Michio Kaku

Introduction to Superstrings and M-Theory

Second Edition

Springer

Michio Kaku

Introduction to Superstrings and M-Theory

Second Edition

With 45 Illustrations

Springer

Michio Kaku
Department of Physics
City College of the
City University of New York
New York, NY 10031, USA

Library of Congress Cataloging-in-Publication Data
Kaku, Michio.
Introduction to superstrings and M-theory / Michio Kaku. —2nd ed.
p. cm.—(Graduate texts in contemporary physics)
Includes bibliographical references and index.
ISBN 0-387-98589-1 (alk. paper)
1. Superstring theories. I. Title. II. Series.
QC794.6.S85K35 1998
539.7′258—dc21 98-26976

Printed on acid-free paper.

Production managed by A. Orrantia; manufacturing supervised by Joe Quatela.
Photocomposed copy prepared in LaTeX by The Bartlett Press, Inc., Marietta, GA.
Printed and bound by R.R. Donnelly and Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3

ISBN 0-387-98589-1 SPIN 10796467

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

Contents

Part I

First Quantization and Path Integrals

CHAPTER 1

Path Integrals and Point Particles

1.1 Why Strings?

One of the greatest scientific challenges of our time is the struggle to unite the two fundamental theories of modern physics, quantum field theory and general relativity, into one theoretical framework. Remarkably, these two theories together embody the sum total of all human knowledge concerning the most fundamental forces of Nature. Quantum field theory, for example, has had phenomenal success in explaining the physics of the microcosm, down to distances less than 10^{-15} cm. General relativity, on the other hand, is unrivaled in explaining the large-scale behavior of the cosmos, providing a fascinating and compelling description of the origin of the Universe itself. The astonishing success of these two theories is that together they can explain the behavior of matter and energy over a staggering 40 orders of magnitude, from the subnuclear to the cosmic domain.

The great mystery of the past five decades, however, has been the total incompatibility of these two theories. It's as if Nature had two minds, each working independently of the other in its own particular domain, operating in total isolation of the other. Why should Nature, at its deepest and most fundamental level, require two totally distinct frameworks, with two sets of mathematics, two sets of assumptions, and two sets of physical principles?

Ideally, we would want a unified theory to unite these two fundamental theories:

$$\left.\begin{array}{l}\text{Quantum field theory}\\ \text{General relativity}\end{array}\right\} \text{Unified field theory.}$$

However, the history of attempts over the past decades to unite these two theories has been dismal. They have inevitably been riddled with infinities or

have violated some of the cherished principles of physics, such as causality. The powerful techniques of renormalization theory developed in quantum field theory over the past decades have failed to eliminate the infinities of quantum gravity. Apparently, a fundamental piece of the jigsaw puzzle is still missing.

Although quantum field theory and general relativity seem totally incompatible, the past two decades of intense theoretical research have made it increasingly clear that the secret to this mystery most likely lies in the power of *gauge symmetry*. One of the most remarkable features of Nature is that its basic laws have great unity and symmetry when expressed in terms of group theory. Unification through gauge symmetry, apparently, is one of the great lessons of physics. In particular, the use of local symmetries in Yang–Mills theories has had enormous success in banishing the infinities of quantum field theory and in unifying the laws of elementary particle physics into an elegant and comprehensive framework. Nature, it seems, does not simply incorporate symmetry into physical laws for aesthetic reasons. Nature *demands* symmetry.

The problem has been, however, that even the powerful gauge symmetries of Yang–Mills theory and the general covariance of Einstein's equations are insufficient to yield a finite quantum theory of gravity.

At present, the most promising hope for a truly unified and finite description of these two fundamental theories is superstring theory and its latest formulation, M-theory. [1–12]. Superstrings possess by far the largest set of gauge symmetries ever found in physics, perhaps even large enough to eliminate all divergences of quantum gravity. Not only does the superstring's symmetry include that of Einstein's theory of general relativity and the Yang–Mills theory, it also includes supergravity and the Grand Unified Theories (GUTs) [13] as subsets.

Roughly speaking the way in which superstring theory solves the riddle of infinities can be visualized as in Fig. 1.1, where we calculate the scattering of two point particles by summing over an infinite set of Feynman diagrams with loops. These diagrams, in general, have similarities that correspond to "pinching" one of the internal lines until the topology of the graph is altered.

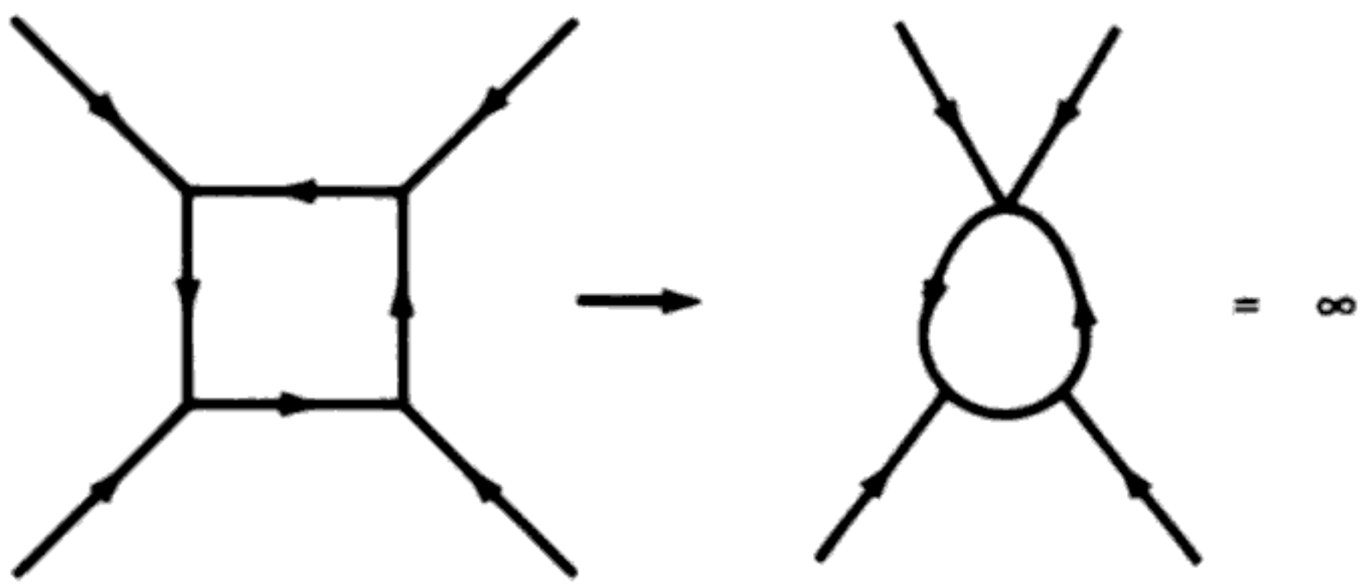

FIGURE 1.1. Single-loop Feynman diagram for four-particle scattering. The ultraviolet divergence of this diagram corresponds to the pinching of one internal leg, i.e., when one internal line shrinks to a point.

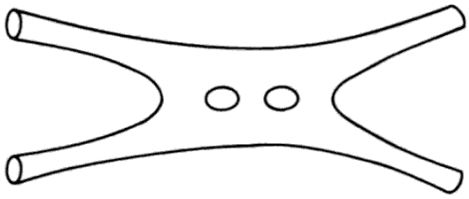

FIGURE 1.2. Two-loop Feynman diagram for closed string scattering. The diagram is ultraviolet finite because it cannot be pinched as in the point particle case. From topological arguments alone, we can see that string theory is less divergent than point particle theory. Infrared divergences, however, may still exist.

By contrast, in Fig. 1.2 we have the single-loop contribution to the scattering of two closed string states. Notice that we cannot "pinch" one of the internal lines as in the point particle case. Thus, we naively expect that the superstring theory is less divergent or even finite because of the symmetries that forbid this topological deformation.

Any theory that can simultaneously eliminate the infinities of the S-matrix and incorporate quantum mechanics, the general theory of relativity, GUT theory, and supergravity obviously possesses mathematics of breathtaking beauty and complexity. In fact, even the mathematicians have been startled at the mathematics emerging from the superstring theory, which links together some of the most dissimilar, far-ranging fields of mathematics, such as Kac–Moody algebras, Riemann surfaces and Teichmüller spaces, modular groups, and even Monster group theory.

The great irony of string theory, however, is that the theory itself is not unified. To someone learning the theory for the first time, it is often a frustrating collection of folklore, rules of thumb, and intuition. At times, there seems to be no rhyme or reason for many of the conventions of the model. For a theory that makes the claim of providing a unifying framework for all physical laws, it is the supreme irony that the theory itself appears so disunited! The secrets of the model, at its most fundamental level, are still being pried loose.

Usually, when we write down a quantum theory, we start with the geometry or symmetry of the theory and then write down the action. From the action, in turn, we derive all the predictions of the model, including the unitary S-matrix. Thus, a *second quantized* action is the proper way in which to formulate a quantum field theory. The fundamental reason why superstring theory seems, at times, to be a loose collection of apparently random conventions is that it is usually formulated as a first quantized theory. Because of this, we must appeal to intuition and folklore in order to construct all the Feynman diagrams for a unitary theory. We hope that M-theory (or perhaps evern a more advanced theory) will be able to unify superstring theory into a simple, coherent formalism. Already, M-theory can unify the five different superstring theories into a single theory. Ultimately, there may be a single equation (perhaps no more than an inch long) which will unify the entire theory.

