1
A first look at strings

1.1 Why strings?

One of the main themes in the history of science has been unification. Time and again diverse phenomena have been understood in terms of a small number of underlying principles and building blocks. The principle that underlies our current understanding of nature is quantum field theory, quantum mechanics with the basic observables living at spacetime points. In the late 1940s it was shown that quantum field theory is the correct framework for the unification of quantum mechanics and electromagnetism. By the early 1970s it was understood that the weak and strong nuclear forces are also described by quantum field theory. The full theory, the $SU(3) \times SU(2) \times U(1)$ *Model* or *Standard Model,* has been confirmed repeatedly in the ensuing years. Combined with general relativity, this theory is consistent with virtually all physics down to the scales probed by particle accelerators, roughly 10^{-16} cm. It also passes a variety of indirect tests that probe to shorter distances, including precision tests of quantum electrodynamics, searches for rare meson decays, limits on neutrino masses, limits on axions (light weakly interacting particles) from astrophysics, searches for proton decay, and gravitational limits on the couplings of massless scalars. In each of these indirect tests new physics might well have appeared, but in no case has clear evidence for it yet been seen; at the time of writing, the strongest sign is the solar neutrino problem, suggesting nonzero neutrino masses.

The Standard Model (plus gravity) has a fairly simple structure. There are four interactions based on local invariance principles. One of these, gravitation, is mediated by the spin-2 graviton, while the other three are mediated by the spin-1 $SU(3) \times SU(2) \times U(1)$ gauge bosons. In addition, the theory includes the spin-0 Higgs boson needed for symmetry breaking, and the quarks and leptons, fifteen multiplets of spin-$\frac{1}{2}$ fermions in three generations of five. The dynamics is governed by a Lagrangian that depends upon roughly twenty free parameters such as the gauge and Yukawa couplings.

In spite of its impressive successes, this theory is surely not complete. First, it is too arbitrary: why does this particular pattern of gauge fields and multiplets exist, and what determines the parameters in the Lagrangian? Second, the union of gravity with quantum theory yields a nonrenormalizable quantum field theory, a strong signal that new physics should appear at very high energy. Third, even at the classical level the theory breaks down at the singularities of general relativity. Fourth, the theory is in a certain sense unnatural: some of the parameters in the Lagrangian are much smaller than one would expect them to be. It is these problems, rather than any positive experimental evidence, that presently must guide us in our attempts to find a more complete theory. One seeks a principle that unifies the fields of the Standard Model in a simpler structure, and resolves the divergence and naturalness problems.

Several promising ideas have been put forward. One is grand unification. This combines the three gauge interactions into one and the five multiplets of each generation into two or even one. It also successfully predicts one of the free parameters (the weak mixing angle) and possibly another (the bottom-tau mass ratio). A second idea is that spacetime has more than four dimensions, with the additional ones so highly curved as to be undetectable at current energies. This is certainly a logical possibility, since spacetime geometry is dynamical in general relativity. What makes it attractive is that a single higher-dimensional field can give rise to many four-dimensional fields, differing in their polarization (which can point along the small dimensions or the large) and in their dependence on the small dimensions. This opens the possibility of unifying the gauge interactions and gravity (the Kaluza–Klein mechanism). It also gives a natural mechanism for producing generations, repeated copies of the same fermion multiplets. A third unifying principle is supersymmetry, which relates fields of different spins and statistics, and which helps with the divergence and naturalness problems.

Each of these ideas — grand unification, extra dimensions, and supersymmetry — has attractive features and is consistent with the various tests of the Standard Model. It is plausible that these will be found as elements of a more complete theory of fundamental physics. It is clear, however, that something is still missing. Applying these ideas, either singly or together, has not led to theories that are substantially simpler or less arbitrary than the Standard Model.

