E \sim n$$

in string units. The size of the random walk can be estimated by the mean square distance to the centre of mass that we suppose at the origin:

$$(\text{Size})^2 \sim \left\langle \sum_{i=1}^{d-1} \sum_{s=1}^{n} (X_s^i)^2 \right\rangle,$$

where s is an index for the steps. We assume that at each step the path makes a random choice among nearest neighbour lattice sites. Then the random variables X_s^i are independent and

$$(\text{Size})^2 \sim d \, n \, \langle \, (X_s^i)^2 \, \rangle \sim d \, n \, \ell_s^2$$

Therefore, the volume W of the random walk scales like

$$W \sim n^{(d-1)/2}.$$

The number of configurations with a fixed initial point and an arbitrary final point are of order

$$\exp(C\, n)$$

for some positive constant C. This is because the probability of any given configuration is

$$P \sim p^n$$

with $p = e^{-C} < 1$ the probability of one single step in a certain direction.

Now, the random walk must close on itself because the string was closed. That means that our first estimate with free endpoint overcounts by a factor of the order of the volume of the random walk. We have an extra overcounting by a factor of n from the fact that the random walk can start at any point in the string. Finally we have the possibility of translating the centre of mass of the random walk, giving a factor of the volume V. All in all we have

$$\omega(n) \sim V\, \frac{1}{n}\, \frac{1}{W(n)}\, e^{C\, n} \sim V\, \frac{e^{C\, n}}{n^{(d+1)/2}}.$$

Putting now $E \sim n\, m_s$ we get the previous result that was found explicitly in Problem 3. This means that the random walk model is a good model of the typical configuration of strings at very high energy.

The important lesson from this model is that the typical high-energy string has a size that grows with the energy. Thus, high-energy strings are soft rather than hard probes of space-time. There is a minimal length that can be probed by perturbative string scattering and this is the string length scale. This modified "uncertainty principle" is one of the landmarks of the physics of strings and is ultimately responsible for the succesful smearing of the gravitational interaction.

Into the Black Hole

The exponential degeneracy of states of strings is so large that one wonders whether it might be related to the degeneracy of black hole states that follow from the Bekenstein–Hawking formula for the entropy of a black hole (in four dimensions)

$$S_{\rm BH} = \frac{A_H}{4G_{\rm N}}, \tag{27}$$

where A_H is the area of the event horizon

$$A_H = 4\pi\, R_H^2 = 16\pi\, (G_{\rm N}\, M)^2$$

We see that the asymptotic growth of the black hole entropy $S_{\rm BH} \sim (\ell_P M)^2$ is stronger than that of strings $S_s \sim \ell_s\, M$. However, notice that the black hole entropy scales like the inverse Newton constant. Therefore, there is a value of the coupling at which highly excited strings and black holes of the same mass have the same entropy:

$$\ell_s\, M \sim \ell_s^2\, g_s^2\, M^2$$

This gives the scale of "correspondence" between strings and black holes

$$M_{\rm corr} \sim \frac{m_s}{g_s^2}$$

The Schwarzschild radius of such a black hole is

$$R_{\rm corr} = 2\, G_{\rm N}\, M_{\rm corr} \sim \ell_s^2\, g_s^2\, M_{\rm corr} \sim \ell_s$$

This is the scale at which the curvature at the horizon becomes of order one in string units and the low-energy gravity used to derive the Bekenstein–Hawking formula breaks down. This coincidence gives support to the idea that excited string states really become black holes for $M > M_{\rm corr}$. This has been checked to $O(1)$ accuracy for all known black holes with arbitrary charges in arbitrary dimensions.

For black holes with sufficient supersymmetry, the entropy is protected by algebraic constraints (the BPS-saturation to be discussed below) and the matching between the Bekenstein–Hawking and the string determination should be exact. Indeed, it has been possible to match exactly the factor of $1/4$ in (27) for such black holes. This is arguably the most significative quantitative success of string theory.

The fundamental importance of these results cannot be over-emphasized. If we were to interpret (27) *a la* Boltzmann, i.e. microscopically, we would write something like:

$$S(M)_{\rm BH} = \log\, \dim\, \mathcal{H}_{E=M}$$

This is the standard definition of the microcanonical entropy in statistical mechanics, as a measure of the dimension of the Hilbert space of states with a given total energy. In this case, the Hilbert space of microstates with a given black hole mass. However, a look at (27) reveals that the value of the entropy is about one bit of information per unit Planck area of the horizon. Hence, the Hilbert space needed is neccessarily a Hilbert space of quantum gravity!

This simple argument shows that the Bekenstein–Hawking entropy formula is one of the most important pieces of information that we have in developing a theory of quantum gravity. The fact that string theory deals succesfully with this fundamentally non-perturbative quantity lends strong support to the view that string theory is the correct way of quantizing gravity.

Holography

The so-called "holographic principle" of 't Hooft and Susskind builds upon these considerations the key to the dynamics of quantum gravity.

The basic idea is very simple. In a local description based on QFT the dimension of the Hilbert space (with Planck-scale regularization) grows exponentially with the volume. If we look at a QFT on a box of size L and consider states of total energy E much larger than any mass scale in the theory, including the box mass-gap $1/L$, the density of states can be estimated by computing the entropy of thermal radiation at temperature T:

$$S \sim V\, T^3.$$

On dimensional grounds it also follows that the energy is

$$E \sim V\, T^4$$

¿From here we can eliminate the temperature and write:

$$S(E) \sim (L\, E)^{3/4}$$

This density of states is smaller than the Hagedorn law $S \sim E$ or the black-hole law $S \sim E^2$. At any rate, the typical states will be extensive on the box and the entropy grows linearly with the volume. However, we know that for a fixed volume, there is a maximal energy density such that the radiation is stable towards collapse into a black hole.

To see this, consider a system of a black hole of mass M in the box together with thermal radiation of energy $E - M$. The total entropy is approximated by the sum of radiation and black-hole entropy:

$$S = C_1 \, (LE - LM)^{3/4} + C_2 \, G_{\rm N} \, M^2.$$

One can see that at energies of order

$$E \sim \left(\frac{V}{G_{\rm N}^4} \right)^{1/5}$$

it becomes entropically favourable to nucleate a stable black hole in equilibrium with the radiation. As the total energy continues to raise the size of the black hole $R_H \sim G_{\rm N} \, M$ becomes of the order of the size of the box. At this point we have reached the maximum "capacity" of the box for physical states.

The holographic principle states that, since the maximal capacity of information is reached with a black hole of the size of the system, and this only uses a finite number of degrees of freedom per Planck *area*, the rest of the Hilbert-space of the local QFT must be redundant. So one conjectures that in quantum gravity nonperturbative effects are characterized by degrees of freedom residing on the boundary of space rather than the bulk. According to this idea, locality in QFT is a "holographic" illusion of working with "diluted" states.

This very radical proposal has been put on a firm ground through string-theory models that realize it, notably the Matrix Model but especially the AdS/CFT correspondence. The ideas around the holographic principle are operating a significative change of paradigm in our thinking about quantum gravity at the nonperturbative level. In particular, very fundamental issues such as the precise relation between short-distance physics and long-distance physics in quantum gravity (basic to the cosmological-constant problem) could take an entirely new shape.

Concluding Remarks

We conclude with a positive message. The evidence for the deep connection between string theory and quantum gravity is by now quite significative. It is also clear that deep physical principles of an entirely different nature are being uncovered. The fact that the simplest classical solutions of string theory give good qualitative models of the low-energy world, including many of the ingredients of the Standard Model and beyond, lends support to the idea that strings could realize the unification paradigm. Unfortunately, our present techniques do not cover the most interesting calculations. On the positive side, we also see that some significant pieces of low-energy physics, such as the black hole entropy formula, can guide us in the search for the fundamental dynamical rules of the theory.

Notation and Conventions

In the lectures and problems, I try to use consistently the metric signature

$$\eta_{\mu\nu} = \mathrm{diag}\,(-1, 1, 1, \ldots, 1),$$

where $\mu, \nu = 0, 1, \ldots, d-1$ and the minus sign corresponds to the time coordinate.

The volume measure in integrals is frequently omitted when clear from the context:

$$\int \equiv \int d^d x \, \sqrt{-g}$$

For integrals in flat space I also use

$$\int_x \equiv \int d^d x$$

or, in momentum space

$$\int_p \equiv \int d^d p$$

Fourier transforms are defined as

$$\phi(x) = \int \frac{d^d x}{(2\pi)^{d/2}} \, \phi(p) \, e^{ipx}$$

Other notational devices: we use κ^2 for the gravitational coupling, i.e. the Einstein–Hilbert action is written as

$$\frac{1}{2\kappa_d^2} \int d^d x \, \sqrt{-g} \, \mathcal{R} + \ldots$$

The Newton constant is defined in any dimension by

$$16\,\pi\, G_d = 2\,\kappa_d^2$$

Historically, the Planck mass and length in four dimensions are defined as

$$M_P = 1/\ell_P = (G_{\mathrm{N}})^{-1/2} \sim 10^{19} \;\; \mathrm{GeV}$$

However, it is more convenient for the discussion of string dualities to adopt another definition with slightly different factors of 2π:

$$2\kappa_d^2 = 16\pi G_d = (2\pi)^{d-3} \, \ell_p^{d-2}, \qquad m_p = 1/\ell_p$$

The fundamental string length and mass scales and the string tension are given by

$$\ell_s = \sqrt{\alpha'}, \qquad m_s = 1/\ell_s, \qquad T_{\mathrm{F1}} = \frac{1}{2\pi\alpha'}$$

The relation between the string scale, the string coupling constant and the Planck length in the $d = 10$ uncompactified string theories (type I, IIA, IIB and heterotic) is

$$\ell_p^4 = g_s \, \ell_s^4$$

Problems

Problem 1

In this problem we elaborate on the Polyakov action

$$S_{\mathrm{P}} = -\frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{-h}\; h^{ab}\,\partial_a X^\mu\,\partial_b X^\nu\;\eta_{\mu\nu}. \tag{28}$$

and its relation to the Nambu–Goto action

$$S_{\mathrm{NG}} = -\frac{1}{2\pi\alpha'}\int d^2\sigma\sqrt{-\det(\partial_a X^\mu\;\partial_b X^\nu)\,\eta_{\mu\nu}}. \tag{29}$$

The main purpose of the problem is to introduce the extra symmetry of the Polyakov action and the emergence of a *conformal symmetry* after gauge-fixing.

Prove that both actions are on-shell equivalent by evaluating the Polyakov action at the solution of the h_{ab} equation of motion.

The variation of the Polyakov action with respect to the two-dimensional metric h_{ab} is:

$$\delta S_{\mathrm{P}} = -\frac{1}{4\pi\alpha'}\int \sqrt{-h}\,\delta h^{ab}\left(\partial_a X\cdot\partial_b X - \tfrac{1}{2}\,h_{ab}\,h^{cd}\,\partial_c X\cdot\partial_d X\right),$$

where we have used the general formula for the variation of the determinant:

$$\delta\, h = \delta\det(h_{ab}) = \delta\prod_n \lambda_n = \sum_n \frac{h}{\lambda_n}\,\delta\,\lambda_n = -\sum_n h\,\lambda_n\,\delta\left(\frac{1}{\lambda_n}\right),$$

where we have diagonalized the metric in eigenvalues λ_n. The two last equalities give

$$\delta\, h = h\, h^{ab}\,\delta\, h_{ab} = -h\, h_{ab}\,\delta\, h^{ab}$$

and

$$\delta\sqrt{-h} = -\frac{1}{2\sqrt{-h}}\delta\, h = -\tfrac{1}{2}\,\sqrt{-h}\,h_{ab}\,\delta\, h^{ab}.$$

Now setting $\delta S_{\mathrm{P}} = 0$ yields

$$\partial_a X\cdot\partial_b X = \tfrac{1}{2}\;h_{ab}\,h^{cd}\,\partial_c X\cdot\partial_d X \tag{30}$$

Taking determinants

$$\det\left(\partial_a X\cdot\partial_b X\right) = h\;\left(\tfrac{1}{2}\;h^{cd}\,\partial_c X\cdot\partial_d X\right)^2 \tag{31}$$

Equivalently

$$\tfrac{1}{2}\;\sqrt{-h}\,h^{ab}\,\partial_a X\,\partial_b X = \sqrt{-\det\left(\partial X\cdot\partial X\right)},$$

which proves the statement.