Unfortunately, the geometry of the superstring and membranes are some of the last features of the model to be developed. In fact, as seen from this perspective, the model has been developing *backward* for the past 30 years, beginning with the accidental discovery of its quantum theory in 1968!

By contrast, when Einstein first discovered general relativity, he started with physical principles, such as the equivalence principle, and formulated it in the language of general covariance. Once the geometry was established, he then wrote down the action as the unique solution to the problem. Later, classical solutions to the equations were discovered in terms of curved manifolds, which provided the first successful theoretical models for the large-scale behavior of the Universe. Finally, the last step in the evolution of general relativity is the development of a quantum theory of gravity. The crucial steps in the historical evolution of general relativity can thus be represented as

$$\text{Geometry} \rightarrow \text{Action} \rightarrow \text{Classical theory} \rightarrow \text{Quantum theory}.$$

Furthermore, both general relativity and Yang–Mills theory are mature theories: they both can be formulated from first principles, which stresses the geometry and the physical assumptions underlying the theory. Superstring theory and M-theory are just beginning to reach that stage of development.

Remarkably, Yang–Mills theory and gravity theory are the *unique* solution to two simple geometric statements:

(1) Global Symmetry
The free theory must propagate pure ghost-free spin-1 and spin-2 fields transforming as irreducible representations of SU(N) and the Lorentz group.

(2) Local Symmetry
The theory must be locally SU(N) and generally covariant.

What is remarkable is that the coupled Yang–Mills gravity action is the unique solution of these two simple principles

$$L = -\tfrac{1}{4}\sqrt{-g}\,F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\kappa^2}\sqrt{-g}\,R_{\mu\nu}g^{\mu\nu}. \qquad (1.1.1)$$

(The first principle contains the real physics of the theory. It cannot be included as a subset of the second principle. There is an infinite number of generally covariant and SU(N) symmetric invariants, so we need the first principle to input the physics and select the irreducible representations of the basic fields. By "pure" fields, we mean ghost-free fields that have at most two derivatives, which rules out R^2 and F^4 higher derivatives theories.)

The question remains: *What is the counterpart to these two simple principles for superstring theory and M-theory*?

The plan of this book, of course, must reflect the fact that the theory has been evolving backward. For pedagogical reasons, we will mostly follow the historical development of the theory. Thus, Part I of the book, which introduces the first quantized theory, will at times appear to be a loose collection of

However, the Yukawa theory could not explain the avalanche of "elementary" particles that were being discovered in particle accelerators. J. Robert Oppenheimer even suggested that the Nobel Prize in Physics should go to the physicist who *didn't* discover a particle that year. Furthermore, the quark model, which seemed to fit data much better than it had any right to, was plagued with the fact that quarks were never seen experimentally.

(3) *The gravitational force.* Gravity research was totally uncoupled from research in the other interactions. Classical relativists continued to find more and more classical solutions in isolation from particle research. Attempts to canonically quantize the theory were frustrated by the presence of the tremendous redundancy of the theory. There was also the discouraging realization that even if the theory could be successfully quantized, it would still be nonrenormalizable.

This bleak landscape changed dramatically in the early 1970s with the coming of the gauge revolution. One of the great achievements of the past 25 years has been the development of a fully renormalizable theory of spin-1 gauge particles in which, for the first time, physicists could actually calculate realistic S-matrix elements. Thus, it took over 100 years to advance beyond the original gauge theory first proposed by Maxwell in the 1860s! (See the Appendix for an elementary introduction to gauge theories and group theory.)

Apparently the key to eliminating the divergences of relativistic quantum mechanics is to go to larger and more sophisticated gauge groups. Symmetry, instead of being a purely aesthetic feature of a particular model, now becomes its most important feature.

For example, Maxwell's equations, which provided the first unification of the electric force with the magnetic force, has a gauge group given by U(1). The unification of the weak and electromagnetic forces into the electroweak force requires $SU(2) \otimes U(1)$. The forces that bind the quarks together into the hadrons, or quantum chromodynamics (QCD), are based on SU(3). All of elementary particle physics, in fact, is compatible with the minimal theory of $SU(3) \otimes SU(2) \otimes U(1)$.

Although the verdict is still not in on the GUTs, which are supposed to unite the electroweak force with the strong force, once again the unifying theme is gauge symmetry, with such proposals as SU(5), O(10), etc., symmetry.

Although the gauge revolution is perhaps one of the most important developments in decades, it is still not enough. There is a growing realization that the Yang–Mills theory by itself cannot push our understanding of the physical universe beyond the present level. Not only do the GUTs fail to explain important physical phenomena, but also there is the crucially important problem of formulating a quantum theory of gravity.

Grand Unified Theories, first of all, cannot be the final word on the unification of all forces. There are several features of GUTs that are still unresolved:

If we set

$$\frac{h}{2\pi} = 1, \qquad c = 1, \tag{1.2.3}$$

there still remains a dimensional constant even in the Newtonian theory of gravity, the gravitational constant G:

$$F = G, \frac{m_1 m_2}{r^2}, \tag{1.2.4}$$

which has dimensions of centimeters squared. When we power expand the metric tensor $g_{\mu\nu}$ around a flat square with the metric $\eta_{\mu\nu} = (- + + +)$, we introduce the coupling constant κ, which has dimensions of centimeters:

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}. \tag{1.2.5}$$

Therefore

$$G \sim \kappa^2. \tag{1.2.6}$$

In this system of units, where the only unit is the centimeter, this coupling constant κ becomes the Planck length, 10^{-33} cm or 10^{19} GeV, which is far beyond the reach of experimentation!

Renormalization theory, however, is founded on the fundamental premise that we can eliminate all divergences with an infinite redefinition of certain constants. Having a negative dimensional coupling constant means that this complicated reshuffling and resuming of graphs is impossible. Negative coupling constants mean that we can always insert the interaction term into a Feynman diagram and increase its power of divergence. This means that any graph can be made arbitrarily divergent by multiple insertions. *This means that general relativity cannot be a renormalizable theory*. The amplitude for graviton–graviton scattering, for example, is now a power expansion in a dimensional parameter (see Fig. 1.3):

$$A = \sum_{n=2}^{\infty} \kappa^n A_n, \tag{1.2.7}$$

where we are no longer able to shuffle graphs in the usual manner to cancel the infinities, which is the heart of renormalization theory. Thus, renormalization theory breaks down.

Because general relativity is hopelessly outside the domain of conventional renormalization theory, we must reconsider Dirac's fundamental objection. It was Dirac who said that the success of quantum mechanics was based on approximation schemes where each correction term was increasingly small. But renormalization theory is flawed because it maximally violates this principle and manipulates infinite quantities and discards them at the end.

One solution might be to construct a theory of gravity that is finite to every order in the coupling constant, with no need for renormalization at all. For a while, one bright hope was supergravity [14, 15], based on the local gauge

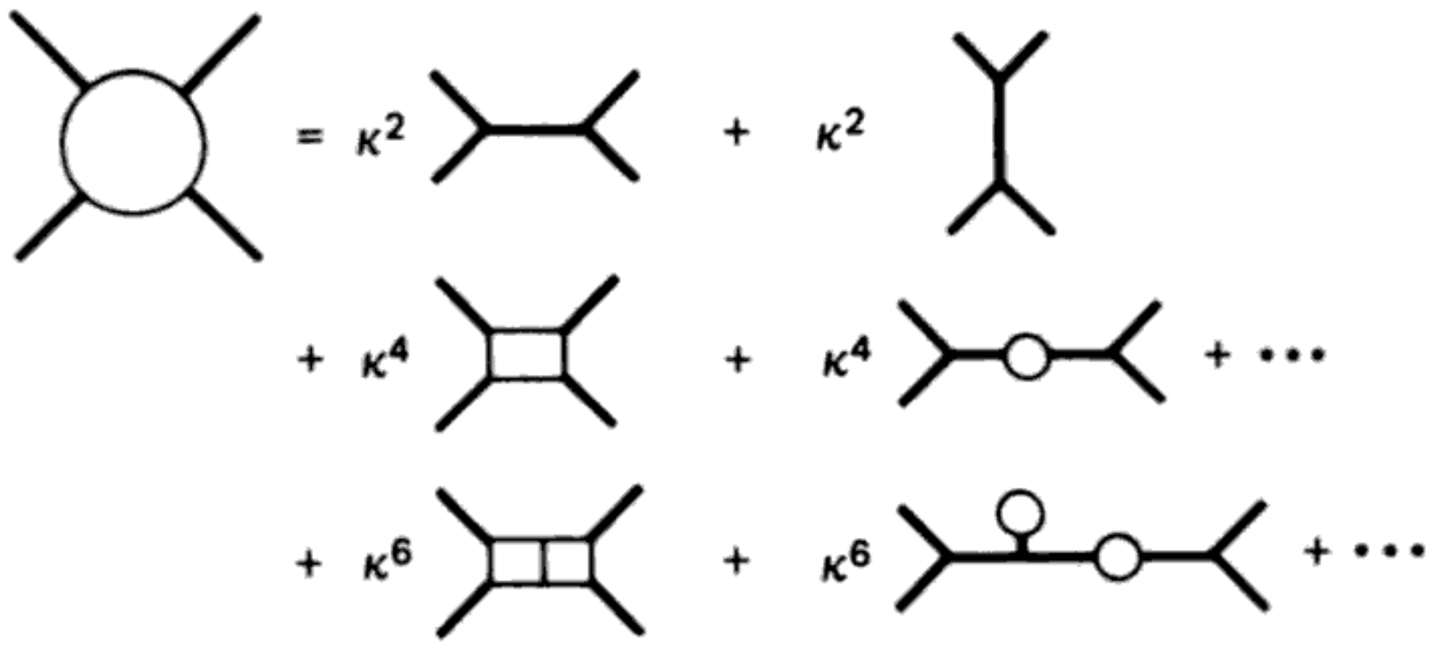

FIGURE 1.3. Scattering amplitude for graviton–graviton scattering. Because the coupling constant has negative dimension any graph can be made arbitrarily divergent, thereby requiring an infinite number of counterterms. Thus, theories containing quantum gravity must be either divergent or completely finite order-by-order. Pure quantum gravity has been shown on computer to diverge at the two-loop level. Counterterms have also been found for quantum gravity coupled to lower-spin particles. Thus, superstring theory is the only candidate for a finite theory.