Short-distance divergences have been an important issue many times in quantum field theory. For example, they were a key clue leading from the Fermi theory of the weak interaction to the Weinberg–Salam theory. Let

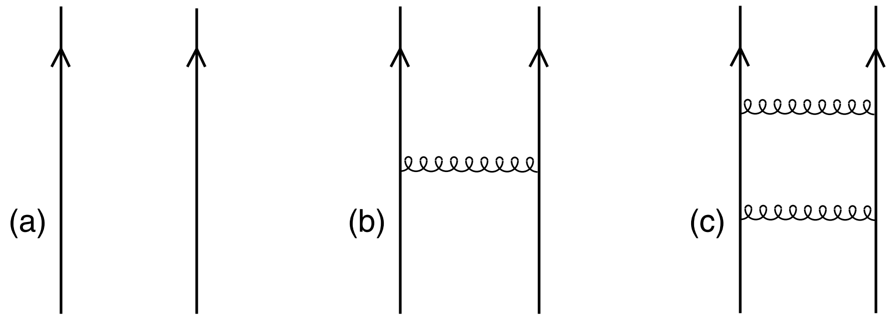

Fig. 1.1. (a) Two particles propagating freely. (b) Correction from one-graviton exchange. (c) Correction from two-graviton exchange.

us look at the short-distance problem of quantum gravity, which can be understood from a little dimensional analysis. Figure 1.1 shows a process, two particles propagating, and corrections due to one-graviton exchange and two-graviton exchange. The one-graviton exchange is proportional to Newton's constant G_{N}. The ratio of the one-graviton correction to the original amplitude must be governed by the dimensionless combination $G_{\text{N}} E^2 \hbar^{-1} c^{-5}$, where E is the characteristic energy of the process; this is the only dimensionless combination that can be formed from the parameters in the problem. Throughout this book we will use units in which $\hbar = c = 1$, defining the Planck mass

$$M_{\text{P}} = G_{\text{N}}^{-1/2} = 1.22 \times 10^{19}\,\text{GeV} \tag{1.1.1}$$

and the Planck length

$$M_{\text{P}}^{-1} = 1.6 \times 10^{-33}\,\text{cm}\ . \tag{1.1.2}$$

The ratio of the one-graviton to the zero-graviton amplitude is then of order $(E/M_{\text{P}})^2$.

From this dimensional analysis one learns immediately that the quantum gravitational correction is an *irrelevant* interaction, meaning that it grows weaker at low energy, and in particular is negligible at particle physics energies of hundreds of GeV. By the same token, the coupling grows stronger at high energy and at $E > M_{\text{P}}$ perturbation theory breaks down. In the two-graviton correction of figure 1.1(c) there is a sum over intermediate states. For intermediate states of high energy E', the ratio of the two-graviton to the zero-graviton amplitude is on dimensional grounds of order

$$G_{\text{N}}^2 E^2 \int dE'\, E' = \frac{E^2}{M_{\text{P}}^4} \int dE'\, E'\ , \tag{1.1.3}$$

which diverges if the theory is extrapolated to arbitrarily high energies. In position space this divergence comes from the limit where all the graviton vertices become coincident. The divergence grows worse with each additional graviton — this is the problem of nonrenormalizability.

There are two possible resolutions. The first is that the divergence is due to expanding in powers of the interaction and disappears when the theory is treated exactly. In the language of the renormalization group, this would be a nontrivial UV fixed point. The second is that the extrapolation of the theory to arbitrarily high energies is incorrect, and beyond some energy the theory is modified in a way that smears out the interaction in spacetime and softens the divergence. It is not known whether quantum gravity has a nontrivial UV fixed point, but there are a number of reasons for concentrating on the second possibility. One is history — the same kind of divergence problem in the Fermi theory of the weak interaction was a sign of new physics, the contact interaction between the fermions resolving at shorter distance into the exchange of a gauge boson. Another is that we need a more complete theory in any case to account for the patterns in the Standard Model, and it is reasonable to hope that the same new physics will solve the divergence problem of quantum gravity.