Generalize the Polyakov action S_{P} to a $(p+1)$-dimensional world-volume and show that the symmetry under local Weyl rescalings $h_{ab}\to e^{2\omega}h_{ab}$ is only present for $p=1$.

The appropriate generalization of the Polyakov action from the $(1+1)$-dimensional world-sheet of a string to a $(p+1)$-dimensional world-volume of a p-brane is

$$S_p = -\frac{T_p}{2} \int_{\Sigma_{p+1}} \sqrt{-h}\, h^{ab}\, \partial_a X^\mu\, \partial_b X^\nu\; \eta_{\mu\nu},$$

where $\mu, \nu = 0, 1, \ldots d-1$ as before, but now $a, b = 0, 1, \ldots p$. Under a Weyl rescaling

$$h = \det(h_{ab}) \to (e^{2\omega})^{p+1}\; h$$

Since the inverse matrix

$$h^{ab} \to e^{-2\omega}\, h^{ab},$$

the Weyl rescaling amounts to

$$\sqrt{-h}\, h^{ab} \longrightarrow (e^{\omega})^{p-1}\; \sqrt{-h}\, h^{ab}.$$

Hence, it is precisely Weyl-invariant for $p = 1$, strings. This extra symmetry of the string theory as opposed to the more general p-branes is of fundamental importance in the development of the theory. It means that, ultimately, the worldsheet metric was a redundant device. In fact, there is no known consistent quantization of relativistic p-branes for $p > 1$. Perhaps this is not a coincidence, and Weyl symmetry is actually a absolute condition for consistence.

We can use the Diff$\times$Weyl symmetry to set $h_{ab} = \eta_{ab}$. The following identity gives the Weyl transformation of the curvature scalar $\mathcal{R}^{(2)}$ in two dimensions. If

$$h'_{ab} = e^{2\omega}\, h_{ab},$$

the following is true

$$\sqrt{-h'}\, \mathcal{R}^{(2)}(h') = \sqrt{-h}\left(\mathcal{R}^{(2)}(h) - 2D^2\, \omega\right), \tag{32}$$

where D_a is the covariant derivative.

Use this identity to *argue* (which is different from *prove*) that locally the metric can be brought to standard form $h_{ab} \to \eta_{ab}$ by a combined Diff$\times$Weyl transformation.

Equation (32) can be solved for ω, at least locally. It has the form

$$D^2\, \omega = \rho$$

for some function ρ. This is the Poisson equation in two dimensions, so it has the solution

$$\omega(\sigma) = \int d\sigma'\; \rho(\sigma')\; \langle\, \sigma' \,|\, \frac{1}{D^2}\, |\sigma\,\rangle$$

in terms of the Green's function of the two-dimensional Laplacian. This is true for any pair of Weyl-related metrics, so it is true for h'_{ab} such that $\mathcal{R}^{(2)}(h') = 0$. Such a metric is locally flat, and thus in cartesian coordinates it is just η_{ab}.

After Diff $\times$ Weyl gauge-fixing, there is still a lot of gauge freedom, at least locally. These are transformations in Diff$\times$ Weyl that still leave $h_{ab} = \eta_{ab}$ unchanged, and therefore are not fixed by the choice of flat metric.

Introduce locally a pair of light-cone coordinates $\sigma^\pm = \tau \pm \sigma$, so that the flat metric reads:

$$ds^2 = -d\tau^2 + d\sigma^2 = -d\sigma^+ \, d\sigma^- .$$

Find the residual group.

Consider reparametrizations of the form

$$\sigma^\pm \to f^\pm(\sigma^\pm),$$

with $f^\pm$ arbitrary. They serve as new coordinates in which the metric reads

$$ds^2 = -d\sigma^+ \, d\sigma^- = -\left(\partial_+ f^+ \, \partial_- f^-\right)^{-1} \; df^+ \, df^- = g_{f^+ f^-} \; df^+ \, df^-,$$

with $g_{f^+ f^+} = g_{f^- f^-} = 0$. This reparametrization conserves the conformal gauge but it is equivalent to a Weyl transformation with

$$e^{-2\omega} = \partial_+ f^+ \, \partial_- f^-$$

An additional Weyl transformation by a factor of $e^{-2\omega}$ leaves the metric invariant *in the new coordinates*:

$$ds^2 \to e^{-2\omega} \, ds^2 = -df^+ \, df^-$$

This is a *conformal* transformation. Notice that it changes the physical distances (rescales ds^2).

Find the infinitesimal generators of the conformal transformations and the algebra that they satisfy.

To do this, write

$$f^\pm(\sigma^\pm) = \sigma^\pm + v^\pm(\sigma^\pm)$$

and drop all terms of higher than linear order in $v^\pm$. In addition, assume that the world-sheet is locally a cylinder with periodic σ with period 2π. Then one can Fourier-analyze $v^\pm$ in:

$$v^\pm(\sigma^\pm) = \sum_{n \in \mathbf{Z}} v_n^\pm \, e^{in\sigma^\pm}$$

The generators can be found by their action on a scalar field.

A scalar field satisfies $\phi(\sigma^a) = \phi'(f^a)$. The variation of the field at a fixed coordinate point is

$$\delta\,\phi(\sigma^a) = \phi'(\sigma^a) - \phi(\sigma^a) = \phi'(f^a - v^a) - \phi(\sigma^a) = \phi'(f^a) - v^b \partial_b \, \phi'(f^a) - \phi(\sigma^a) + \ldots = -v^+ \partial_+ \phi - v^- \partial_- \phi + \ldots$$

The differential operators generating the conformal transformations are

$$-v^\pm \, \partial_\pm = i \sum_n v_n^\pm \, i \, e^{in\sigma^\pm} \, \partial_\pm = i \sum_n v_n^\pm \, \ell_n^\pm$$

The operators

$$\ell_n^\pm = i \, e^{in\sigma^\pm} \, \partial_\pm$$

satisfy the so-called *Virasoro algebra*:

$$\left[\ell_n^\pm, \ell_m^\pm\right] = (n - m) \, \ell_{n+m}^\pm$$

Thus, the gauge-fixed theory on the world-sheet of a string is a *conformal field theory*, locally invariant under two copies of the Virasoro algebra.

Problem 2

In this problem we discuss the relation between the graviton and the dilaton at the level of the linearized theory. The purpose is to derive the low-energy effective action of massless modes of closed string theory in the linearized approximation. In the process, it will become clear that the expectation value of the dilaton is related to the string coupling constant by

$$g_s = e^{\langle \phi \rangle}$$

The physical transverse polarizations of massless closed strings were found to be given by a general tensor t_{ij} with indices in the transverse spatial space, transforming with respect to the group of transverse rotations $SO(d-2)$. This tensor splits in irreducible representations of $SO(d-2)$ as a symmetric traceless tensor (the graviton), an antisymmetric tensor and the pure trace (the dilaton).

$$t_{ij} \to h_{ij} \oplus b_{ij} \oplus \phi$$

Show that these degrees of freedom result from a d-dimensional general tensor $t_{\mu\nu}$ with gauge invariance

$$t_{\mu\nu} \longrightarrow t_{\mu\nu} + \partial_\mu \lambda^L_\nu + \partial_\nu \lambda^R_\mu,$$

and subject to transversality conditions:

$$\partial^\mu t_{\mu\nu} = \partial^\nu t_{\mu\nu} = 0$$

Hint: go to the special frame $(p^\mu) = (1, 1, 0, \ldots, 0)$ and use the gauge symmetry plus the transversality conditions to cancel all components except the transverse ones t_{ij}, with $i, j = 2, \ldots, d-1$.

For massless fields $p^2 = 0$ we may go to the special frame

$$p = (1, 1, 0_\perp)$$

In momentum space, the gauge symmetry and transversality conditions are

$$t_{\mu\nu} \to t_{\mu\nu} + p_\mu \lambda^L_\nu + p_\nu \lambda^R_\mu, \qquad p^\mu t_{\mu\nu} = t_{\mu\nu} p^\nu = 0.$$

The gauge variation of the time-time component is

$$t^{00} \to t^{00} + \lambda^0_L p^0 + \lambda^0_R p^0 = t^{00} + \lambda^0_L + \lambda^0_R$$

We can set $t^{00} \to 0$ by choosing

$$\lambda^0_L + \lambda^0_R = -t^{00}$$

With the same argument, we see that choosing

$$\lambda^0_L + \lambda^1_R = -t^{01}, \qquad \lambda^1_L + \lambda^0_R = -t^{10}, \qquad \lambda^1_L + \lambda^1_R = -t^{11}$$

eliminates all $t^{\mu\nu}$ with $\mu, \nu = 0, 1$. By the same trick, we can also make t^{0i} and t^{i0} vanish by choosing

$$\lambda^i_R = -t^{0i}, \qquad \lambda^i_L = -t^{i0}$$

At this point, we have used all $2d$ components of the gauge functions λ_L^μ and λ_R^μ. The remaining non-zero components of $t^{\mu\nu}$ are t^{1i}, t^{i1}, t^{ij} with i, j running in the $d-2$ transverse directions. However, the transversality conditions in this frame read

$$p_\mu \, t^{\mu\nu} = p_1 \, t^{1\nu} = t^{1\nu} = 0, \qquad t^{\mu\nu} \, p_\nu = t^{\mu 1} \, p_1 = t^{\mu 1} = 0$$

Therefore, we have cancelled all $t^{0\mu}, t^{1\mu}, t^{\mu 0}, t^{\mu 1}$ components and we are left only with transverse components.

In order to derive an action for the massless fields, we work in covariant notation, with all components of $t_{\mu\nu}$. We will enforce the transversality conditions by a projector. Let

$$P_{\mu\nu} = \eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}$$

Then the projected tensor

$$(t_P)_{\mu\nu} \equiv P_\mu{}^\rho \, t_{\rho\sigma} \, P^\sigma{}_\nu$$

is transverse and gauge-invariant.

Verify that $P_{\mu\nu}$ is a projector and that $(t_P)_{\mu\nu}$ is fully transverse and gauge-invariant

---

For $P_{\mu\nu}$ to be a projector it must square to itself:

$$P_{\mu\nu} \, P^{\nu\rho} = \left(\eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right)\left(\eta^{\nu\rho} - \frac{p^\nu p^\rho}{p^2}\right) = \delta^\rho_\mu - \frac{p_\nu p^\rho}{p^2} = P_\nu^\rho$$

Transversality is immediate:

$$p^\mu \, P_{\mu\nu} = p_\nu - p_\nu = 0 = P_{\nu\mu} \, p^\mu,$$

and from this it follows that t_P is gauge invariant, since the gauge variation of $t_{\mu\nu}$ involves terms proportional to p_μ and p_ν.