group Osp(N/4) (see Appendix), which was the first nontrivial extension of Einstein's equations in 60 years. The hope was that this gauge group would offer us a large enough set of Ward–Takahashi identities to cancel a large class of divergent diagrams. The larger the gauge group, the more likely troublesome infinities would cancel (see Fig. 1.4):

Theory	Gauge group
Electromagnetism	U(1)
Electroweak	SU(2) ⊗ U(1)
Strong	SU(3)
GUT(?)	SU(5), O(10)
Gravity(?)	GL(4), O(3, 1)
Supergravity(?)	Osp(N/4)

The basic strategy being pursued was

Gauge symmetry → Ward–Takahashi identities
→ Cancellation of graphs → Renormalizable theory.

For example, even Einstein's theory of gravity can be shown to be trivially finite at the first loop level. There exists a remarkable identity, called the Gauss–Bonnet identity, which immediately shows that all one-loop graphs in general relativity (which would take a computer to write down) sum to zero. In fact, the super-Gauss–Bonnet identities eliminate many of the divergences of supergravity, but probably not enough to make the theory finite.

The largest and most promising of the supergravities, the O(8) supergravity, is probably divergent. Unfortunately, it is possible to write down locally supersymmetric counterterms at the seventh loop level. It is highly unlikely that the coefficients of this and probably an infinite number of other counterterms can all vanish without appealing to an even higher symmetry. This is discourag-

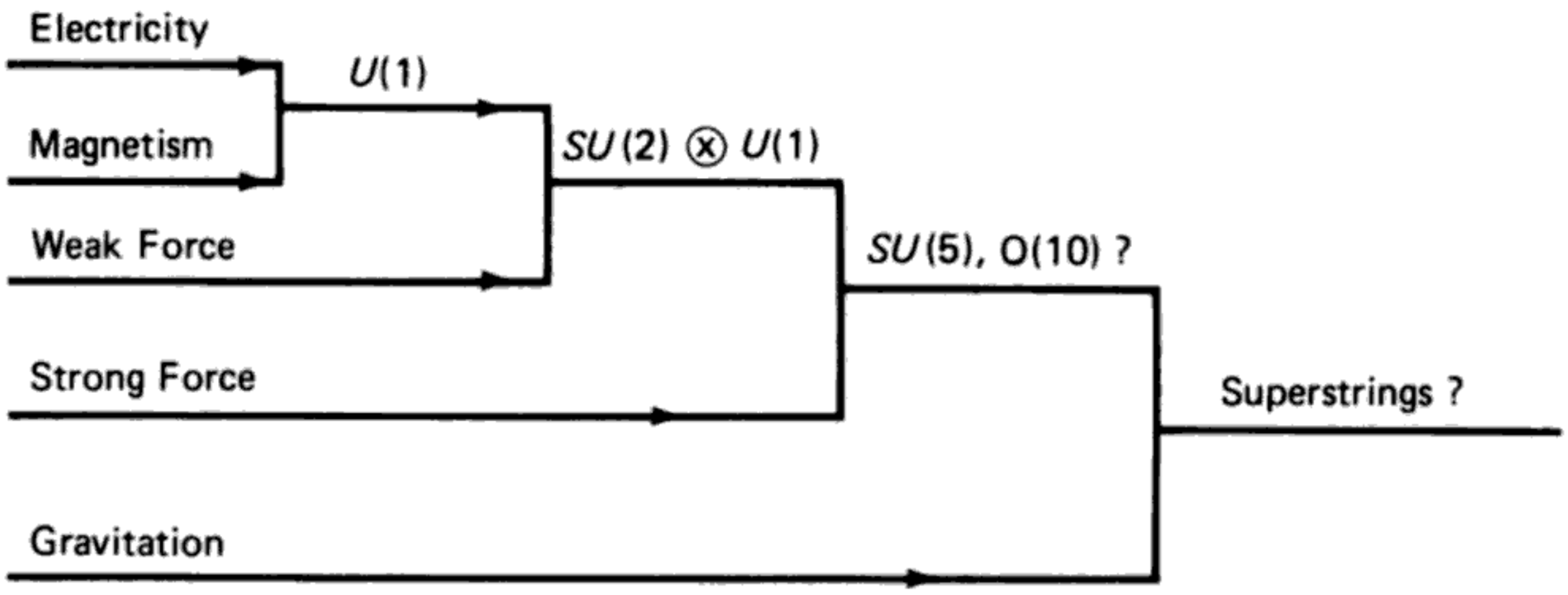

FIGURE 1.4. Chart showing how gauge theories based on Lie groups have united the fundamental forces of Nature. Maxwell's theory, based on U(1), unites electricity and magnetism. The Weinberg–Salam model, based on SU(2) ⊗ U(1), unites the weak force with the electromagnetic force. GUTs (based on SU(5), O(10), or larger groups) are the best candidate to unite the strong force with the electroweak force. Superstring theory is the only candidate for a gauge theory that can unite gravity with the rest of the particle forces.

ing, because it means that the gauge group of the largest supergravity theory, Osp(8/4), is still too small to eliminate the divergences of general relativity.

Furthermore, the O(8) gauge group is too small to accommodate the minimal SU(3) ⊗ SU(2) ⊗ U(1) of particle physics. If we go to higher groups beyond O(8), we find that we must incorporate higher and higher spins into the theory. However, an interacting spin-3 theory is probably not consistent, making one suspect that O(8) is the limit to supergravity theories.

In conclusion, supergravity must be ruled out for two fundamental reasons:

(1) It is probably not a finite theory because the gauge group is not large enough to eliminate all possible supersymmetric counterterms. There is a possible counterterm at the seventh loop level.
(2) Its gauge group O(8) is not large enough to accommodate the minimal symmetry of particle physics, namely, SU(3) ⊗ SU(2) ⊗ U(1); nor can the theory accommodate chiral fermions.

Physicists, faced with these and other stumbling blocks over the years, have concluded that perhaps one or more of our cherished assumptions about our Universe must be abandoned. Because general relativity and quantum mechanics can be derived from a small set of postulates, one or more of these postulates must be wrong. The key must be to drop one of our commonsense assumptions about Nature on which we have constructed general relativity and quantum mechanics. Over the years, several proposals have been made to drop some of our commonsense notions about the Universe:

(1) Continuity
This approach assumes that space–time must be granular. The size of these grains would provide a natural cutoff for the Feynman integrals, allowing

us to have a finite S-matrix. Integrals like

$$\int_{\varepsilon}^{\infty} d^4x \tag{1.2.8}$$

would then diverge as ε^{-n}, but we would never take the limit as ε goes to zero. Lattice gravity theories are of this type. In Regge calculus [16], for example, we latticize Riemannian space with discrete four-simplexes and replace the curvature tensor by the angular deficit calculated when moving in a circle around a simplex:

$$-\frac{1}{2\kappa^2}\sqrt{-g}R \rightarrow \text{angular deficit.}$$

(In flat space, there is no angular deficit when walking around a closed path, and the action collapses.) Usually, in lattice theories, we take the limit as the lattice length goes to zero. Here, however, we keep it fixed at a small number [17]. At present, however, there is no experimental evidence to support the idea that space–time is granular. Although we can never rule out this approach, it seems to run counter to the natural progression of particle physics, which has been to postulate larger and more elegant groups.

(2) Causality

This approach allows small violations in causality. Theories that incorporate the Lee–Wick mechanism [18] are actually renormalizable, but permit small deviations from causality. These theories make the Feynman diagrams converge by adding a fictitious Pauli–Villars field of mass M that changes the ultraviolet behavior of the propagator. Usually, the Feynman propagator converges as p^{-2} in the ultraviolet limit. However, by adding a fictitious particle, we can make the propagator converge even faster, like p^{-4}:

$$\frac{1}{p^2+m^2} - \frac{1}{p^2+M^2} \rightarrow \frac{1}{p^4}. \tag{1.2.9}$$

Notice that the Pauli–Villars field is a ghost because of the -1 that appears in the propagator. (This means that the theory will be riddled with negative probabilities.) Usually, we let the mass of the Pauli–Villars field tend to infinity. However, here we keep it finite, letting the pole go out onto the unphysical sheet. Investigations of the structure of the resulting Feynman diagrams show, however, that causality is violated; that is, you can meet your parents before you are born.