In quantum field theory it is not easy to smear out interactions in a way that preserves the consistency of the theory. We know that Lorentz invariance holds to very good approximation, and this means that if we spread the interaction in space we spread it in time as well, with consequent loss of causality or unitarity. Moreover we know that Lorentz invariance is actually embedded in a local symmetry, general coordinate invariance, and this makes it even harder to spread the interaction out without producing inconsistencies.

In fact, there is presently only one way known to spread out the gravitational interaction and cut off the divergence without spoiling the consistency of the theory. This is string theory, illustrated in figure 1.2. In this theory the graviton and all other elementary particles are one-dimensional objects, strings, rather than points as in quantum field theory. Why this should work and not anything else is not at all obvious *a priori,* but as we develop the theory we will see how it comes about.[1] Perhaps we merely suffer from a lack of imagination, and there are many other consistent theories of gravity with a short-distance cutoff. However, experience has shown that divergence problems in quantum field theory

[1] There is an intuitive answer to at least one common question: why not membranes, two- or higher-dimensional objects? The answer is that as we spread out particles in more dimensions we reduce the spacetime divergences, but encounter new divergences coming from the increased number of *internal* degrees of freedom. One dimension appears to be the unique case where both the spacetime and internal divergences are under control. However, as we will discuss in chapter 14, the membrane idea has resurfaced in somewhat transmuted form as *matrix theory*.

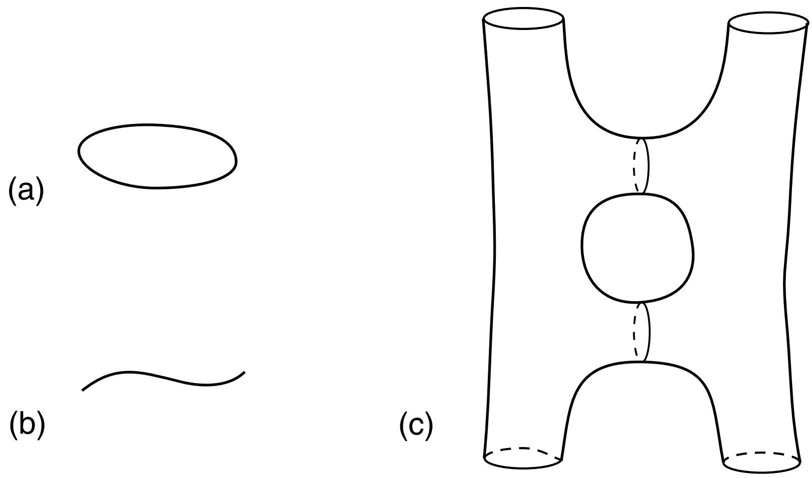

Fig. 1.2. (a) Closed string. (b) Open string. (c) The loop amplitude of fig. 1.1(c) in string theory. Each particle world-line becomes a cylinder, and the interactions no longer occur at points. (The cross-sections on the intermediate lines are included only for perspective.)

are not easily resolved, so if we have even one solution we should take it very seriously. Indeed, we are fortunate that consistency turns out to be such a restrictive principle, since the unification of gravity with the other interactions takes place at such high energy, $M_{\rm P}$, that experimental tests will be difficult and indirect.

So what do we find if we pursue this idea? In a word, the result is remarkable. String theory dovetails beautifully with the previous ideas for explaining the patterns in the Standard Model, and does so with a structure more elegant and unified than in quantum field theory. In particular, if one tries to construct a consistent relativistic quantum theory of one-dimensional objects one finds:

1. *Gravity*. Every consistent string theory must contain a massless spin-2 state, whose interactions reduce at low energy to general relativity.

2. *A consistent theory of quantum gravity, at least in perturbation theory*. As we have noted, this is in contrast to all known quantum *field* theories of gravity.

3. *Grand unification*. String theories lead to gauge groups large enough to include the Standard Model. Some of the simplest string theories lead to the same gauge groups and fermion representations that arise in the unification of the Standard Model.