With these ingredients we can construct a gauge-invariant free action for all the degrees of freedom contained in t_P by writing:

$$S_{\text{free}} = \frac{1}{4} \int_x (t_P)^{\mu\nu} \, \partial^2 \, (t_P)_{\mu\nu} = -\frac{1}{4} \int_p t_P^{\mu\nu}(-p) \, p^2 \, t^P_{\mu\nu}(p) \tag{33}$$

We can now decompose the complete tensor into symmetric and antisymmetric parts:

$$t_{\mu\nu} = h_{\mu\nu} + b_{\mu\nu},$$

with $b_{\mu\nu} = -b_{\nu\mu}$, and $h_{\mu\nu} = h_{\nu\mu}$. Therefore, $h_{\mu\nu}$ should contain the graviton and the dilaton degrees of freedom. First we handle the antisymmetric part.

Extract from (33) the action of the antisymmetric field $b_{\mu\nu}$. Show that it can be written as

$$S[b_{\mu\nu}] = -\frac{1}{12} \int H_{\mu\nu\rho} \, H^{\mu\nu\rho},$$

where

$$H_{\mu\nu\rho} = \partial_\mu \, b_{\nu\rho} + \partial_\nu \, b_{\rho\mu} + \partial_\rho \, b_{\mu\nu}$$

It is convenient to manipulate (33) a bit further in the general case. Let us obviate the p^2 term and concentrate on the contraction of Lorentz indices. We can write

$$(t_P)_{\mu\nu}\,(t_P)^{\mu\nu} = \text{tr}\;\, t_P\,(t_P)^t$$

in matrix notation, where $(t_P)^t$ denotes the transposed matrix. Then we can put $t_P = P\,t\,P$ and

$$(t_P)_{\mu\nu}\,(t_P)^{\mu\nu} = \text{tr}\;P\,t\,P\,(P\,t\,P)^t = \text{tr}\;t\,P\,(P\,t)^t = (t\,P)_{\mu\nu}\,(P\,t)^{\mu\nu}.$$

Hence we find the structure

$$-\frac{1}{4}\,t_P^{\mu\nu}(-p)\,p^2\,t^P_{\mu\nu}(p) = -\frac{1}{4}\left[t_{\mu\nu}\,p^2\,t^{\mu\nu} - 2\,p^\mu\,t_{\mu\nu}\,(t^t)^{\nu\beta}\,p_\beta + t_{\mu\alpha}\,\frac{p^\alpha\,p^\nu\,p^\mu\,p^\beta}{p^2}\,t_{\beta\nu}\right]. \qquad (34)$$

We see that, unlike the case of the electromagnetic field in the notes, there is a surviving non-local term (non-analytic in momenta).

We can now project onto the antisymmetric part $t_{\mu\nu} \to b_{\mu\nu}$. Using the antisymmetry of $b_{\mu\nu}$ we find that all non-local terms cancel and we are left with

$$-\frac{1}{4}\,\left[b^{\mu\nu}\,p^2\,b_{\mu\nu} + 2\,p^\nu\,b_{\nu\mu}\,b^{\mu\sigma}\,p_\sigma\right] = -\frac{1}{12}\,H_{\mu\nu\rho}(-p)\,H^{\mu\nu\rho}(p), \qquad (35)$$

where

$$H_{\mu\nu\rho}(p) = i\,p_\mu\,b_{\nu\rho}(p) + i\,p_\nu\,b_{\rho\mu}(p) + i\,p_\rho\,b_{\mu\nu}(p)$$

Using the linearized approximation of the Ricci tensor:

$$\mathcal{R}^{(1)}_{\mu\nu}[g=\eta+h] = \frac{1}{2}\left(\partial^2\,h_{\mu\nu} - \partial^\lambda\,\partial_\mu\,h_{\lambda\nu} - \partial^\lambda\,\partial_\nu\,h_{\mu\lambda} + \partial_\mu\,\partial_\nu\,h^\lambda_\lambda\right) + O(h^2) \qquad (36)$$

and the Einstein–Hilbert Lagrangian

$$S_{\text{EH}} = \int d^dx\,\sqrt{-g}\,\mathcal{R},$$

with $\mathcal{R} = g^{\mu\nu}\mathcal{R}_{\mu\nu}$, find the free Lagrangian of a graviton, the so-called Fierz–Pauli action:

$$S_{\text{FP}} = -\frac{1}{4}\int_x h^{\mu\nu}\,\left(-\eta_{\mu\alpha}\,\eta_{\nu\beta}\,\partial^2 + 2\,\eta_{\mu\alpha}\,\partial_\nu\,\partial_\beta - 2\,\eta_{\alpha\beta}\,\partial_\mu\,\partial_\nu + \eta_{\mu\nu}\,\eta_{\alpha\beta}\,\partial^2\right)\,h^{\alpha\beta} \qquad (37)$$

Hint: the first variation of the EH Lagrangian under

$$g_{\mu\nu} = \eta_{\mu\nu} + \delta\,g_{\mu\nu}$$

is given by

$$\delta S_{\text{EH}} = \int \sqrt{-g}\,\left(\mathcal{R}_{\mu\nu} - \tfrac{1}{2}\,g_{\mu\nu}\,\mathcal{R}\right)\,\delta\,g^{\mu\nu} + \text{total derivatives}$$

At the level of the linearized theory, we simply need to evaluate the equation of motion to linear order in $h_{\mu\nu}$. The Ricci scalar to linear order is given by

$$\mathcal{R}^{(1)} = \eta^{\mu\nu}\,\mathcal{R}^{(1)}_{\mu\nu} = \partial^2\,h^\lambda_\lambda - \partial^\mu\,\partial^\nu\,h_{\mu\nu}$$

Thus, the quadratic Lagrangian from which these equations of motion derive is

$$S_{\rm FP} = \frac{1}{2}\int h^{\mu\nu}\,\left(\mathcal{R}^{(1)}_{\mu\nu}[h] - \tfrac{1}{2}\,\eta_{\mu\nu}\,\mathcal{R}^{(1)}[h]\right) = -\frac{1}{4}\int_p \left(h^{\mu\nu}\,p^2\,h_{\mu\nu} - 2\,p_\mu\,h^{\mu\nu}\,h_{\nu\lambda}\,p^\lambda + 2\,h^\lambda_\lambda\,p_\mu\,p_\nu\,h^{\mu\nu} - h^\lambda_\lambda\,p^2\,h^\lambda_\lambda\right),$$

which is the momentum space version of (37).

Show that the *local* part of (33) projected on the graviton degrees of freedom coincides with the Fierz–Pauli Lagrangian. Write the nonlocal part as a term quadratic in the linearized Ricci scalar.

The answer is

$$S_{\rm nonlocal} = \frac{1}{4}\int_x \mathcal{R}^{(1)}\,\frac{1}{\partial^2}\,\mathcal{R}^{(1)}.$$

Manipulating the action of the symmetric part just like before for the case of $b_{\mu\nu}$ we obtain

$$-\frac{1}{4}\left[h^{\mu\nu}\,p^2\,h_{\mu\nu} - 2\,p_\mu\,h^{\mu\nu}\,h_{\nu\lambda}\,p^\lambda + h_{\mu\nu}\,\frac{p^\mu p^\lambda p^\nu p^\sigma}{p^2}\,h_{\lambda\sigma}\right]$$

The main difference with (35) is that now the nonlocal terms do not cancel out of the symmetry of the field. Adding and subtracting the term

$$-\frac{1}{2}\,h^\lambda_\lambda\,p_\mu\,p_\nu\,h^{\mu\nu} + \frac{1}{4}\,h^\lambda_\lambda\,p^2\,h^\lambda_\lambda,$$

we recover the total symmetric Lagrangian as

$$S[h] = S_{\rm FP} + S_{\rm nonlocal},$$

where $S_{\rm FP}$ is the Fierz–Pauli action and

$$S_{\rm nonlocal} = \frac{1}{4}\int_x \left(\partial^\mu\,\partial^\nu\,h_{\mu\nu} - \partial^2\,h^\lambda_\lambda\right)\,\frac{1}{\partial^2}\,\left(\partial^\mu\,\partial^\nu\,h_{\mu\nu} - \partial^2\,h^\lambda_\lambda\right) = \frac{1}{4}\int_x \mathcal{R}^{(1)}\,\frac{1}{\partial^2}\,\mathcal{R}^{(1)} \qquad (38)$$

Starting from the action

$$S_{\rm eff} = \int d^dx\,\sqrt{-g}\,\mathcal{R} + \int d^dx\,\sqrt{-g}\,\left(\chi\,\mathcal{R} + (\partial\chi)^2\right)$$

with an extra scalar field χ, expand it in the linearized approximation and integrate-out χ in the tree-level approximation. The result should be (38).

Expanding to second order in any combination of $h_{\alpha\beta}$ or χ one finds

$$S_{\rm eff} = S_{\rm FP} + \int \left(\chi\,\mathcal{R}^{(1)} + (\partial\chi)^2\right)$$

Now the equations of motion of χ are:

$$\mathcal{R}^{(1)} - 2\,\partial^2\,\chi = 0$$

We can formally solve them as

$$\chi_{cl} = \frac{1}{2\partial^2}\,\mathcal{R}^{(1)},$$

so that, plugging them back into the effective action one obtains

$$S(\chi_{cl})_{\text{eff}} = S_{\text{FP}} + \frac{1}{2}\int \mathcal{R}^{(1)}\,\frac{1}{\partial^2}\,\mathcal{R}^{(1)} - \frac{1}{4}\int \partial^2\,\frac{1}{\partial^2}\,\mathcal{R}^{(1)}\,\frac{1}{\partial^2}\,\mathcal{R}^{(1)} = S_{\text{FP}} + \frac{1}{4}\int \mathcal{R}^{(1)}\,\frac{1}{\partial^2}\,\mathcal{R}^{(1)}$$

We conclude that, at the level of the linearized approximation, the non-local term in the effective Lagrangian for the graviton is due to having integrated out a massless scalar field with linear coupling to the curvature. This is the *dilaton* field. We can now put

$$\chi = -2\phi,$$

and write a non-linear extension of the action:

$$S = \int d^dx\,\sqrt{-g}\,e^{-2\phi}\left(\mathcal{R} + 4(\partial\phi)^2 - \frac{1}{12}H^2\right), \tag{39}$$

where

$$H^2 = g^{\mu\alpha}\,g^{\nu\beta}\,g^{\rho\gamma}\,H_{\mu\nu\rho}\,H_{\alpha\beta\gamma}, \qquad (\partial\phi)^2 = g^{\mu\nu}\,\partial_\mu\phi\,\partial_\nu\phi.$$

This is the classical effective action of the massless fields in the closed-string sector. According to this, the gravitational constant is proportional to the exponential of the dilaton expectation value:

$$\kappa \propto e^{\langle\phi\rangle}$$

This means that we can define the string coupling constant as

$$g_s = e^{\langle\phi\rangle}$$

The kinetic term of the dilaton in (39) has the "wrong" sign. This is because it mixes with the trace of the metric. We can disentangle it by going to the so-called Einstein frame. This is just a Weyl rescaling of the metric of the form

$$g_{\mu\nu} \longrightarrow e^{2\omega}\,g_{\mu\nu},$$

under which the Ricci scalar transforms as

$$\mathcal{R} \to e^{-2\omega}\left(\mathcal{R} - 2(d-1)\,D^2\omega - (d-1)(d-2)(\partial\omega)^2\right)$$

The function ω is chosen so that the resulting Newton's constant is field-independent.