(3) Unitarity

We can replace Einstein's theory, which is based on the curvature tensor, with a conformal theory based on the Weyl tensor:

$$\sqrt{-g}R_{\mu\nu}g^{\mu\nu} \rightarrow \sqrt{-g}C^2_{\mu\nu\rho\sigma}, \tag{1.2.10}$$

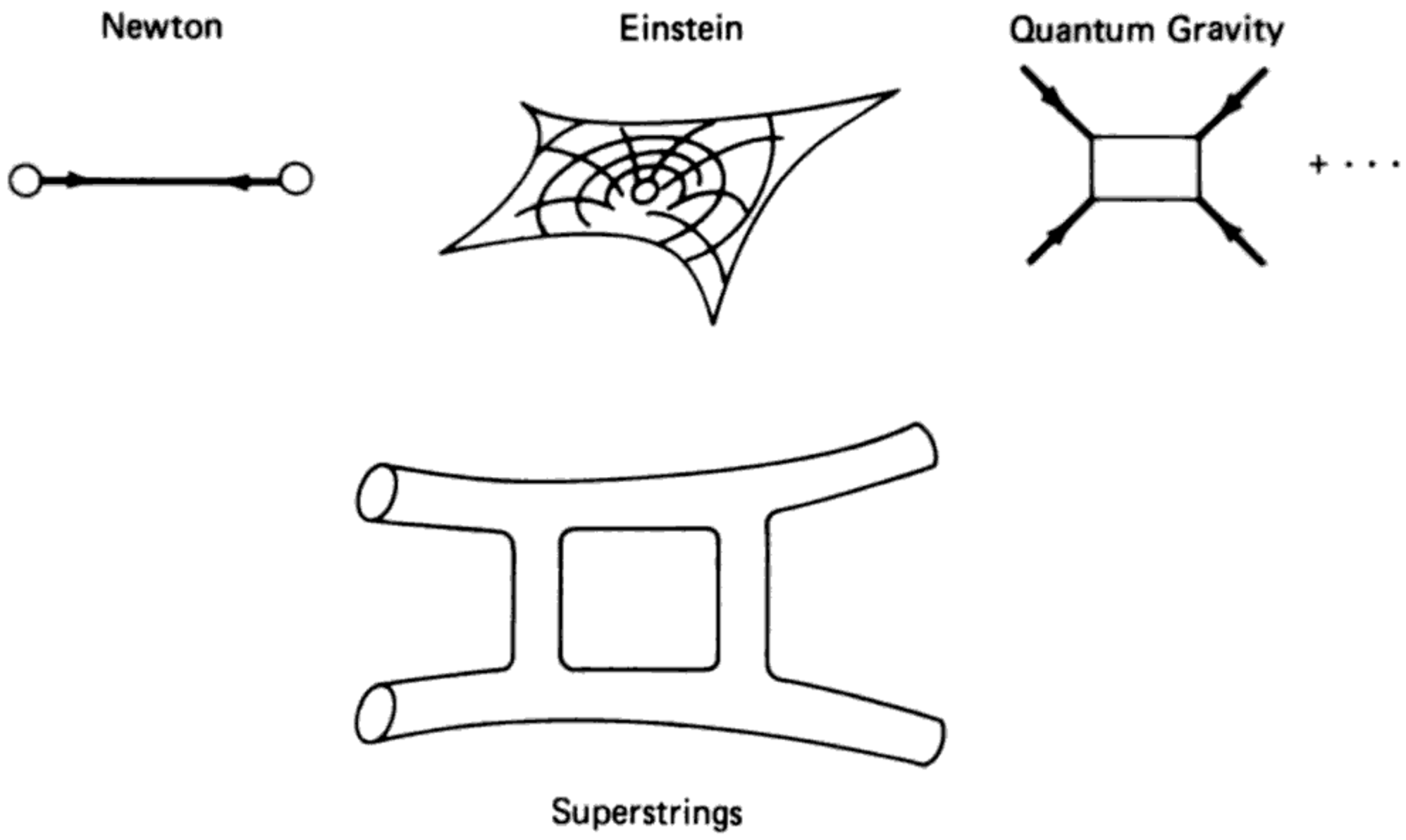

FIGURE 1.5. Steps in the evolution of the theory of gravitation. Each step in this chart builds on the successes of the previous step. Newton thought gravity was a force that acted instantly over a distance. Einstein proposed that gravitation was caused by the curvature of space–time. The naive merger of general relativity and quantum mechanics produces a divergent theory, quantum gravity, which assumes that gravitation is caused by the exchange of particle-like gravitons. Superstring theory proposes that gravitation is caused by the exchange of closed strings.

Superstring theory, however, is quite unlike its predecessors in its historical development. Unlike other physical theories, superstring theory has perhaps one of the strangest histories in science, with more twists and turns than a roller coaster.

First, two young physicists Veneziano and Suzuki [21, 22], independently discovered its quantum theory when they were thumbing through a mathematics book and accidentally noted that the Euler Beta function satisfied all the postulates of the S-matrix for hadronic interactions (except unitarity). Neveu, Schwarz, and Ramond [23–25] quickly generalized the theory to include spinning particles. To solve the problem of unitarity, Kikkawa, Sakita, and Virasoro [26] proposed that the Euler Beta function be treated as the Born term to a perturbation series. Finally, Kaku, Yu, Lovelace, and Alessandrini [27–33] completed the quantum theory by calculating bosonic multiloop diagrams. The theory, however, was still formulated entirely in terms of on-shell S-matrix amplitudes.

Next, Nambu and Goto [34, 35] realized that lurking behind these scattering amplitudes was a classical relativistic string. In one sweep, they revolutionized the entire theory by revealing the unifying, classical picture behind the theory. The relationship between the classical theory and the quantum theory was quickly made by Goldstone, Goddard, Rebbi, and Thorn [36] and further developed by Mandelstam [37]. The theory, however, was still formulated as a

self-consistent. The symmetries of the superstring theory, by a series of "miracles," can cancel all of its potential anomalies.

(3) Powerful arguments from the theory of Riemann surfaces indicate that the theory is finite to all orders in perturbation theory (although a rigorous proof is still lacking).

(4) There is very little freedom to play with. Superstring models are notoriously difficult to tinker with without destroying their miraculous properties. Thus, we do not have the problem of 19 arbitrary coupling constants.

(5) The theory includes GUTs, super-Yang–Mills, supergravity, and Kaluza–Klein theories as subsets. Thus, many of the features of the phenomenology developed for these theories carry over into the string theory.

Superstring theory, crudely speaking, unites the various forces and particles in the same way that a violin string provides a unifying description of the musical tones. By themselves, the notes A, B, C, etc., are not fundamental. However, the violin string is fundamental; one physical object can explain the varieties of musical notes and even the harmonies we can construct from them. In much the same way, the superstring provides a unifying description of elementary particles and forces. In fact, the "music" created by the superstring is the forces and particles of Nature.

Although superstring theory, because of its fabulously large set of symmetries, has "miraculous" cancellations of anomalies and divergences, we must also present a balanced picture and point out its shortcomings. To be fair we must also list the potential problems of the theory that have been pointed out by critics of the model:

(1) It is impossible experimentally to reach the tremendous energies found at the Planck scale. Therefore, the theory is in some sense untestable. A theory that is untestable is not an acceptable physical theory.

(2) Not one shred of experimental evidence has been found to confirm the existence of supersymmetry, let alone superstrings.

(3) It is presumptuous to assume that there will be no surprises in the "desert" between 100 and 10^{19} GeV. New, totally unexpected phenomena have always cropped up when we have pushed the energy scale of our accelerators. Superstring theory, however, makes predictions over the next 17 orders of magnitude, which is unheard of in the history of science.

(4) The theory does not explain why the cosmological constant is zero. Any theory that claims to be a "theory of everything" must surely explain the puzzle of a vanishing cosmological constant, but it is not clear how superstrings solve this problem.

(5) The theory has an embarrassment of riches. There are apparently *millions* of ways to break down the theory to low energies. Which is the correct vacuum? Although the superstring theory can produce the minimal theory of $SU(3) \otimes SU(2) \otimes U(1)$, it also predicts many other interactions that have not yet been seen.

(6) No one really knows how to break a 10-dimensional theory down to four dimensions.

Of these six objections to the model, the most fundamental is the last, the inability to calculate dimensional breaking. The reason for this is simple: to every order in perturbation theory, the dimension of space–time is stable. Thus, in order to have the theory spontaneously curl up into four- and six-dimensional universes, we must appeal to nonperturbative, dynamical effects, which are notoriously difficult to calculate. This is why the search for the geometry underlying the theory is so important. The geometric formulation of the model may give us the key insight into the model that will allow us to make nonperturbative calculations and make definite predictions with the theory.

Thus, the criticism that the model cannot be tested at the Planck length is actually slightly deceptive. *The superstring theory, if it could be successfully broken dynamically, should be able to make predictions down to the level of everyday energies.* For example, it should be able to predict the masses of the quarks. Therefore, we do not have to wait for several centuries until we have accelerators that can reach the Planck length.

Thus, the fundamental problem facing superstrings is not necessarily an experimental one. It is mainly theoretical. The outstanding problem of the theory is to calculate dynamical symmetry breaking, so that its predictions can be compared with experimental data at ordinary energies.