4. *Extra dimensions.* String theory requires a definite number of spacetime dimensions, ten.[2] The field equations have solutions with four large flat and six small curved dimensions, with four-dimensional physics that resembles the Standard Model.

5. *Supersymmetry.* Consistent string theories require spacetime supersymmetry, as either a manifest or a spontaneously broken symmetry.

6. *Chiral gauge couplings.* The gauge interactions in nature are parity asymmetric (chiral). This has been a stumbling block for a number of previous unifying ideas: they required parity symmetric gauge couplings. String theory allows chiral gauge couplings.

7. *No free parameters.* String theory has no adjustable constants.

8. *Uniqueness.* Not only are there no continuous parameters, but there is no discrete freedom analogous to the choice of gauge group and representations in field theory: there is a unique string theory.

In addition one finds a number of other features, such as an axion, and hidden gauge groups, that have played a role in ideas for unification.

This is a remarkable list, springing from the simple supposition of one-dimensional objects. The first two points alone would be of great interest. The next four points come strikingly close to the picture one arrives at in trying to unify the Standard Model. And as indicated by the last two points, string theory accomplishes this with a structure that is tighter and less arbitrary than in quantum field theory, supplying the element missing in the previous ideas. The first point is a further example of this tightness: string theory *must* have a graviton, whereas in field theory this and other fields are combined in a mix-and-match fashion.

String theory further has connections to many areas of mathematics, and has led to the discovery of new and unexpected relations among them. It has rich connections to the recent discoveries in supersymmetric quantum field theory. String theory has also begun to address some of the deeper questions of quantum gravity, in particular the quantum mechanics of black holes.

Of course, much remains to be done. String theory may resemble the real world in its broad outlines, but a decisive test still seems to be far away. The main problem is that while there is a unique *theory,* it has an enormous number of classical solutions, even if we restrict attention

[2] To be precise, string theory modifies the notions of spacetime topology and geometry, so what we mean by a dimension here is generalized. Also, we will see that ten dimensions is the appropriate number for weakly coupled string theory, but that the picture can change at strong coupling.

to solutions with four large flat dimensions. Upon quantization, each of these is a possible ground state (vacuum) for the theory, and the four-dimensional physics is different in each. It is known that quantum effects greatly reduce the number of stable solutions, but a full understanding of the dynamics is not yet in hand.

It is worth recalling that even in the Standard Model, the dynamics of the vacuum plays an important role in the physics we see. In the electroweak interaction, the fact that the vacuum is less symmetric than the Hamiltonian (spontaneous symmetry breaking) plays a central role. In the strong interaction, large fluctuating gauge fields in the vacuum are responsible for quark confinement. These phenomena in quantum field theory arise from having a quantum system with many degrees of freedom. In string theory there are seemingly many more degrees of freedom, and so we should expect even richer dynamics.

Beyond this, there is the question, 'what is string theory?' Until recently our understanding of string theory was limited to perturbation theory, small numbers of strings interacting weakly. It was not known how even to define the theory at strong coupling. There has been a suspicion that the degrees of freedom that we use at weak coupling, one-dimensional objects, are not ultimately the simplest or most complete way to understand the theory.

In the past few years there has been a great deal of progress on these issues, growing largely out of the systematic application of the constraints imposed by supersymmetry. We certainly do not have a complete understanding of the dynamics of strongly coupled strings, but it has become possible to map out in detail the space of vacua (when there is enough unbroken supersymmetry) and this has led to many surprises. One is the absolute uniqueness of the theory: whereas there are several weakly coupled string theories, all turn out to be limits in the space of vacua of a single theory. Another is a limit in which spacetime becomes eleven-dimensional, an interesting number from the point of view of supergravity but impossible in weakly coupled string theory. It has also been understood that the theory contains new extended objects, D-branes, and this has led to the new understanding of black hole quantum mechanics. All this also gives new and unexpected clues as to the ultimate nature of the theory.