Express the string effective action in the Einstein frame. Check that the dilaton's kinetic term comes out with the right sign.

Answer:

$$S_{\text{eff}} = \frac{1}{2\kappa^2}\int d^dx\,\sqrt{-g}\,\left(\mathcal{R} - \frac{4}{d-2}(\partial\phi)^2 - \frac{1}{12}\,e^{-8\phi/(d-2)}\,H^2\right).$$

Let us write $\phi \to \langle\phi\rangle + \phi$, so that the effective action after the Weyl rescaling will start with the term

$$\frac{1}{2\kappa^2}\int \sqrt{-g}\,\mathcal{R} + \ldots$$

In order to cancel the dilaton dependence in this term we need to choose the rescaling function in

$$g_{\mu\nu} \to e^{2\omega}\, g_{\mu\nu},$$

so that

$$\sqrt{-g}\, e^{-2\phi}\,\mathcal{R} \longrightarrow e^{d\omega}\, e^{-2\phi}\, e^{-2\omega}(\sqrt{-g}\,\mathcal{R} + \ldots) = \sqrt{-g}\,\mathcal{R} + \ldots$$

This fixes

$$\omega = \frac{2\phi}{d-2}$$

Then, the full curvature term scales:

$$\sqrt{-g}\, e^{-2\phi}\,\mathcal{R} \to \sqrt{-g}\left(\mathcal{R} - \frac{8(d-1)}{d-2}\, D^2\phi - \frac{4(d-1)}{d-2}(\partial\phi)^2\right)$$

We will neglect in the sequel the total covariant derivative that only leads to boundary terms. Thus we get the standard Einstein–Hilbert Lagrangian plus a contribution to the dilaton Lagrangian equal to

$$-4\frac{d-1}{d-2}\sqrt{-g}\,(\partial\phi)^2$$

The antisymmetric tensor term yields

$$\sqrt{-g}\, e^{-2\phi}\, H^2 \to e^{2d\phi/(d-2)}\, e^{-2\phi}\, e^{-6\phi/(d-2)}\,\sqrt{-g}\, H^2 = e^{-8\phi/(d-2)}\,\sqrt{-g}\, H^2$$

Finally, the scalar term in the original action gives

$$+4\sqrt{-g}\, e^{-2\phi}\, g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi \to 4\, e^{2d\phi/(d-2)}\, e^{-2\phi}\, e^{-4\phi/(d-2)}\,\sqrt{-g}\,(\partial\phi)^2 = 4\,\sqrt{-g}\,(\partial\phi)^2$$

Adding this contribution to the dilaton's kinetic term to the one coming from the transformation of the curvature we find the total action in the Einstein frame:

$$S_{\text{eff}} = \frac{1}{2\kappa^2}\int d^dx\,\sqrt{-g}\left(\mathcal{R} - \frac{4}{d-2}\,(\partial\phi)^2 - \frac{1}{12}\, e^{-8\phi/(d-2)}\, H^2\right)$$

On dimensional grounds $\kappa^2 \sim (\alpha')^4\, g_s^2$. The standard normalization of g_s is such that, for the maximally supersymmetric string theories in $d = 10$ one has:

$$2\kappa^2 = (2\pi)^7\, (\alpha')^4\, g_s^2$$

Problem 3

In this problem we count the asymptotic density of mass levels of a string theory. We consider first the case bosonic closed strings. The mass formula is

$$M^2 = \frac{2}{\alpha'}\left(N_L + N_R - \frac{d-2}{12}\right),$$

with $N_{L,R}$ the total oscillation numbers for left and right oscillators. The level matching constraint implies $N_L = N_R$. Otherwise the oscillators are independent. Therefore, the number of states with total oscillator number $N_L + N_R = 2N$ is given by $p(N)^2$ where $p(N)$ is the degeneracy due to purely left- or right-moving oscillators.

We start by counting oscillator states built from a single left-moving oscillator of a given frequency n and a given transverse direction. These are states of the form

$$(a_n^\dagger)^k \, |0\rangle$$

for any positive k. The total oscillator level is

$$N = n \cdot k$$

Obviously, there is only one such state up to bosonic symmetry. However, we will write this number in a barroque way:

$$1 = p(N) = \sum_k \delta_{(nk,N)} = \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} \sum_k x^{nk} = \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} \frac{1}{1-x^n},$$

i.e. we obtain a formula for $p(N)$ in terms of a generating functional.

Complete the calculation of the full generating functional by including all $d-2$ transverse oscillators of a given frequency, for all possible frequencies.

Answer:

$$p(N) = \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} \left(\prod_{n=1}^{\infty} \frac{1}{1-x^n} \right)^{d-2}.$$

Suppose now that we include all $d-2$ transverse oscillators of a given frequency n. Then the total oscillator level is

$$N = n \sum_j k_j, \qquad j = 2, 3, \ldots, d-1$$

The same as before applies with the replacement of a single sum by $d-2$ sums:

$$\sum_k \longrightarrow \sum_{k_2} \sum_{k_3} \cdots \sum_{k_{d-1}}$$

Accordingly

$$\sum_k x^{nk} \longrightarrow \sum_{\{k_j\}} x^{\sum_j nk_j} = \sum_{\{k_j\}} \prod_j x^{nk_j} = \prod_j \sum_{k_j} x^{nk_j} = \left(\frac{1}{1-x^n} \right)^{d-2}$$

Finally, we allow the frequency n to vary as well. This gives the same effect as before, except that the range of n goes up to infinity. Therefore we find

$$p(N) = \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} \left(\prod_{n=1}^{\infty} \frac{1}{1-x^n} \right)^{d-2}$$

The infinite product blows up near $x = 1$. To see this, write

$$\prod_n \frac{1}{1-x^n} = \exp\left(-\sum_n \log(1-x^n)\right) = \exp\left(\sum_{m,n=1}^{\infty} \frac{x^{mn}}{m}\right) = \exp\left(\sum_{m=1}^{\infty} \frac{x^m}{m(1-x^m)}\right) \tag{40}$$

Now, close to $x = 1$ the last exponent may be approximated by

$$\frac{1}{1-x} \sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{\pi^2}{6(1-x)}$$

Therefore, since the infinite product blows-up very fast when close to $x = 1$, and x^{N+1} is very small for large N and $x < 1$, there should be a sharp saddle point for the contour integral near $x = 1$.

Localize the saddle point and calculate the asymptotic form of $p(N)$ at large N. In order to get the subleading terms all right you will need the Hardy–Ramanujan formula:

$$f(x) = \left(-\frac{2\pi}{\log x}\right)^{-1/2} x^{1/24}\, y^{-1/12}\, f(y^2),$$

where

$$f(x) \equiv \prod_n \frac{1}{1-x^n}, \qquad y \equiv \exp\left(\frac{2\pi^2}{\log x}\right)$$

Answer:

$$p(N) \propto \frac{1}{N^{\frac{d+1}{4}}} \exp\left(4\pi\sqrt{\frac{d-2}{24}}\,\sqrt{N}\right).$$

As $y \to 0$, or $x \to 1$, the function $f(y^2)$ obviously tends to unity. On the other hand:

$$\log x \to x - 1$$

and therefore

$$y^{-1/12} \to \exp\left(\frac{\pi^2}{6(1-x)}\right)$$

Finally

$$\left(-\frac{2\pi}{\log x}\right)^{-1/2} \to \sqrt{\frac{1-x}{2\pi}}$$

All in all we get

$$[f(x)]^{d-2} \sim (1-x)^{(d-2)/2} \exp\left(\frac{\pi(d-2)}{6(1-x)}\right)$$

The exponent in the contour integral is then

$$\exp\left(-\frac{\pi^2(d-2)}{6\log x} - (N+1)\log x\right),$$

where we have replaced again $1 - x \to -\log x$. It is convenient to change variables to $x = e^w$ and write

$$p(N) \propto \oint \frac{dw}{2\pi i} (-w)^{(d-2)/2} e^{g(w)},$$

where

$$g(w) = -\frac{\pi^2(d-2)}{6\,w} - N\,w$$

There are saddle-points at the solutions of $g'(w_c) = 0$:

$$w_c = \pm\frac{\pi}{\sqrt{N}}\sqrt{\frac{d-2}{6}},$$

with classical action

$$g(w_c) = \mp 2\pi\sqrt{\frac{d-2}{6}}\,\sqrt{N}$$

Thus, only the one at negative w_c gives a dominant contribution at large N. The saddle point approximation to the integral around this gives

$$p(N) \propto (-w_c)^{\frac{d}{2}-1}\,\frac{1}{\sqrt{|g''(w_c)|}}\,e^{g(w_c)},$$

where

$$g''(w_c) = \frac{N^{3/2}}{\pi}\sqrt{\frac{d-2}{24}}$$

Since $g''(w_c)$ is positive, we have to rotate the contour so that $w - w_c$ becomes pure imaginary near the saddle, thus cancelling the factor of i in the integral measure. The final result is

$$p(N) \propto \frac{1}{N^{\frac{d+1}{4}}}\,\exp\left(4\pi\sqrt{\frac{d-2}{24}}\,\sqrt{N}\right)$$

Use the value found for $p(N)$ to compute the density of states as a function of the mass $\rho(M)$. This means that $\rho(M)\,dM$ gives the number of states with masses between M and $M + dM$.

The result is:

$$\rho(M) \propto \sqrt{\alpha'}\,\frac{e^{\beta_s M}}{(\alpha' M^2)^{d/2}},$$

where $\beta_s = 1/T_s$ is the inverse of the so-called Hagedorn temperature:

$$\beta_s = 4\pi\sqrt{\frac{d-2}{24}}\,\sqrt{\alpha'}.$$

We use the mass formula of the closed bosonic string

$$M^2 = \frac{4}{\alpha'}\left(N - \frac{d-2}{24}\right)$$

and the definition of the density of states:

$$\rho(M) = \sum_f \delta(M_f - M) = \sum_N p(N)^2\,\delta\left(M - \frac{2}{\sqrt{\alpha'}}\sqrt{N - \frac{d-2}{24}}\right),$$

where the index f runs over the infinite set of particle states of the string. Since $N \gg 1$ we can approximate the sum by an integral and write

$$\rho(M) \propto \sqrt{\alpha'} \int \frac{dN}{N^{\frac{d+1}{2}}} \, e^{8\pi\sqrt{\frac{d-2}{24}}\sqrt{N}} \, \delta\left(M - \sqrt{\frac{4N}{\alpha'}}\right)$$

The final result is

$$\rho(M) \propto \frac{\sqrt{\alpha'}}{(\alpha' M^2)^{d/2}} \, e^{\beta_s M},$$

with β_s as above.

Calculate the full asymptotic density of energy levels, i.e. the number $\omega(E)\, dE$ of states with energy in the interval $[E, E + dE]$ at $E\sqrt{\alpha'} \gg 1$.