A fundamental theory at Planck energies is also a fundamental theory at ordinary energies. Thus, the main stumbling block to the development of the theory is an understanding of its nonperturbative behavior.

In Part I of this book, however, we will follow historical precedent and present the first quantized formulation of the model. As we will stress throughout this book, the first quantized theory seems to be a loose collection of random facts. As a consequence, we have emphasized the path integral formulation (first written down for the Veneziano model by Hsue, Sakita, and Virasoro [42, 43]) as the most powerful method of formulating the first quantized theory. Although the path integral approach cannot reveal the underlying geometric formulation of the model, it provides the most comprehensive formulation of the first quantized theory.

We will not turn to the functional formulation [44] of point particle theory, which can be incorporated almost directly into the string theory.

1.3 Path Integrals and Point Particles

Let us begin our discussion by analyzing the simplest of all possible systems, the classical nonrelativistic point particle. Surprisingly, much of the analysis of this simple dynamical system carries over directly to the superstring theory. The language we will use is the formalism of path integrals, which is so versatile that

it can accommodate both first quantized point particles and second quantized gauge fields with equal ease.

As in classical mechanics, the starting point is the Lagrangian for a point particle:

$$L = \tfrac{1}{2}m\dot{x}_i^2 - V(x), \tag{1.3.1}$$

where the particle is moving in an external potential. The real physics is contained in the statement that the action S must be minimized. The equations of motion can be derived by minimizing the action:

$$S = \int L(x_i, \dot{x}_i, t)\, dt,$$
$$\delta S = 0. \tag{1.3.2}$$

To calculate the equations of motion, let us make a small variation in the path of the particle given by

$$\delta x_i, \qquad \delta \dot{x}_i. \tag{1.3.3}$$

Under this small variation, the action varies as follows:

$$\int dt \left\{ \frac{\delta L}{\delta x_i}\delta x_i + \frac{\delta L}{\delta \dot{x}_i}\delta \dot{x}_i \right\} = 0. \tag{1.3.4}$$

Integrating by parts, we arrive at the Euler–Lagrange equations:

$$\frac{\delta L}{\delta x_i} - \frac{d}{dt}\frac{\delta L}{\delta \dot{x}_i} = 0. \tag{1.3.5}$$

For our point particle, the equations of motion become

$$m\frac{d^2x_i}{dt^2} = -\frac{\partial V(x)}{\partial x_i} \tag{1.3.6}$$

which correspond to the usual classical Newtonian equations of motion.

In addition to the Lagrangian formulation of classical mechanics, there is also the Hamiltonian form. Instead of introducing the position and the velocities as fundamental objects, we now introduce the position and the momentum:

$$p_i = \frac{\delta L}{\delta \dot{x}_i}. \tag{1.3.7}$$

With this definition of the conjugate variable, we have

$$H = p_i\dot{x}_i - L,$$
$$H(p_i, x_i) = \frac{p_i^2}{2m} + V(x). \tag{1.3.8}$$

Finally, the Poisson brackets between the momenta and the coordinates are given by

$$[p_i, x_j]_{\mathrm{PB}} = -\delta_{ij}. \tag{1.3.9}$$

A celebrated theorem in classical mechanics states that the equations of motion of Newton and the action principle method can be shown to be identical. Beginning with the action principle, we can derive Newton's laws of motion, and vice versa

$$\text{Equations of motion} \leftrightarrow \text{Action principle.}$$

This equivalence, however, breaks down at the quantum level. Quantum mechanically, there is a fundamental difference between the two, with the equations of motion being only an approximation to the actual quantum behavior of matter. *Thus, the action principle is the only acceptable framework for quantum mechanics.*

Let us now reformulate the principles of quantum mechanics in terms of Feynman path integrals [44]:

(1) The probability $P(a, b)$ of a particle moving from point a to point b is the square of the absolute value of a complex number, the transition function $K(a, b)$:

$$P(a, b) = |K(a, b)|^2. \tag{1.3.10}$$

(2) The transition function is given by the sum of a certain phase factor, which is a function of the action S, taken over all possible paths from a to b:

$$K(a, b) = \sum_{\text{paths}} k e^{i2\pi S/h}, \tag{1.3.11}$$

where the constant k can be fixed by

$$K(a, c) = \sum_{\text{paths}} K(a, b)K(b, c), \tag{1.3.12}$$

and the intermediate sum is taken over paths that go through all possible intermediate points b.

The second principle says that a particle "sniffs out" all possible paths from point a to point b, no matter how complicated the paths may be. We calculate this phase factor for each of this infinite number of paths. Then the transition factor for the path between a and b is calculated by summing over all possible phase factors (see Fig. 1.6).

Remarkably, *the essence of quantum mechanics is captured in these two principles*. All the profoundly important implications of quantum mechanics, which represent a startling departure from classical mechanics, can be derived from these two innocent-sounding principles! In particular, these two principles summarize the essence of the quantum interpretation of the double-slit experiment, which, in turn, summarizes the essence of quantum mechanics itself.

It is apparent at this point that the results of classical mechanics can be reproduced from our two assumptions in a certain approximation. Notice that,

and

$$\sum_{\text{paths}} \to \int Dx = \lim_{N\to\infty} \int \prod_{i=1}^{3} \prod_{n=1}^{N} dx_{i,n} \,, \tag{1.3.17}$$

where the index n labels N intermediate points that divide the interval between the initial and the final coordinate. We will take the limit when N approaches infinity.

It is absolutely essential to understand that the integration Dx is not the ordinary integration over x. In fact, it is the product of all possible integrations over all intermediate points $x_{i,n}$ between points a and b. This crucial difference between ordinary integration and functional integration goes to the heart of the path integral formalism.

This infinite series of integrations, in turn, is equivalent to summing over all possible paths between a and b. Thus, we will have to be careful to include normalization factors when performing an integration over an infinite number of intermediate points.

If we take the simple case where $L = \frac{1}{2}m\dot{x}_i^2$, all functional integrations can actually be performed exactly. The integral in question is a Gaussian, which is fortunately one of the small number of functional integrals that can actually be performed. One of the great embarrassments of the method of path integrals is that one of the few integrals that can actually be performed is

$$\int_{-\infty}^{\infty} dx x^{2n} e^{-r^2x^2} = \frac{\Gamma(n+\frac{1}{2})}{r^{2n+1}}. \tag{1.3.18}$$

We will be using this formula throughout the entire book.

Let us now break up the path into an infinite number of intermediate points, $x_{i,n}$. (Notice that the functional expression integrates over all possible values of the intermediate point $x_{i,n}$, so we cannot expect that $x_{i,n}$ and $x_{i,n+1}$ are close to each other even for small time separations.) Let us write

$$\begin{aligned} dt &\to \varepsilon \\ \tfrac{1}{2}m\dot{x}_1^2\, dt &\to \tfrac{1}{2}m(x_n - x_{n+1})_i^2 \varepsilon^{-1}. \end{aligned} \tag{1.3.19}$$

In order to perform the functional integral over an infinite number of intermediate points, we will repeatedly use the following Gaussian integration:

$$\begin{aligned} &\int_{-\infty}^{\infty} dx_2 \exp[-a(x_1 - x_2)^2 - a(x_2 - x_3)^2] \\ &\qquad = \sqrt{\frac{\pi}{2a}} \exp\left[-\tfrac{1}{2}a(x_1 - x_3)^2\right]. \end{aligned} \tag{1.3.20}$$

One of the crucial points to observe here is that the integration over a Gaussian in one of the intermediate points yields another Gaussian with that intermediate point removed. This is the fundamental reason why we can perform the functional integration over an infinite number of intermediate points.

With these eigenstates, we can now rewrite the expression for the Green's function for going from point x_1 to x_N.

$$K(1, N) = \langle x_1, t_1 | x_N, t_N \rangle. \tag{1.3.27}$$

In order to derive the previous expression (1.3.22) for transition amplitude, let us insert a complete set of intermediate states at every intermediate point between x_1 and x_N:

$$\begin{aligned}\langle x_1, t_1 | x_N, t_N \rangle = \langle x_1, t_1 | x_2, t_2 \rangle \int dx_2 \, \langle x_2, t_2 | \int dx_2 \\ \cdots | x_{N-1}, t_{N-1} \rangle \int dx_{N-1} \langle x_{N-1}, t_{N-1} | x_N, t_N \rangle.\end{aligned} \tag{1.3.28}$$

Now let us examine each infinitesimal propagator in terms of the Hamiltonian, which we write as a function of the coordinates and derivatives:

$$H = H(x, \partial_x). \tag{1.3.29}$$

Then the transition for an infinitesimal interval is given by

$$\begin{aligned}\langle x_1, t_1 | x_2, t_2 \rangle &= \langle x_1 | \, e^{-iH(x,\partial_x)\delta t} \, | x_2 \rangle \\ &= e^{-iH(x,\partial_{\dot{x}})\delta t} \langle x_1 | x_2 \rangle \\ &= e^{-iH(x,\partial_{\dot{x}})\delta t} \langle x_1 | p \rangle \int dp \, \langle p | x_2 \rangle \\ &= e^{-iH(x,p)\delta t} \int \frac{dp}{2\pi} e^{ip(x_2 - x_1)} \\ &= e^{-iH(x,p)\delta t} \int \frac{dp}{2\pi} e^{ip\dot{x}\delta t}.\end{aligned} \tag{1.3.30}$$

It is very important to notice that path integrals have made it possible to make the transition from classical to quantum commutators. The Hamiltonian can be expressed either as a function of derivatives with respect to the position or as a function of the canonical momenta because of the identity:

$$\partial_x e^{ipx} = ipe^{ipx}. \tag{1.3.31}$$

This allows us to make the important identification:

$$\begin{cases} H(x, p) \leftrightarrow H(x, \partial_x), \\ p \leftrightarrow -i \dfrac{\delta}{\delta x}. \end{cases} \tag{1.3.32}$$

In the functional formalism, the important correspondence between momenta and partial derivatives arises because of this identity.