In summary, we are fortunate that so many approaches seem to converge on a single compelling idea. Whether one starts with the divergence problem of quantum gravity, with attempts to account for the patterns in the Standard Model, or with a search for new symmetries or mathematical structures that may be useful in constructing a unified theory, one is led to string theory.

Outline

The goal of these two volumes is to provide a complete introduction to string theory, starting at the beginning and proceeding through the compactification to four dimensions and to the latest developments in strongly coupled strings.

Volume one is an introduction to bosonic string theory. This is not a realistic theory — it does not have fermions, and as far as is known has no stable ground state. The philosophy here is the same as in starting a course on quantum field theory with a thorough study of scalar field theory. That is also not the theory one is ultimately interested in, but it provides a simple example for developing the unique dynamical and technical features of quantum field theory before introducing the complications of spin and gauge invariance. Similarly, a thorough study of bosonic string theory will give us a framework to which we can in short order add the additional complications of fermions and supersymmetry.

The rest of chapter 1 is introductory. We present first the action principle for the dynamics of string. We then carry out a quick and heuristic quantization using light-cone gauge, to show the reader some of the important aspects of the string spectrum. Chapters 2–7 are the basic introduction to bosonic string theory. Chapter 2 introduces the needed technical tools in the world-sheet quantum field theory, such as conformal invariance, the operator product expansion, and vertex operators. Chapters 3 and 4 carry out the covariant quantization of the string, starting from the Polyakov path integral. Chapters 5–7 treat interactions, presenting the general formalism and applying it to tree-level and one-loop amplitudes. Chapter 8 treats the simplest compactification of string theory, making some of the dimensions periodic. In addition to the phenomena that arise in compactified field theory, such as Kaluza–Klein gauge symmetry, there is also a great deal of 'stringy' physics, including enhanced gauge symmetries, T-duality, and D-branes. Chapter 9 treats higher order amplitudes. The first half outlines the argument that string theory in perturbation theory is finite and unitary as advertised; the second half presents brief treatments of a number of advanced topics, such as string field theory. Appendix A is an introduction to path integration, so that our use of quantum field theory is self-contained.

Volume two treats supersymmetric string theories, focusing first on ten-dimensional and other highly symmetric vacua, and then on realistic four-dimensional vacua.

In chapters 10–12 we extend the earlier introduction to the supersymmetric string theories, developing the type I, II, and heterotic superstrings and their interactions. We then introduce the latest results in these subjects. Chapter 13 develops the properties and dynamics of D-branes, still using the tools of string perturbation theory as developed earlier in the book. Chapter 14 then uses arguments based on supersymmetry to understand strongly coupled strings. We find that the strongly coupled limit of any string theory is described by a dual weakly coupled string theory, or by a new eleven-dimensional theory known provisionally as M-theory. We discuss the status of the search for a complete formulation of string theory and present one promising idea, m(atrix) theory. We briefly discuss the quantum mechanics of black holes, carrying out the simplest entropy calculation. Chapter 15 collects a number of advanced applications of the various world-sheet symmetry algebras.

Chapters 16 and 17 present four-dimensional string theories based on orbifold and Calabi–Yau compactifications. The goal is not an exhaustive treatment but rather to make contact between the simplest examples and the unification of the Standard Model. Chapter 18 collects results that hold in wide classes of string theories, using general arguments based on world-sheet and spacetime gauge symmetries. Chapter 19 consists of advanced topics, including (2,2) world-sheet supersymmetry, mirror symmetry, the conifold transition, and the strong-coupling behavior of some compactified theories.

Annotated reference lists appear at the end of each volume. I have tried to assemble a selection of articles, particularly reviews, that may be useful to the student. A glossary also appears at the end of each volume.

1.2 Action principles

We want to study the classical and quantum mechanics of a one-dimensional object, a string. The string moves in D flat spacetime dimensions, with metric $\eta_{\mu\nu} = \text{diag}(-, +, +, \cdots, +)$.