Answer:

$$\omega(E) \propto m_s^{d-2} \, V \, \frac{e^{\beta_s E}}{(\ell_s E)^{\frac{d+1}{2}}}.$$

We have

$$\omega(E) = \sum_{f,\vec{p}} \delta(E - E_{f,\vec{p}}) = \sum_f \sum_{\vec{p}} \delta\left(E - \sqrt{\vec{p}^{\,2} + M_f^2}\right) = V \int \frac{d\vec{p}}{(2\pi)^{d-1}} \int dM \rho(M) \delta\left(E - \sqrt{\vec{p}^{\,2} + M^2}\right)$$

We can now write

$$dM = \frac{\sqrt{\vec{p}^{\,2} + M^2}}{\sqrt{E^2 - \vec{p}^{\,2}}} \, d\sqrt{\vec{p}^{\,2} + M^2}$$

and solve the delta function to get

$$\omega(E) \sim V \int d\vec{p} \, \frac{E}{(E^2 - \vec{p}^{\,2})^{\frac{d+1}{2}}} \, \exp\left(\beta_s \sqrt{E^2 - \vec{p}^{\,2}}\right)$$

Now, most of the contribution comes from the exponential degeneracy of mass levels $\rho(M) \sim \exp(\beta_s M)$ at $M \gg m_s$. Therefore we shall assume

$$E^2 - \vec{p}^{\,2} = M^2 \gg m_s^2,$$

and expand in powers of $\vec{p}^{\,2}/E^2$:

$$\sqrt{E^2 - \vec{p}^{\,2}} \approx E - \frac{\vec{p}^{\,2}}{2E} + \ldots,$$

so that the leading contribution to the density is

$$\omega(E) \sim V \, E \int d\vec{p} \, \frac{e^{\beta_s E}}{E^{d+1}} \, e^{-\beta_s \vec{p}^{\,2}/2E} \sim V \, E^{-\frac{d+1}{2}} \, e^{\beta_s E}$$

Restoring dimensional analysis we find the final result

$$\omega(E) = m_s^{d-2} \, V \, \frac{e^{\beta_s E}}{(\ell_s E)^{\frac{d+1}{2}}}$$

Prove that open strings and closed strings of the same type have the *same* Hagedorn temperature.

The Hagedorn temperature is determined by the exponetial growth of mass levels, which in turn is determined by the growth of oscillator states with oscillator number N. For open strings, there is only one set of oscillators of total level N (a combination of left- and right-moving oscillators). In closed strings, we had both left- and right-moving oscillators, each at the same total level N, because of the level matching condition. Therefore, for open strings we have a single factor of $p(N)$ instead of $p(N)^2$. This means that the exponential growth of mass levels

$$\rho(M)_{\text{open}} \sim p(N) \sim e^{C\sqrt{N}},$$

whereas we had for closed strings

$$\rho(M)_{\text{closed}} \sim p(N)^2 \sim e^{2C\sqrt{N}}.$$

This seems to imply that the Hagedorn temperatures differ by a factor of two. However, the relation between M and N is now $\ell_s M_{\text{open}} \sim \sqrt{N}$ instead of $\ell_s M_{\text{closed}} \sim \sqrt{4N} = 2\sqrt{N}$. Therefore, the two effects cancel one another and

$$\beta_s(\text{open}) = \beta_s(\text{closed})$$

What is the Hagedorn temperature of a string theory with supersymmetric spectrum?

Hint: Find the generating functional that counts both bosonic and fermionic oscillators and then reason like in (40). You will need the sum

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$$

The answer for the generating functional is

$$p(N) = \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} D_0 \left(\prod_{n=1}^{\infty} \frac{1+x^n}{1-x^n} \right)^{d-2},$$

where D_0 is the (finite) degeneracy of the oscillator ground state. The inverse Hagedorn temperature turns out to be

$$\beta_s(\text{susy}) = \pi \sqrt{d-2}\,\sqrt{\alpha'}.$$

With a supersymmetric spectrum most of the previous results go through with little difference, since we are just counting states. In the supersymmetric case the mass formula is

$$M^2_{\text{open}} = \frac{1}{\alpha'} (N_B + N_F),$$

where N_B and N_F denote the total oscillator numbers of bosonic and fermionic states respectively. For closed strings

$$M^2_{\text{closed}} = \frac{4}{\alpha'} (N_B + N_F),$$

where N_B, N_F correspond now to the purely left oscillator levels. As before, the important quantity is the exponential growth of $p(N)$ with $N = N_B + N_F$. Since we have now fermionic oscillators, the total level for a single frequency n and a single spin degree of freedom is

$$N = k \cdot n + s \cdot n$$

from

$$(a_n^\dagger)^k \, (b_n^\dagger)^s \, |0\rangle,$$

where k is a positive integer giving the bosonic occupation number of the oscillator with frequency n, but $s = 0, 1$ depending on whether the fermionic oscillator is (singly) occuppied or not. Therefore, in the generating functional

$$\sum_k x^{kn} \longrightarrow \sum_{k=0}^{\infty} \sum_{s=0,1} x^{kn+sn} = \frac{1+x^n}{1-x^n}$$

¿From here we get the basic modification:

$$p(N) = \frac{1}{2\pi i} \oint \frac{dx}{x^{N+1}} \, D_0 \left(\prod_{n=1}^{\infty} \frac{1+x^n}{1-x^n} \right)^{d-2},$$

where D_0 is the (finite) degeneracy of the ground states. To get the leading exponential asymptotics of $p(N)$ we estimate the infinite product near $x = 1$, just as before.

$$\prod_n \left(\frac{1+x^n}{1-x^n} \right) = \exp \left[\sum_n \log \left(\frac{1+x^n}{1-x^n} \right) \right] = \exp \left(2 \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \frac{x^{n(2k+1)}}{(2k+1)} \right) = \exp \left(\sum_k \frac{2x^{2k+1}}{(2k+1)(1-x^{2k+1})} \right)$$

Expanding the exponent close to $x = 1$ we get

$$\frac{2}{1-x} \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{4(1-x)}$$

Thus

$$\left(\prod_n \frac{1+x^n}{1-x^n} \right)^{d-2} \sim \exp \left(\frac{(d-2)\pi^2}{4(1-x)} \right)$$

This is exactly the same result of the bosonic string up to a factor of $\sqrt{3/2}$. Therefore, this factor propagates down to the calculation of the Hagedorn temperature and we have

$$\beta_s(\text{susy}) = \pi \sqrt{d-2} \sqrt{\alpha'}$$

Problem 4

In this problem we develop the basic facts about the Kaluza–Klein models. One starts by assuming that a $(d+1)$-dimensional space-time M_{d+1} has the structure

$$M_{d+1} = \mathbf{R}^d \times \mathbf{S}^1,$$

where the circle has radius R. We split the coordinates x^M as

$$(x^M) = (x^\mu, y), \qquad x \in \mathbf{R}^d, \qquad y \in [0, 2\pi R]$$

We assume that the circle is *isometric*, which means that translations along it are “rigid”:

$$\partial_y \, g_{MN} = 0$$

The most general $(d+1)$-dimensional metric with this property can be conveniently written as

$$ds^2_{d+1} = g_{MN}\, dx^M\, dx^N = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu + e^{2\sigma(x)}\; (dy + R\, A_\mu(x)\, dx^\mu)^2\,, \tag{41}$$

where the factor of R in front of A_μ is a conventional normalization. Thus, a $(d+1)$-dimensional metric (a symmetric tensor) splits into a d-dimensional metric, plus a d-dimensional vector and a d-dimensional scalar. The scalar $\sigma(x)$ is called *radion* and measures the physical size of $\mathbf{S}^1$ at a given point $x \in \mathbf{R}^d$:

$$\ell(\mathbf{S}^1_x) = \int_0^{2\pi R} dy\, \sqrt{g_{yy}(x)} = e^{\sigma(x)} \cdot 2\pi R$$

If σ has a constant expectation value $\langle\sigma\rangle$, we can absorb it into the definition of the radius R, so that we can set $\langle\sigma\rangle = 0$. If $\langle\sigma(x)\rangle$ is x-dependent, we have what is called “warped compactification” in modern parlance.

The main fact about the Kaluza–Klein set-up that makes it interesting is that general covariance of M_{d+1} descends to general covariance of $\mathbf{R}^d$ plus *ordinary* gauge symmetry. To see this, notice that since $\mathbf{S}^1$ is invariant under translations y, we can actually do this at each $x \in \mathbf{R}^d$ independently:

$$y \to y + R\,\lambda(x), \qquad dy \to dy + R\, \partial_\mu\, \lambda\, dx^\mu.$$

Therefore

$$dy + R\, A_\mu\, dx^\mu \longrightarrow dy + R(A_\mu + \partial_\mu \lambda)\, dx^\mu$$

So, the whole effect of the local $\mathbf{S}^1$-translation on the fields defined on $\mathbf{R}^d$ is just a gauge transformation of the vector field:

$$A_\mu \to A_\mu + \partial_\mu \lambda,$$

which becomes a $U(1)$ (Maxwell) gauge field. This can be generalized to any compact manifold with isometry group G. Thus, we have a geometrical origin of the gauge symmetry.

The most characteristic prediction of KK models is the existence of an infinite tower of massive “KK resonances”. Let ϕ be a bosonic scalar field defined on $\mathbf{R}^d \times \mathbf{S}^1$. It is periodic on the circle, so that we can do Fourier analysis and write:

$$\phi(x,y) = \sum_{n \in \mathbf{Z}} \phi_n(x)\; e^{iny/R}$$

If the field is massless in the higher-dimensional space:

$$\partial_M \partial^M\, \phi = 0,$$

then we have

$$0 = \partial_M \partial^M\, \phi = \partial_\mu \partial^\mu\, \phi + \partial_y^2\, \phi = \sum_n \left(\partial_\mu \partial^\mu\, \phi_n - \frac{n^2}{R^2}\, \phi_n \right)\; e^{iny/R}$$

Hence, the massless field in $d+1$ dimensions decomposes into an infinite set of “normal modes” on the circle, giving *massive* fields ϕ_n on $\mathbf{R}^d$ with mass

$$M_n = \frac{|n|}{R}$$

Prove that ϕ_n is charged with respect to A_μ and find the charge

---

Under $y \to y' = y + R\,\lambda(x)$ the higher-dimensional field $\phi(x,y)$ is a scalar:

$$\phi(x,y) \to \phi'(x,y') = \phi(x, y' - R\lambda(x))$$

and, to linear order in $\lambda(x)$:

$$\delta\,\phi(x,y) = \phi'(x,y) - \phi(x,y) = -R\,\lambda(x)\,\partial_y\,\phi(x,y).$$

Now write

$$\delta\,\phi(x,y) = \sum_n e^{iny/R}\,(\delta\phi)_n(x), \qquad \phi(x,y) = \sum_n e^{iny/R}\phi_n(x).$$

Combining both equations we obtain

$$(\delta\phi)_n(x) = -in\,\lambda(x)\,\phi_n(x),$$

which is nothing but the infinitesimal version of

$$\phi_n(x) \to e^{-in\lambda(x)}\,\phi_n(x),$$

a $U(1)$ gauge transformation with charge $Q_n = n$. The covariant derivative transforms as

$$(\partial_\mu + iQ_n\,A_\mu)\,\phi_n \to (\partial_\mu + iQ_n\,A'_\mu)\,\phi'_n,$$

where

$$(\partial_\mu + iQ_n\,A'_m u)\phi'_n = e^{-in\lambda}\,(\partial_\mu - in\,\partial_\mu\lambda + iQ_n\,A_\mu + iQ_n\,\partial_\mu\lambda)\;\phi_n = e^{-in\lambda}(\partial_\mu + iQ_n\,A_\mu)\phi_n$$

precisely if

$$Q_n = n$$

This construction generalizes to all higher-dimensional fields, including the graviton. In particular, the zero-modes of the graviton (corresponding to $n = 0$) yield the previously studied A_μ and σ.