Putting everything together, we can now write the complete transition amplitude as

$$\langle x_1, t_1 | x_N, t_N \rangle = \int_{x_1}^{x_N} DpDx \exp\left\{ i \int_{t_1}^{t_N} [p\dot{x} - H(p, x)] \, dt \right\}, \tag{1.3.33}$$

where

$$H = \frac{p_i^2}{2m} + V(x). \tag{1.3.34}$$

(As usual, we have dropped all the intermediate normalizations, which are just factors of 2π.) Notice that the functional integral, which was once only a function of the coordinates, is now a function of both the momenta and the coordinates.

In order to retrieve the original Lagrangian, we can perform the p integration exactly, because it is a simple Gaussian integral, and we arrive at

$$\langle x_1, t_1 | x_N, t_N \rangle = \int_{x_1}^{x_N} Dx \exp\left\{ i \int_{t_1}^{t_N} \left[\tfrac{1}{2} m \dot{x}_i^2 - V(x) \right] dt \right\}. \tag{1.3.35}$$

We have thus made the transition between the Lagrangian and the Hamiltonian formalism using functional methods. We can use either:

$$L = \tfrac{1}{2} m \dot{x}_i^2 - V(x) \quad \leftrightarrow \quad H = \frac{p_i^2}{2m} + V(x). \tag{1.3.36}$$

Functionally, the only difference between these two expressions is whether we integrate over the coordinates or a combination of the coordinates and the momenta. The transition probability can be represented as

$$\begin{aligned} K(a,b) &= \int_{x_a}^{x_b} Dx \exp\left\{ i \int_{t_a}^{t_b} dt \left[\tfrac{1}{2} m x_i^2 - V(x) \right] \right\} \\ &= \int_{x_a}^{x_b} Dx\, Dp \exp\left\{ i \int_{t_a}^{t_b} dt \left[p \dot{x}_i - \frac{p_i^2}{2m} - V(m) \right] \right\}. \end{aligned} \tag{1.3.37}$$

1.4 Relativistic Point Particles

So far, our discussion has been limited to nonrelativistic particles, where all degrees of freedom are physical. However, nontrivial complications occur when we generalize our previous discussion to the case of relativistic particles. In particular, the (-1) appearing in the Lorentz metric will, in general, cause nonphysical states to propagate in the theory. These nonphysical "ghost" states, which have negative probability, must be eliminated carefully to ensure a sensible causal theory free of negative norm states.

For the relativistic case, let us assume that the location of a point particle is given by a four-vector:

$$x_\mu(\tau), \tag{1.4.1}$$

where parametrization τ does *not* necessarily refer to the time. The action is particulary simple, being proportional to the four-dimensional path length:

$$S = -m \int ds = -m \text{ (length)}. \tag{1.4.2}$$

The path length ds can be written in terms of the coordinates:

$$ds = \sqrt{-\dot{x}_\mu^2}\, d\tau, \tag{1.4.3}$$

where the dot refers to differentiation with respect to the parameter τ. This action, unlike the previous nonrelativistic action, is invariant under reparametrizations of the fictitious parameter τ. Let us make a change of coordinates from τ to $\tilde{\tau}$:

$$\tau \to \tilde{\tau}(\tau). \tag{1.4.4}$$

Then we find

$$\begin{aligned} d\tau &= \frac{d\tau}{d\tilde{\tau}} d\tilde{\tau}, \\ \frac{dx}{d\tau} &= \frac{dx}{d\tilde{\tau}} \frac{d\tilde{\tau}}{d\tau}, \\ \left\{ \left(\frac{dx_\mu}{d\tau} \right)^2 \right\}^{1/2} d\tau &= \left\{ \left(\frac{dx_\mu}{d\tilde{\tau}} \right)^2 \right\}^{1/2} d\tilde{\tau}. \end{aligned} \tag{1.4.5}$$

Thus, the action is invariant under an arbitrary reparametrization of the variable τ.

This can be written infinitesimally as

$$\begin{cases} \tau \to \tau + \delta\tau, \\ \delta x_\mu = \dot{x}_\mu \delta\tau. \end{cases} \tag{1.4.6}$$

As before, we can now introduce canonical conjugates:

$$p_\mu = \frac{\delta L}{\delta \dot{x}^\mu} = \frac{m\dot{x}_\mu}{\sqrt{-\dot{x}_\mu^2}}. \tag{1.4.7}$$

The crucial difference, however, from our previous discussion of the nonrelativistic point particle is that not all the canonical momenta are independent. In fact, we find a constraint among them:

$$p_\mu^2 + m^2 \equiv 0. \tag{1.4.8}$$

Thus, the mass shell condition arises as an exact constraint among the momenta. If we calculate the Hamiltonian associated with this system, we find that

$$H = p^\mu \dot{x}_\mu - L \equiv 0. \tag{1.4.9}$$

The Hamiltonian vanishes identically.

These unusual features, the vanishing of the Hamiltonian and the constraints among the momenta, are typical of systems with redundant gauge degrees of freedom. The invariance under reparametrization, for example, tells us that the path integral that we wrote earlier actually diverges:

$$\int Dx e^{iS} = \infty. \tag{1.4.10}$$

This is because there is a separate contribution from each particular parametrization. But since Dx is parametrization invariant, this means that we are summing over an infinite number of copies of the same thing. Thus, the integral must diverge.

Dirac, however, explained how to quantize systems with redundant gauge degrees of freedom. For example, let us introduce canonical momenta p and impose the constraint condition via a Lagrange multiplier as follows:

$$L = p_\mu \dot{x}^\mu - \tfrac{1}{2} e(p_\mu^2 + m^2). \tag{1.4.11}$$

The constraint equation (1.4.8) is imposed here as a classical equation of motion. By varying e, we recover the constraint on the momenta. Quantum mechanically, however, this constraint is imposed by functionally integrating out over e. In the path integral, we have

$$\int De \exp\left[-i \int d\tau \tfrac{1}{2} e(p^2 + m^2)\right] \sim \delta(p^2 + m^2), \tag{1.4.12}$$

where we have used the fact that the integral over e^{ikx} (or the Fourier transform of the number 1) is equal to $\delta(x)$. Notice that the new Lagrangian (1.4.11) still possesses the gauge degree of freedom. It is invariant under

$$\begin{aligned} \delta x_\mu &= \varepsilon \dot{x}_\mu, \\ \delta p_\mu &= \varepsilon \dot{p}_\mu, \\ \delta e &= \frac{d(\varepsilon e)}{d\tau}. \end{aligned} \tag{1.4.13}$$

The advantage that this action has over the previous one is that all variables occur linearly. We do not have to worry about complications caused by the square root. (The field e that we have introduced will become the metric tensor g_{ab} when we generalize this action to the string.)

Let us now functionally integrate over the p variable. Because the integration is again a Gaussian, we have no problem in performing the p integration:

$$\begin{aligned} &\int Dp \exp\left\{i \int d\tau \left[p\dot{x} - \tfrac{1}{2} e(p^2 + m^2)\right]\right\} \\ &\qquad \sim \exp\left\{i \int d\tau \tfrac{1}{2}(e^{-1}\dot{x}^2 - em^2)\right\}. \end{aligned} \tag{1.4.14}$$

Thus, we have now obtained a third version of the point particle action. The advantage of this action is that it is linear in the coordinates and is invariant under

$$\begin{cases} \delta x_\mu = \varepsilon \dot{x}_\mu, \\ \delta e = \dfrac{d(\varepsilon e)}{d\tau}. \end{cases} \tag{1.4.15}$$

In summary, we have found three equivalent ways to express the relativistic point particle. The "second-order" Lagrangian (1.4.14) is expressed in terms of

action reduces to

$$L = -m \int \sqrt{1 - v_i^2} \quad dt. \tag{1.5.2}$$

In the limit of velocities small compared to the velocity of light, we have

$$L \sim \tfrac{1}{2} m \dot{x}_i^2 \tag{1.5.3}$$

as before, so that the functional integral is modified to

$$\int Dx_\mu \delta(x_0 - t) e^{iS} = \int Dx_i \exp\left(i \int \tfrac{1}{2} m \dot{x}_i^2 \, dt \right). \tag{1.5.4}$$

For the case of the string, this simple example will lay the basis for the light cone quantization. The advantage of the Coulomb gauge is that all ghosts have been explicitly removed from the theory, so we are dealing only with physical quantities. The other advantage is that the zeroth component of the position vector is now explicitly defined to be the time variable. The parametrization of the point particle is now given in terms of the physical time.