It is useful to review first the classical mechanics of a zero-dimensional object, a relativistic point particle. We can describe the motion of a particle by giving its position in terms of $D-1$ functions of time, $\mathbf{X}(X^0)$. This hides the covariance of the theory though, so it is better to introduce a parameter τ along the particle's world-line and describe the motion in spacetime by D functions $X^\mu(\tau)$. The parameterization is arbitrary: a different parameterization of the same path is physically equivalent, and all physical quantities must be independent of this choice. That is, for any monotonic function $\tau'(\tau)$, the two paths X'^μ and X^μ are the same, where

$$X'^\mu(\tau'(\tau)) = X^\mu(\tau) \ . \tag{1.2.1}$$

We are trading a less symmetric description for a more symmetric but redundant one, which is often a useful step. Figure 1.3(a) shows a parameterized world-line.

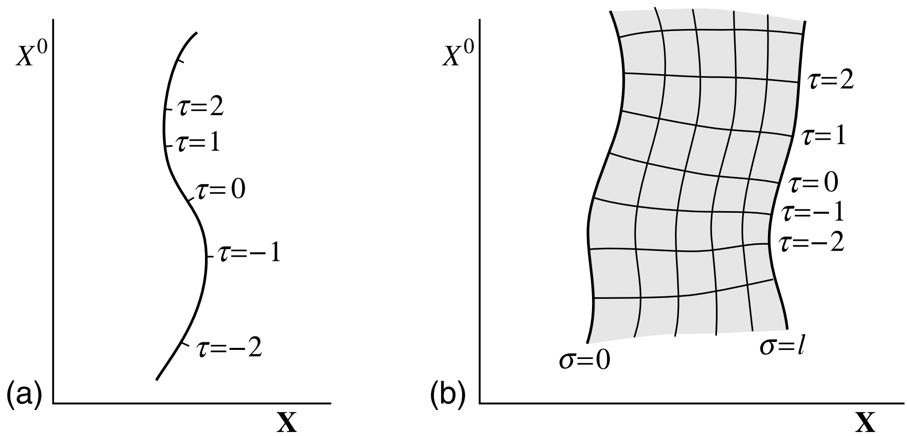

Fig. 1.3. (a) Parameterized world-line of a point particle. (b) Parameterized world-sheet of an open string.

The simplest Poincaré-invariant action that does not depend on the parameterization would be proportional to the proper time along the world-line,

$$S_{\mathrm{pp}} = -m \int d\tau \, (-\dot{X}^\mu \dot{X}_\mu)^{1/2} \, , \tag{1.2.2}$$

where a dot denotes a τ-derivative. The variation of the action, after an integration by parts, is

$$\delta S_{\mathrm{pp}} = -m \int d\tau \, \dot{u}_\mu \delta X^\mu \, , \tag{1.2.3}$$

where

$$u^\mu = \dot{X}^\mu (-\dot{X}^\nu \dot{X}_\nu)^{-1/2} \tag{1.2.4}$$

is the normalized D-velocity. The equation of motion $\dot{u}^\mu = 0$ thus describes free motion. The normalization constant m is the particle's mass, as can be checked by looking at the nonrelativistic limit (exercise 1.1).

The action can be put in another useful form by introducing an additional field on the world-line, an independent world-line metric $\gamma_{\tau\tau}(\tau)$. It will be convenient to work with the tetrad $\eta(\tau) = (-\gamma_{\tau\tau}(\tau))^{1/2}$, which is defined to be positive. We use the general relativity term *tetrad* in any number of dimensions, even though its root means 'four.' Then

$$S'_{\mathrm{pp}} = \frac{1}{2} \int d\tau \left(\eta^{-1} \dot{X}^\mu \dot{X}_\mu - \eta m^2 \right) . \tag{1.2.5}$$

This action has the same symmetries as the earlier action S_{pp}, namely Poincaré invariance and world-line reparameterization invariance. Under