There is an interesting interplay with supersymmetry, namely the KK reduction of massless states gives the simplest example of BPS-saturated states. Assume that the $(d+1)$-dimensional theory has a Majorana spinor-valued conserved charge Q_α satisfying a standard supersymmetry algebra

$$\{Q_\alpha,\,Q_\beta\} = C\,(\gamma^A)_{\alpha\beta}\,p_A,$$

where p_A are the momentum components in the $(d+1)$-dimensional space, γ^A are Dirac matrices:

$$\{\gamma^A,\,\gamma^B\} = 2\,\eta^{AB},$$

and C is a charge-conjugation matrix. We assume that the Dirac matrices can be chosen real (Majorana representation) and $C = \gamma^0$.

Consider a massive state in $d+1$ dimensions, satisfying the dispersion relation

$$p_A\,p^A + M_{d+1}^2 = 0.$$

For a state with momentum

$$(p_A) = (p_\mu, p_d) = (p_\mu, n/R)$$

we have

$$p_\mu p^\mu + \frac{n^2}{R^2} + M_{d+1}^2 = 0,$$

so that there is an effective d-dimensional mass given by

$$M^2 = M_{d+1}^2 + \frac{n^2}{R^2}$$

Thus, we find a relation with the structure of a BPS bound:

$$M \geq \frac{|n|}{R}$$

Prove that positivity of $\{Q, Q\}$ implies this BPS bound for all states with charge n. The bound is saturated by massless states in $d+1$ dimensions, if and only if some supersymmetry is unbroken. Find the broken and unbroken supercharges.

Answer: the unbroken supercharges satisfy

$$\gamma^0 \gamma^d Q_u = s(n) Q_u,$$

with $s(n) = \text{sign}\,(n)$. The broken supercharges, Q_b, are the complement:

$$\gamma^0 \gamma^d Q_b = -s(n) Q_b.$$

In the rest frame in $\mathbf{R}^d$, the momentum of the massive particle can be chosen as

$$P^A = (M, 0, 0, \ldots, n/R)$$

The Susy algebra on this frame reduces to

$$\{Q, Q\} = M + \gamma^0 \gamma^d \frac{n}{R}$$

¿From the Dirac algebra we have

$$\text{Tr}\, \gamma^0 \gamma^d = \frac{1}{2}\, \text{Tr}\, \{\gamma^0, \gamma^d\} = 0, \qquad (\gamma^0 \gamma^d)^2 = -(\gamma^0)^2 (\gamma^d)^2 = 1$$

Therefore, the matrix $\gamma^0 \gamma^d$ has an equal number of ± 1 eigenvalues. Positivity of $\{Q, Q\}$ on the eigenvalue basis then implies

$$M \pm \frac{n}{R} \geq 0,$$

which proves the BPS bound. The bound is saturated for massless fields in $(d+1)$-dimensions, so that

$$\frac{n}{R} = s(n)\, M,$$

where $s(n) = \pm 1$ is the sign of n. In this case the algebra is

$$\{Q, Q\} = M\, \left[1 + s(n)\, \gamma^0 \gamma^d\right]$$

The matrix in brackets is now proportional to a projector, with half of the eigenvalues vanishing. Let us write

$$Q_u = P_u\, Q$$

for the unbroken supercharges, with P_u a projector. For the unbroken supercharges, the matrix

$$\{Q_u,\, Q_u\} = P_u\, \{Q,\, Q\}\, P_u^t$$

vanishes over the BPS state. Therefore, P_u must be the complement of the projector appearing in $\{Q,\, Q\}$, i.e.

$$P_u = \tfrac{1}{2}\,(1 - s(n)\,\gamma^0\,\gamma^d)$$

Thus, the unbroken supercharges satisfy

$$\gamma^0\,\gamma^d\, Q_u = s(n)\, Q_u,$$

whereas the broken supercharges, Q_b, are the complement:

$$\gamma^0\,\gamma^d\, Q_b = -s(n)\, Q_b$$

With fermions there is more freedom in the KK reduction. A massless Dirac fermion in $(d+1)$ dimensions has equation:

$$i\,\partial\!\!\!/\,\psi = 0 = (i\,\partial\!\!\!/)^2\,\psi = -\partial^2\,\psi$$

If the action is invariant under a “lepton number” symmetry:

$$\psi \to e^{i\alpha}\,\psi$$

(this is always true at least for $\alpha = \pi$), then more general boundary conditions (other than periodic) are possible, and we can have a tower of fermions in $\mathbf{R}^d$ with α-dependent masses.

Find the generalized boundary conditions and the KK spectrum for arbitrary $\alpha \in [0, 2\pi]$.

Answer:

$$M_n = \frac{1}{R}\left|n + \frac{\alpha}{2\pi}\right|$$

The phase symmetry of the action means that we can identify the fermion after a circulation of $\mathbf{S}^1$ up to a phase transformation:

$$\psi(x, y + 2\pi R) = e^{i\alpha}\,\psi(x, y)$$

Notice, incidentally, that we still keep α an x-independent constant. The normal mode decomposition compatible with this boundary conditions is

$$\psi(x, y) = \sum_{n\in\mathbf{Z}} \psi_n(x)\; e^{i(n+\alpha/2\pi)y/R}$$

Hence, the field equation becomes

$$\partial^2\,\psi = \sum_n e^{iny/R}\left[-\frac{1}{R^2}\,\left(n + \frac{\alpha}{2\pi}\right)^2 + \partial_\mu\partial^\mu\right]\;\psi_n = 0,$$

which leads to the mass spectrum:

$$M_n = \frac{1}{R}\left|n + \frac{\alpha}{2\pi}\right|$$

In particular, $\alpha = 0$ is the supersymmetric case. Thus, turning on α is a way of enforcing a "soft breaking" of supersymmetry. This is called the Scherk–Schwarz mechanism.

Since the $(d+1)$-dimensional metric induces d-dimensional metric, gauge field and scalar degrees of freedom, we expect that the Einstein–Hilbert action on the whole space will generate the corresponding Maxwell term in the d-dimensional action.

Find the action for $g_{\mu\nu}, A_\mu, \sigma$ that is induced on $\mathbf{R}^d$. Present the result in the Einstein frame, where the Newton's constant is really *constant*.

In order to prove this, use the following identities for the Ricci scalar. For the Kaluza–Klein metric decomposition (41) one has

$$\mathcal{R}_{(d+1)} = \mathcal{R} - 2\, e^{-\sigma} D^2 e^{\sigma} - \frac{R^2}{4} e^{2\sigma} F^2,$$

where

$$F^2 = g^{\mu\nu} g^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta}, \qquad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$$

Second, under a Weyl rescaling of the d-dimensional metric,

$$g_{\mu\nu} \to e^{2\omega} g_{\mu\nu},$$

the d-dimensional Ricci scalar transforms as

$$\mathcal{R} \to e^{-2\omega}\left(\mathcal{R} - 2(d-1)\, D^2\omega - (d-1)(d-2)(\partial\omega)^2\right)$$

Answer:

$$S_d = \frac{2\pi R}{2\kappa_{d+1}^2} \int_{\mathbf{R}^d} \left(\mathcal{R} - \left(\frac{d-1}{d-2}\right)(\partial\,\sigma)^2 - \frac{R^2}{4} e^{2(d-1)\sigma/(d-2)} F^2\right)$$

Taking the determinant of the higher-dimensional metric:

$$\det(g_{MN}) = e^{2\sigma} \det(g_{\mu\nu}) = g\, e^{2\sigma}$$

Thus, one finds

$$\frac{1}{2\kappa_{d+1}^2} \int_0^{2\pi R} dy \int d^dx \sqrt{-\det(g_{MN})}\, \mathcal{R}_{(d+1)} = \frac{2\pi R}{2\kappa_{d+1}^2} \int d^dx \sqrt{-g}\, e^{\sigma} \left(\mathcal{R} - 2e^{-\sigma} D^2 e^{\sigma} - \frac{R^2}{4} e^{2\sigma} F^2\right)$$

Neglecting total derivatives on $\mathbf{R}^d$ of the form

$$\int_{\mathbf{R}^d} \sqrt{-g}\, D_\mu (\ldots)^\mu = \int_{\mathbf{R}^d} \partial_\mu (\ldots)^\mu$$

we find

$$\frac{\pi R}{\kappa_{d+1}^2} \int \left(e^{\sigma} \mathcal{R} - \frac{R^2}{4} e^{3\sigma} F^2\right)$$

Even if we assume $\langle\sigma\rangle = 0$, the effective gravitational constant in $\mathbf{R}^d$ still depends on the σ field. In order to remove it we perform a Weyl rescaling

$$g_{\mu\nu} \to e^{2\omega}\, g_{\mu\nu}$$

in such a way that

$$e^{\sigma}\sqrt{-g}\,\mathcal{R} \to e^{\sigma}\,(e^{\omega})^d\, e^{-2\sigma}\,\sqrt{-g}\,\mathcal{R} + \ldots = \sqrt{-g}\,\mathcal{R} + \ldots$$

Therefore, we need

$$\omega = -\frac{\sigma}{d-2}$$

and the full transformation of the Einstein–Hilbert term is

$$e^{\sigma}\sqrt{-g}\,\mathcal{R} \longrightarrow \sqrt{-g}\,\left(\mathcal{R} + \frac{2(d-1)}{d-2}\,D^2\sigma - \frac{d-1}{d-2}\,(\partial\sigma)^2\right).$$

The Maxwell term transforms as

$$e^{3\sigma}\sqrt{-g}g^{\mu\nu}g^{\alpha\beta}F_{\mu\alpha}F_{\nu\beta} \to e^{3\sigma}\,e^{-\sigma d/(d-2)}\left(e^{-2\sigma/(d-2)}\right)^{-2}\sqrt{-g}\,F^2 = e^{2(d-1)\sigma/(d-2)}\sqrt{-g}\,F^2$$

So, neglecting total derivatives we finally get

$$S_d = \frac{2\pi R}{2\kappa_{d+1}^2}\int_{\mathbf{R}^d}\left(\mathcal{R} - \left(\frac{d-1}{d-2}\right)(\partial\,\sigma)^2 - \frac{R^2}{4}\,e^{2(d-1)\sigma/(d-2)}\,F^2\right)$$

Now the gauge fields have a field-dependent coupling. If the radion was to obtain an effective mass by some quantum corrections, so that we can replace σ by its expectation value, then we would have standard electromagnetism below the energy scale of the mass of σ.

Use the previous action to find a relation between the effective Planck mass in d dimensions and the gauge coupling. By comparing the mass and the charge of KK resonances, show that a realization of electromagnetism *a la* Kaluza–Klein is too naive an idea.

The previous effective action is valid at energy scales

$$E \ll \min\,(M_P, 1/R),$$

with M_P the Planck mass of the effective theory on $\mathbf{R}^d$. The M_P threshold is from the non-renormalizability of gravity, and the threshold $1/R$ is from our neglect of the higher KK resonances. We define the Planck mass as

$$M_P = \frac{1}{\ell_P} = \left(\frac{1}{G}\right)^{\frac{1}{d-2}} = \left(\frac{8\pi}{\kappa^2}\right)^{\frac{1}{d-2}}$$

The parameters of the KK model are the $(d+1)$-dimensional gravitational coupling κ_{d+1} and the radius of the extra dimension R (assuming $\langle\sigma\rangle = 0$). The predictions for the d-dimensional gravitational and gauge couplings are:

$$\kappa^2 = \frac{\kappa_{d+1}^2}{2\pi R}, \qquad g^2 = \frac{\kappa^2}{\pi R^3} \tag{42}$$

We can eliminate the higher-dimensional gravitational constant and derive a relation between the KK radius, the Newton constant and the fine structure constant $\alpha = g^2/4\pi$. Using $16\pi G_{\rm N} = 2\kappa^2$ in four dimensions we obtain from (42):

$$\alpha = \frac{4G_{\rm N}}{R^2}$$

Plugging the four-dimensional values we obtain a formula for any KK radius that would be related to the electromagnetic coupling:

$$\frac{1}{R} \sim 2\sqrt{\alpha} \cdot 10^{19}\,\text{GeV}$$

Since $\alpha_{em}(E) > 10^{-2}$ for $E > 1\text{TeV}$ we find that

$$\frac{1}{R} > 10^{17}\ \text{GeV}$$

Therefore, the mass of electrically charged particles cannot be of KK origin.