The disadvantage of the Coulomb formalism, however, is that manifest Lorentz symmetry is broken and we have to check explicitly that the quantized Lorentz generators close correctly. Although this is trivial for the point particle, surprising features will emerge for the quantum string, fixing the dimension of space–time to be 26.

Gupta–Bleuler Quantization

This approach tries to maintain Lorentz invariance. This means, of course, that particular care must be taken to prevent the negative norm states from spoiling the physical properties of the S-matrix. The Gupta–Bleuler method keeps the action totally relativistic, but imposes the constraint (1.4.8) on state vectors:

$$[p_\mu^2 + m^2] \, |\phi\rangle = 0. \tag{1.5.5}$$

(Notice that the above equation is a ghost-killing constraint, because we can use it to eliminate p_0.) This formalism allows us to keep the commutators fully relativistic:

$$[p_\mu, x_\nu] = -i \eta_{\mu\nu}, \tag{1.5.6}$$

where we choose $\eta_{\mu\nu} = (- + + + \cdots)$. Notice that this gauge constraint naturally generalizes to the Klein–Gordon equation:

$$[\Box - m^2] \phi(x) = 0. \tag{1.5.7}$$

The Gupta–Bleuler formalism is an important one because most of the calculations in string theory have been carried out in this formalism.

determinant is gauge independent of any particular parametrization:

$$\Delta_{\mathrm{FP}}(x) = \Delta_{\mathrm{FP}}(x^{\varepsilon}). \tag{1.6.4}$$

Let us now insert the number 1 into the functional integral and make a gauge transformation to reabsorb the ε dependence in x:

$$Z = \int Dx\, \Delta_{\mathrm{FP}}(x) \int D\varepsilon \delta[F(x^{\varepsilon})] e^{iS}$$

$$= \int Dx\, \Delta_{\mathrm{FP}}(x) \int D\varepsilon \delta[F(x)] e^{iS}. \tag{1.6.5}$$

Notice here the x^{ε} was gauge rotated back into the original variable x. Since all other parts of the functional integral were already gauge independent, we now have

$$Z = \left[\int D\varepsilon\right] \int Dx\, \Delta_{\mathrm{FP}} \delta[F(x)] e^{iS}. \tag{1.6.6}$$

We can now extract out the integral over the gauge parameter, which measures the infinite volume of the group space:

$$\text{volume} = \int D\varepsilon \tag{1.6.7}$$

and obtain a new expression for the functional which no longer has this infinite redundancy:

$$Z = \int Dx\, \Delta_{\mathrm{FP}} \delta[F(x)] e^{iS}. \tag{1.6.8}$$

Notice that a naive quantization of the path integral would simply insert the F constraint and would omit the Faddeev–Popov determinant, which is a new feature that makes the measure come out correctly.

Now let us calculate the Faddeev–Popov determinant, which carries all the information concerning the ghosts of the theory. The trick is to change variables from ε to F. We can do this because both ε and F have the same number of degrees of freedom. Thus, the Jacobian can be calculated:

$$\det\left[\frac{\delta F}{\delta \varepsilon}\right] D\varepsilon = DF. \tag{1.6.9}$$

We can therefore write

$$\Delta_{\mathrm{FP}} = \left\{\int D\varepsilon \delta(F)\right\}^{-1} = \left\{\int DF \det\left[\frac{\delta \varepsilon}{\delta F}\right] \delta(F)\right\}^{-1}$$

$$= \left\{\det\left[\frac{\delta \varepsilon}{\delta F}\right]_{F=0}\right\}^{-1} = \det\left[\frac{\delta F}{\delta \varepsilon}\right]_{F=0}. \tag{1.6.10}$$

Thus, the Faddeev–Popov factor can be expressed as a simple determinant of the variation of the gauge constraint. It is more convenient to introduce this

factor directly into the action by exponentiating it. We use the following trick:

$$\Delta_{\mathrm{FP}} = \int D\theta D\bar{\theta} e^{iS_{\mathrm{gh}}}, \tag{1.6.11}$$

where the new ghost contribution to the action is given by

$$S_{\mathrm{gh}} = \int d\tau \bar{\theta} \left[\frac{\delta F}{\delta \varepsilon}\right]_{F=0} \theta, \tag{1.6.12}$$

where the θ variables are anticommuting c-numbers called *Grassmann numbers*. Normally, when performing functional integrations, we expect to find the determinant of the inverse of a matrix. With functional integration over Grassmann numbers, the determinant occurs in the *numerator*, not the denominator. Grassmann numbers have the strange property that

$$\theta_i \theta_j = -\theta_j \theta_i. \tag{1.6.13}$$

In particular, this means

$$\theta^2 = 0. \tag{1.6.14}$$

Normally, this would mean that θ vanishes. However, this is not the case for a Grassmann number. Thus, we also have the strange identity

$$e^{\theta} = 1 + \theta. \tag{1.6.15}$$

This identity makes the integration over exponentials of Grassmann-valued fields in the functional integral rather easy, because they are simply polynomials. More identities on Grassmann numbers are presented in the Appendix, where we show that

$$\int \prod_{i=1}^{N} d\theta_i \, d\bar{\theta}_i \exp\left[\sum_{i,j=1}^{N} \bar{\theta}_i A_{ij} \theta_j\right] = \det(A_{ij}). \tag{1.6.16}$$

This identity verifies that integration over Grassmann variables yields determinant factors in the numerator, not the denominator, so that we can express the Faddeev–Popov determinant in (1.6.11) as a Grassmann integral.

Now that we have developed the apparatus of Faddeev–Popov quantization, let us return to the BRST approach, where we wish to impose the gauge condition

$$e = 1 \tag{1.6.17}$$

(we omit some subtleties with respect to this gauge). In this gauge, we should be able to recover the usual covariant Feynmann propagator. To show this, notice that our action (1.4.14) becomes

$$L = \tfrac{1}{2}(\dot{x}_{\mu}^2 - m^2). \tag{1.6.18}$$

Given this Lagrangian, our Green's function for the propagation of a point particle from one point to another is now given by

$$\Delta_F(x_1, x_2) = \langle x_1 | \frac{1}{\Box - m^2} | x_2 \rangle = \langle x_1 | \int_0^\infty d\tau e^{-\tau(\Box - m^2)} | x_2 \rangle$$
$$= \int_0^\infty d\tau \int_{x_1}^{x_2} Dx \exp\left(-\frac{1}{2} \int_0^\tau d\bar{\tau} (\dot{x}_\mu^2 - m^2) \right). \quad (1.6.19)$$

Notice that this is the usual covariant Feynman propagator rewritten in first quantized path integral language.

Originally, before gauge fixing, our action was invariant under

$$\delta e = \frac{d(\varepsilon e)}{d\tau}. \quad (1.6.20)$$

Thus, the Faddeev–Popov determinant associated with the gauge choice $e = 1$ is the determinant of the derivative. We now use a Gaussian integral over Grassmann states to represent the determinant, using (1.6.10):

$$\Delta_{\mathrm{FP}} = \det |\partial_\tau| = \int D\theta D\bar{\theta} \exp\left(i \int d\tau \bar{\theta} \partial_\tau \theta \right). \quad (1.6.21)$$

(If we had used ordinary real fields instead of Grassmann-valued fields, the determinant would have come out with the wrong power.)

Putting everything together, we find that our final action can be represented as

$$L = p_\mu \dot{x}^\mu - \tfrac{1}{2}(p_\mu^2 + m^2) - i\bar{\theta}\partial_\tau \theta. \quad (1.6.22)$$

The essence of the BRST approach is to notice that this gauge-fixed action has the additional symmetry:

$$\begin{aligned} \delta x_\mu &= i\varepsilon\theta \dot{x}_\mu, \\ \delta p_\mu &= i\varepsilon\theta \dot{p}_\mu, \\ \delta\theta &= i\varepsilon\theta\dot{\theta}, \\ \delta\bar{\theta} &= i\varepsilon\theta\dot{\bar{\theta}} + \tfrac{1}{2}\varepsilon(p_\mu^2 + m^2). \end{aligned} \quad (1.6.23)$$

At first, we may wonder why yet another symmetry appears after we have already fixed the gauge degree of freedom. However, this extra symmetry is *global* and hence does not allow us to impose any constraints on the theory. This symmetry, therefore, is different from the ones found earlier and cannot be used to eliminate gauge fields from the action.

We can summarize the BRST approach by extracting an operator Q that will generate the symmetry found earlier:

$$\begin{aligned} \delta\phi &= [\varepsilon Q, \phi], \\ Q &= \theta(\Box - m^2), \\ Q^2 &= 0. \end{aligned} \quad (1.6.24)$$

The physical states satisfy

$$Q\,|\phi\rangle = 0. \tag{1.6.25}$$

Notice that enforcing this constraint recovers the Klein–Gordon equation for on-shell particles:

$$(\Box - m^2)\phi = 0. \tag{1.6.26}$$

1.7 Second Quantization

So far, we have been analyzing only the first quantized approach to quantum particles. We have quantized only the position and momentum vectors:

$$\text{first quantization:} \quad [p_i, x_j] = -i\delta_{ij}. \tag{1.7.1}$$

The limitations of the first quantized approach, however, will soon become apparent when we introduce interactions. Let us say that we wish to describe point particles that can bump into each other and split apart, rather than introduce an external potential. We must now modify the generating functional to include summing over Feynman graphs:

$$Z = \sum_{\text{topologies}} \int Dx e^{-\text{length}}. \tag{1.7.2}$$

(Notice that we have Wick rotated the τ integration so that the exponential converges. It will be clear from the context when the Wick-rotated theory is being used in this book because the exponential becomes real. We will not discuss the delicate question of the convergence of path integrals.)