What is the "unification" scale between electromagnetism and gravity in the KK scenario?

It should be $M_X \sim 1/R$ since above this scale we only have $(d+1)$-dimensional gravity. To see this explicitly, consider $d = 4$ and the relation

$$g^2 \sim \kappa^2/R^2$$

The dimensionless effective gravitational coupling $\alpha_G(E) \sim \kappa^2 E^2$ becomes of the order of the dimensionless electromagnetic coupling $\alpha_{em} = g^2/4\pi$ at energies of order

$$M_X \sim 1/R$$

The condition for this to occur within the weak-coupling regime of the $U(1)$ theory is that $R \gg \ell_P$.

Problem 5

The purpose of this problem is to give very general arguments on the expected size of nonperturbative effects in string theory. It is a general fact of perturbation theory in field theory that the set of diagrams with n vertices is of order $n!$. This gives the dominant growth of perturbative amplitudes:

$$\mathcal{A} = \sum_n \mathcal{A}_n, \qquad \mathcal{A}_n \sim g^{2n}\, n!$$

In general, such series are neither convergent nor Borel summable. At best, they are asymptotic. We can obtain this estimate by counting the number of Feynman diagrams of a given order. For the purposes of counting diagrams we can just set all propagators to unity, i.e. we can consider a field theory in zero dimensions. For such a theory, the path integral is just an ordinary finite-dimensional integral. Let us take a cubic interaction for definiteness:

$$\int dx\, \exp\left(-\frac{x^2}{2} - g\,\frac{x^3}{3!}\right)$$

The integral only makes sense as a formal expansion in powers of g, but this is enough.

Find the asymptotics of the coefficients of the perturbative expansion of the previous integral. You will need the formula for Euler's Gamma:

$$\Gamma(z) = \int_0^\infty du\, e^{-u}\, u^{z-1},$$

and you also need to derive Stirling's approximation for large $n!$.

Rescaling $x \to g\,x$ we get

$$\frac{1}{g}\int dx\,\exp\left[\frac{1}{g^2}\left(\frac{x^2}{2}+\frac{x^3}{3!}\right)\right],$$

which shows that g^2 is the loop-counting parameter. Since the one-loop term corresponding to the gaussian approximation is of order $g^0 = 1$, we see that the term with power g^{2k} is from $k+1$ loops.

Expanding in powers of the interaction:

$$\sum_{n=0}^{\infty}\frac{(-g)^n}{6^n\,n!}\int dx\,x^{3n}\,e^{-x^2/2} = \sqrt{2\pi}\sum_{k=0}^{\infty} g^{2k}\,\frac{2^k\,\Gamma(3k+\frac{1}{2})}{\sqrt{\pi}\,3^{2k}\,(2k)!}$$

We can now use Stirling's approximation $n! \approx n^n\,e^{-n}$ to obtain

$$\sum_k g^{2k}\,c_k\,k! \sim \sum_k g^{2k}\,c_k\,k^k,$$

where $c_k \sim k^a\,b^k$ for bounded a, b.

To get Stirling's, take logarithms

$$\log\,n! = \sum_n \log\,n \approx \int_0^n dx\,\log\,x = n\,\log\,n - n$$

Asymptotic expansions have a limited accuracy for a given value of the coupling. One sees that higher-order terms give succesive better approximations up to a critical order beyond which the series gets out of control, in the sense that the terms $\mathcal{A}_n$ grow without bound. Therefore, the size of $|\mathcal{A}_n|$ at the critical order at which the series gets "crazy" provides an estimate of the maximum perturbative accuracy that can be achieved with a given value of the coupling constant. Hence this is also a bound on the size of possible nonperturbative effects.

Find the size of nonperturbative ambiguities at weak coupling for a series whose leading growth is

$$\mathcal{A}_n \sim g^{2n}\,(qn)!,$$

with q a positive integer.

To do this, we determine the order in perturbation theory at which two succesive terms are of the same order of magnitude. Given

$$\mathcal{A}_n \sim (qn)!\,g^{2n},$$

the required condition is

$$1 \sim \left|\frac{\mathcal{A}_{n+1}}{\mathcal{A}_n}\right| \sim \frac{(qn+q)!\,g^{2n+2}}{(qn)!\,g^{2n}} \sim (qn+q)(qn+q-1)\cdots(qn+1)\cdot g^2 \sim (qn)^q\,g^2,$$

thus, the critical n_c is given by

$$n_c = \frac{C_q}{q}\left(\frac{1}{g^2}\right)^{1/q}$$

This illustrates a typical property of asymptotic expansions. In order to increase the precision we must lower the size of the coupling.

The nonperturbative ambiguity is determined by the size of the amplitude at the critical order

$$|\mathcal{A}_{n_c}| \sim (qn_c)!\, g^{2n_c} \sim \left[(qn_c)^q\, g^2\right]^{n_c}\; e^{-qn_c} \sim e^{-qn_c} \sim \exp\left(-\frac{C_q}{g^{2/q}}\right)$$

In field theory $q = 1$, but in string theory $q = 2$. The argument for this is roughly the following.

In string theory, the field-theoretical estimate based on three-point vertices is accurate for the case of open-string perturbation theory, because there is an open-string field theory with a cubic vertex. So, open-string nonperturbative amplitudes are of order

$$\exp(-C/g_o^2),$$

where g_o is the open-string coupling. It is related to the closed string coupling by $g_s = g_o^2$. Therefore, the size of nonperturbative effects in closed-string theory is at least

$$\exp(-C/g_s)$$

In turn, this means that the growth of perturbation theory in closed-string theory is of the order of

$$\sum_n (g_s)^{2n}(2n)!,$$

i.e. much harder than in field theory. If we were to realize these effects semiclassically via some tunneling event, the action of the nonperturbative object that dominates the tunneling is of order $1/g_s$. This is precisely the action of D-branes.

Problem 6

In this problem we study the basic example of string duality: the duality relations between the maximally supersymmetric string theories in $d = 10$, that is the so-called type IIA and type IIB string theories, and M-theory in $d = 11$.

The basic assumption for this problem is the equivalence of M-theory on $\mathbf{S}^1_R \times \mathbf{R}^{10}$ and IIA string theory on $\mathbf{R}^{10}$. The eleven-dimensional theory is only parametrized by the Planck length ℓ_p. The ten-dimensional type IIA string theory is parametrized by the string length ℓ_s and the string coupling g_s. The relations that embody the duality are

$$R = g_s\,\ell_s, \qquad \ell_p^3 = g_s\,\ell_s^3 \tag{43}$$

We can motivate this as follows. From the basic rule of Kaluza–Klein reduction on a circle, the Newton constants in eleven and ten dimensions are related by

$$\frac{2\pi R}{G_{(11)}} = \frac{1}{G_{(10)}}$$

Neglecting numerical factors, using that

$$G_{(11)} \sim \ell_p^9, \qquad G_{(10)} \sim g_s^2\,\ell_s^8$$

we derive one relation:

$$\ell_p^9 \sim R\, g_s^2\, \ell_s^8$$

Now, eleven-dimensional supergravity has a solitonic membrane excitation whose tension is governed by the only dimensionfull parameter of the theory:

$$T_{\rm M2} \sim (\ell_p)^{-3}$$

by dimensional analysis. If the M2 wraps the compact circle of radius R, in the low-energy ten-dimensional description it looks like a string. If we identify this with the IIA string, the tensions are related by

$$T_{\rm M2} = 2\pi R\, T_{\rm F1} = \frac{R}{\ell_s^2}$$

This gives the second relation between the parameters of the dual theories. By choosing a normalization convention we obtain the relations (43).

The second assumption for the problem is that IIA and IIB theories in ten dimensions are related by T-duality. The rules for T-duality are

$$\ell_s \to \ell_s, \qquad R \to \ell_s^2/R, \qquad g_s \to g_s\, \ell_s/R,$$

and were derived in the notes.

By iterating these two basic duality transformations, derive a self-duality transformation (S-duality) of IIB strings in $d = 10$. What is the geometrical interpretation of the IIB S-duality in terms of the eleven-dimensional M-theory? What is the physical length scale of type IIB string theory that is left invariant by this S-duality?

We start with M-theory compactified on *two* circles of radii R and R', i.e. a background $\mathbf{S}^1_R \times \mathbf{S}^1_{R'} \times \mathbf{R}^9$. Going down to ten dimensions through $\mathbf{S}^1_{R'}$ gives a IIA string theory on $\mathbf{S}^1_R \times \mathbf{R}^9$ with parameters

$$\ell_s = \ell_p\, (\ell_p/R')^{1/2}, \qquad g_s = (R'/\ell_p)^{3/2}$$

Now we can do T-duality to get a type $\widetilde{\rm IIB}$ theory on $\mathbf{S}^1_{\tilde R}$ with

$$\tilde R = \ell_s^2/R, \qquad \tilde \ell_s = \ell_s, \qquad \tilde g_s = \ell_s\, g_s/R$$

On the other hand, descending through the circle of radius R we land in IIA$'$ theory on $\mathbf{S}^1_{R'} \times \mathbf{R}^9$ with primed parameters

$$\ell_s' = \ell_p\, (\ell_p/R)^{1/2}, \qquad g_s' = (R/\ell_p)^{3/2}$$

Under further T-duality we arrive at a type $\widetilde{\rm IIB}'$ string theory on $\mathbf{S}^1_{\tilde R'} \times \mathbf{R}^9$ with parameters

$$\tilde R' = (\ell_s')^2/R', \qquad \tilde \ell_s' = \ell_s', \qquad \tilde g_s' = \ell_s'\, g_s'/R'$$

Now, from these relations we obtain

$$\tilde g_s = g_s\, \frac{\ell_s}{R} = \frac{R'}{R}, \qquad \tilde g_s' = g_s'\, \frac{\ell_s'}{R'} = \frac{R}{R'},$$

from which we derive the first duality transformation between the two IIB theories:

$$\tilde g_s' = \frac{1}{\tilde g_s}$$

Thus, it is clearly a nonperturbative (S) duality. The other relation can be obtained form the calculation of the Newton constant in the IIB models:

$$\tilde{G}_{(10)} \sim (\tilde{\ell}_s)^8\,(\tilde{g}_s)^2 = \ell_s^8\,g_s^2\,\frac{\tilde{R}}{R}\,\ell_s^8\,g_s^2\,\frac{\ell_s^2}{R^2} = \ell_s^8\frac{R'^2}{R^2} = \left(\frac{(\ell_p)^6}{R'R}\right)^2$$

Since the result is invariant under permutation of the primes, we find that the physical scale left invariant by the S-duality is the ten-dimensional Planck length, or

$$(\tilde{\ell}'_s)^4\,\tilde{g}'_s = (\tilde{\ell}_s)^4\,\tilde{g}_s$$

This, together with the duality of the couplings, gives the other duality relation:

$$\tilde{\ell}'_s = \sqrt{\tilde{g}_s}\,\tilde{\ell}_s$$

The interpretation in the eleven-dimensional M-theory is just as a permutation of the two circles of the compactification. Therefore, the nonperturbative S-duality of type IIB is a completely obvious symmetry in eleven dimensions. This is a general rule of duality symmetries in string theory.

Use the results of this problem to "define" type IIB string theory in $\mathbf{R}^{10}$ in terms of a degenerate limit of a compactification in M-theory.

To define both dual IIB theories in $\mathbf{R}^{10}$ we require taking $\tilde{R}, \tilde{R}' \to \infty$ at fixed values of the string length and string coupling of both dual IIB theories. From the previous relations we learn that this requires $R, R' \to 0$ with a fixed ratio R/R'. Therefore type IIB string theory emerges from M-theory by compactification on a torus of zero size and fixed complex structure.

From the knowledge that type IIB F1-strings and D1-strings are S-dual of one another, calculate the tension of the Dirichlet string in type IIB theory.

The just found IIB S-duality is given by

$$g_s \to 1/g_s, \qquad \ell_s^2 \to g_s\,\ell_s^2$$

Under this mapping, the tension of a fundamental string

$$T_{\rm F1} = \frac{1}{2\pi\ell_s^2}$$

maps to the tension of a D1-string:

$$T_{\rm D1} = \frac{1}{2\pi\,g_s\,\ell_s^2}$$

From the knowledge that type IIA D2-branes and D0-branes are T-dual to type IIB D1-branes, calculate the tensions of all D-branes in either IIA or IIB theory. Check that D0-branes of IIA theory are the KK modes of M-theory on $\mathbf{S}^1 \times \mathbf{R}^{10}$.

Let us wrap a D1-brane on a circle of radius R. Its mass is given by

$$M = 2\pi R T_{\rm D1} = \frac{R}{g_s\, \ell_s^2}$$

Under T-duality $\ell_s \to \ell_s, R \to \ell_s^2/R, g_s \to g_s\ell_s/R$ it must map to the mass of a D0-brane:

$$M_{\rm D0} = \frac{1}{g_s \ell_s} = 2\pi\ell_s\, T_{\rm D1}$$

Now, from the basic M-theory–IIA duality relation $R = g_s\ell_s$ we find that

$$M_{\rm D0} = \frac{1}{R}$$

and indeed the D0-brane is viewed as a KK resonance from the point of view of the eleven-dimensional M-theory.

Now, whenever we T-dualize a wrapped brane we find, by exactly the same manipulation as in the D1–D0 case here, the invariant ratio

$$\frac{T_{{\rm D}p}}{T_{{\rm D}(p-1)}} = \frac{1}{2\pi\ell_s}$$

(remember that ℓ_s is invariant under T-duality). Therefore, applying this relation iteratively we can get the tension of all even (IIA) and odd (IIB) Dp-branes:

$$T_{{\rm D}p} = \frac{2\pi}{g_s(2\pi\ell_s)^p}$$

Problem 7

In this problem we practice with the non-linear effects of the gauge theory that appears on the world-volume of Dp-branes. For a single Dp-brane we have a photon with an effective action of the Dirac–Born–Infeld type:

$$S_{\rm DBI} = -T_{{\rm D}p} \int_{\Sigma_{p+1}} \sqrt{-\det(\eta_{ab} + 2\pi\alpha'\, F_{ab})}$$

There are also other important terms, of Chern–Simons type, that we neglect here. This effective action is valid for constant field strengths. For space-time varying field strengths there are corrections in powers of derivatives of F_{ab}.

Calculate the coupling constant of the world-volume photon field as a function of the fundamental parameters of the string theory m_s and g_s. Compute also the coefficient of the effective higher-dimensional operators with four powers of the field strength.

We simply have to expand the DBI action in powers of the field strength and isolate the Maxwell term

$$\mathcal{L}_{\rm Maxwell} = -\frac{1}{4g^2}\, F_{ab}\, F^{ab}$$

that determines the coupling constant. In order to make things easier, we make a Wick rotation to Euclidean signature and rotate back to Minkowski signature after the expansion has been completed. Under the rotation we have the usual rule

$$i\, S_{\rm DBI} \to -T_{{\rm D}p} \int \sqrt{\det(1+2\pi\alpha' F)},$$

where now the integral is over a positive-signature Euclidean world-volume, $\eta_{ab} \to \delta_{ab}$, and F_{ab} is an antisymmetric matrix. We write

$$\sqrt{\det(\delta_{ab}+2\pi\alpha' F_{ab})} = \exp\left(\tfrac{1}{2}\ {\rm tr}\ \ \log\left(\delta_{ab}+2\pi\alpha' F_{ab}\right)\right),$$

where the trace is over space-time indices. Now it is a simple matter to do a Taylor expansion in powers of F_{ab}, noticing that, because of antisymmetry tr $F^{2n+1}=0$ and we only have even powers. We find

$$1+\frac{1}{4}(2\pi\alpha')^2\, F_{ab}F^{ab} - \frac{1}{8}(2\pi\alpha')^4 \left(F_{ab}F^{bc}F_{cd}F^{da} - \frac{1}{4}(F_{ab}F^{ab})^2\right) + O(F^6)$$

Now we can rotate back to Minkowski signature by raising and lowering indices in the Minkowski metric η_{ab}. The result is

$$S_{\rm DBI} = -\int_{\Sigma_{p+1}} \left[T_{{\rm D}p} + \frac{1}{4g^2}\, F_{ab}F^{ab} + C\, \left(F^4 - \frac{1}{4}(F^2)^2\right)\right] + O(F^6)$$

The first term is just the action of a Dp-brane at rest:

$$-T_{{\rm D}p}\,{\rm Vol}\,(\Sigma_{p+1}) = -\int dt\ M_{{\rm D}p}$$

The second term is the Maxwell action for the gauge field with coupling

$$g^2 = g_s\,(2\pi)^{p-2}\,(m_s)^{3-p}$$

The third term indicates that the operators with four powers of the field strength combine into a single linear combination with the global coefficient

$$C = -\frac{\ell_s^{7-p}}{8(2\pi)^{p-4}\, g_s}$$

Consider a *static* D1-string which is wrapped on a circle of radius R. Show that the DBI action in the temporal gauge $A_0 = 0$ has the form (in units $2\pi\alpha' = 1$):

$$S_{\rm DBI} = -\frac{1}{g_s}\int dt\,dx\,\sqrt{1-\dot{A}_x^2},$$

where A_x is the spatial component of the gauge field.

In the temporal gauge $A_0 = 0$ the only nontrivial component of the field strength is the electric field $F_{01} = -F_{01} = \dot{A}_x$. The matrix in the DBI action is

$$(\eta+F)_{ab} = \begin{pmatrix} -1 & F_{01} \\ -F_{01} & 1 \end{pmatrix}.$$

so that

$$-\det(\eta + F) = 1 - \dot{A}_x^2$$

Show that $\dot{A}_x$ can be taken independent of x and the constant mode of A_x is valued in a *finite* interval $A_x \in [0, 1/R]$.

Even if we can set $A_0 = 0$ as a gauge choice, we should not forget to enforce its equation of motion, i.e. the Gauss constraint on the electric field:

$$\partial_1 F_{01} = \partial_x \dot{A}_x = 0$$

This proves the first part of the statement.

The remaining gauge transformations that are not fixed by $A_0 = 0$ are

$$A_x \longrightarrow A_x + \partial_x \lambda(x)$$

The gauge transformation $\lambda(x)$ is a function of the spatial circle parametrized by the coordinate x to the $U(1)$ group. Since $U(1)$ is the group of phases, the gauge transformation is globally

$$A_x \longrightarrow A_x + \frac{1}{i} U^{-1} \partial_x U, \qquad U = e^{i\lambda}$$

Therefore, λ is only defined up to a shift by an integer multiple of 2π. The gauge transformations are classified topologically by the quasiperiodicity of $\lambda(x)$:

$$\lambda(2\pi R) = \lambda(0) + 2\pi n$$

On a given topologial sector we can write

$$\lambda(x) = \lambda'(x) + \lambda_n(x),$$

where $\lambda'(x)$ is strictly periodic and λ_n is linear:

$$\lambda_n(x) = \frac{n}{R} x$$

Hence, the constant mode of A_x transforms under a linear gauge transformation as

$$A_x \longrightarrow A_x + \frac{n}{R}$$

The primitive identification is then up to constant shifts by $1/R$.

Use this result to prove that the electric field is quantized. Find the quantization rule. Is there a maximum value of the electric field?

The previous result means that we may consider a one-dimensional particle Lagrangian with coordinate $X(t) \equiv A_x(t)$ that is identified by

$$X(t) \equiv X(t) + \frac{1}{R}$$

Therefore, we have a particle effectively living on a circle of radius $1/2\pi R$. The conjugate momentum to this particle P_X is the generator of translations, and thus must be quantized in units of the inverse effective radius. Thus, physical states must have

$$P_X\,|\Psi\rangle = 2\pi\,R\,n\,|\Psi\rangle$$

for integer n. We can find P_X in terms of the gauge field by returning to the Lagrangian:

$$L = -\frac{2\pi R}{g_s}\sqrt{1-\dot{X}^2}$$

So that

$$P_X = \frac{dL}{d\dot{X}} = \frac{2\pi R}{g_s}\frac{\dot{X}}{\sqrt{1-\dot{X}^2}}$$

Since the electric field $\mathcal{E} = \dot{A}_x = \dot{X}$, we have the quantization condition

$$2\pi\alpha'\,\mathcal{E} = \frac{n\,g_s}{\sqrt{1+n^2\,g_s^2}},$$

where we have restored arbitrary units. The magnitude of the electric field is bounded by

$$|\mathcal{E}| \leq \frac{1}{2\pi\alpha'}$$

This is interesting. In string theory we cannot support arbitrarily large electric fields!

Find the tension of the D-string excited to the sector with n units of quantized electric field. Show that IIB S-duality suggests an interpretation in terms of a bound state of a D1-string and n F1-strings. In this case, what is the binding energy to the leading order in the string coupling?

To compute the mass, we calculate the Hamiltonian

$$M = H = P_X\,\dot{X} - L = \frac{2\pi R}{2\pi\alpha' g_s\sqrt{1-\dot{X}^2}} = \frac{2\pi R}{2\pi\alpha'\,g_s}\,\sqrt{1+g_s^2\,n^2}$$

We can write this as $M = 2\pi R\,T_{(1,n)}$ where

$$T_{(1,n)} = \sqrt{(n\,T_{\rm F1})^2 + (T_{\rm D1})^2}$$

Since the tensions of fundamental and D-strings are mapped into one another by the S-duality, the energy of our excited D-string is S-duality invariant. The form of the energy suggests that the result can be interpreted as a bound state of one D-string and n F1-strings.

The binding energy per unit length is given by the difference between the tension of the bound state and the sum of the tensions of separated components:

$$T_{\rm binding} = T_{(1,n)} - T_{\rm D1} - n\,T_{\rm F1} = -n\,T_{\rm F1}\,(1 - \tfrac{1}{2}\,n\,g_s) + O(g_s^3),$$

where we have kept the leading terms in the weak coupling expansion (by expanding the square root to leading order in the ratio $T_{\rm F1}/T_{\rm D1}$).

This result is interesting. It means that at very weak coupling almost *all* the rest mass of the fundamental strings is lost in the binding to the D1-string. This is similar to the interpretation of dyons as bound states of electrons and 't Hooft–Polyakov monopoles.