In other words, we must, by hand, sum over the various particle topologies where point particles can split and reform. Each topology represents the history of the trajectories of the various point particles as they interact. The amplitude for N-particle scattering, with momenta given by $k_1, k_2, \ldots, k_N$, can now be represented as

$$A(k_1, k_2, \ldots, k_N) = \sum_{\text{topologies}} g^n \int Dx\, \Delta_{\text{FP}} \times \exp\left\{ -\int dt L(t) + i \sum_{i=1}^{N} k_\mu x_i^\mu \right\}. \tag{1.7.3}$$

Notice that we are taking the Fourier transform of the Green's function, so that the amplitude is a function of the external momenta. This formula can be more conveniently represented as

$$A_N = \sum_{\text{topologies}} g^n \left\langle \exp\left(i \sum_{i=1}^{N} k_\mu x_i^\mu \right) \right\rangle. \tag{1.7.4}$$

The Green's function satisfies

$$(i\partial_t - H)K(x,t;x',t') = \delta^{(3)}(x-x')\delta(t-t'). \tag{1.7.8}$$

Solving for this Green's function, we find

$$K(a,b) = [i\partial_t - H]^{-1}_{x_a,t_a;x_b,t_b}, \tag{1.7.9}$$

where we are treating the inverse Green's function as if it were a discrete matrix in (x,t) space, and we have dropped trivial normalization factors. This allows us to write the integral in second quantized language. To demonstrate this, we will use the following identities throughout this book:

$$\int \prod_{i=1}^{N} dx_i \exp\left\{\sum_{i,j=1}^{N} -x_i A_{ij} x_j + \sum_{i=1}^{N} J_i x_i\right\}$$
$$= \frac{\pi^{(1/2)N}}{\det|A_{ij}|} \exp\left\{\frac{1}{4}\sum_{i,j=1}^{N} J_i (A^{-1})_{ij} J_j\right\}. \tag{1.7.10}$$

(This integral can easily be derived using our earlier formula for the Gaussian integral (1.3.18). We simply diagonalize the A matrix by making a change of variables in x. Thus, the quadratic term in the integral becomes a function of the eigenvalues of the A matrix. Because all the modes have now decoupled, the Gaussian integral can be performed exactly by completing the square. Finally, we make another similarity transformation to convert the eigenvalues of A back into the A matrix itself.)

From this, we can also derive the following:

$$\int x_n x_m \prod_{i=1}^{N} dx_i \exp\left\{\sum_{i=1}^{N} -x_i A_{ij} x_j + \sum_{i=1}^{N} J_i x_i\right\}$$
$$\sim \left[\frac{\delta}{\delta J_n}\frac{\delta}{\delta J_m}\exp\left\{\frac{1}{4}J_i(A^{-1})_{ij}J_j\right\}\right]_{J=0} \det|A_{ij}|^{-1}$$
$$\sim (A^{-1})_{nm}(\det|A_{ij}|)^{-1}. \tag{1.7.11}$$

These are some of the most important integrals in this book. Using these equations, we can now write the Green's function totally in terms of second quantized fields:

$$K(a,b) = \int \psi^*(x_a,t_a)\psi(x_b,t_b)D\psi^* D\psi \exp\left[i\int dx dt L(\psi)\right], \tag{1.7.12}$$

where

$$L(\psi) = \psi^*(i\partial_t - H)\psi, \tag{1.7.13}$$

where we are again treating $K(a,b)$ as if it were a matrix in discretized (x,t) space.

In order to satisfy the canonical commutation relation (1.8.2), we must have

$$[a, a^\dagger] = 1. \tag{1.8.5}$$

If we insert this expression back into the Hamiltonian, we find

$$H = \tfrac{1}{2}\omega(aa^\dagger + a^\dagger a). \tag{1.8.6}$$

By extracting a c-number term, we can write this in normal ordered fashion:

$$H = \omega(a^\dagger a + h_0), \tag{1.8.7}$$

where h_0 is the zero point energy. We can now introduce the Hilbert space of harmonic oscillators. Let us define the vacuum as

$$a\,|0\rangle = 0. \tag{1.8.8}$$

Then an element of the Fock space of the harmonic oscillator Hamiltonian is given by

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}\,|0\rangle \tag{1.8.9}$$

such that the states form an orthonormal basis:

$$\langle n|m\rangle = \delta_{nm}. \tag{1.8.10}$$

The energy of the system is quantized and given by

$$E_n = (n + \tfrac{1}{2})\omega. \tag{1.8.11}$$

So far, the systems has been presented only in a first quantized formalism. We are quantizing only a single point particle at any time. We would now like to make the transition to the second quantized wave function by introducing

$$|\Phi\rangle = \sum_{n=0}^{\infty} \phi_n\,|n\rangle\,, \tag{1.8.12}$$

where we power expand in the basis states of the harmonic oscillator. Thus, instead of describing a single excited state of a point particle, we are now introducing the wave function, which will be a superposition of an arbitrary number of excited states.

Let us make the important definition

$$\langle x|\Phi\rangle = \Phi(x). \tag{1.8.13}$$

This can be calculated explicitly. Notice that we now have two independent basis states, the harmonic oscillator basis $|n\rangle$ and the position eigenvectors $|x\rangle$. We must now calculate how to go back and forth between these two bases.

Let us first analyze the simplest matrix element:

$$\sigma_0(x) = \langle x|0\rangle. \tag{1.8.14}$$

This matrix element satisfies the equation

$$\begin{aligned} 0 &= \langle x|\, a\, |0\rangle \\ &= \langle x|\, \frac{p - im\omega x}{\sqrt{2m\omega}}\, |0\rangle \\ &= (2m\omega)^{-1/2} \left(-\frac{i\partial}{\partial x} - im\omega x \right) \langle x|0\rangle \\ &= -i(2m\omega)^{-1/2} \left(\frac{\partial}{\partial x} + m\omega x \right) \sigma_0(x). \end{aligned} \tag{1.8.15}$$

This last equation can be solved exactly:

$$\sigma_0(x) = (m\omega/\pi)^{1/4} e^{-(1/2)\xi^2}, \tag{1.8.16}$$

where

$$\xi = (m\omega)^{1/2} x. \tag{1.8.17}$$

It is now a straightforward step to calculate all such matrix elements. Let

$$\begin{aligned} \sigma_n(x) &= \langle x|n\rangle \\ &= \langle x|\, (n!)^{1/2} a^{\dagger n}\, |0\rangle \\ &= (n!)^{-1/2} (2m\omega)^{-n(1/2)} \langle x|\, [p + im\omega x]^n\, |0\rangle \\ &= (n!)^{-1/2} (2m\omega)^{-(1/2)n} \left(-i\frac{\partial}{\partial x} + im\omega x \right)^n \sigma_0(x). \end{aligned} \tag{1.8.18}$$

The solution is therefore

$$\sigma_n(x) = i^n (2^n n!)^{-1/2} (m\omega/\pi)^{1/4} \left(\xi - \frac{\partial}{\partial \xi} \right)^n e^{-(1/2)\xi^2}. \tag{1.8.19}$$

In general, these are nothing but Hermite polynomials H_n. In terms of these polynomials, we can express the eigenstate $|x\rangle$ and $|n\rangle$ in terms of each other:

$$\begin{cases} |x\rangle = \sum_{n=1}^{\infty} |n\rangle \langle n|x\rangle = \sum_{n=1}^{\infty} |n\rangle \sigma_n(x), \\ |n\rangle = |x\rangle \int dx \langle x|n\rangle = \int dx \sigma_n(x) |x\rangle . \end{cases} \tag{1.8.20}$$

Thus, using (1.8.12) and (1.8.20), we have the power expansion of the wave function in terms of a complete set of orthogonal polynomials, the Hermite polynomials:

$$\Phi(x) = \langle x|\Phi\rangle = \langle x| \sum_{n=1}^{\infty} \phi_n\, |n\rangle = \sum_{n=1}^{\infty} \phi_n H_n(\xi) e^{-(1/2)\xi^2}. \tag{1.8.21}$$

Similarly, it is not difficult to calculate the Green's function for the propagation of a point particle in a harmonic oscillator potential. The Green's function

would be the same as if we had started with the second quantized formalism with the action:

$$L = \Phi(x)^* \left(i\partial_t + \frac{1}{2m}\nabla^2 - \tfrac{1}{2}kx^2 \right) \Phi(x). \tag{1.8.22}$$

From this second quantized action, we can therefore derive the equations of motion:

$$\begin{aligned} i\partial_t \Phi(x,t) &= \left[\frac{-1}{2m}\nabla^2 + \tfrac{1}{2}kx^2 \right] \Phi(x,t) \\ &= H\Phi(x,t). \end{aligned} \tag{1.8.23}$$

From this, we can define the canonical momenta conjugate to $\Phi(x,t)$ such that the canonical quantization relations are satisfied:

$$[\Pi(x,t), \Phi(x',t)] = -i\delta(x-x'). \tag{1.8.24}$$

1.9 Currents and Second Quantization

Let us begin with a discussion of the relativistic second quantized theory, which, as we have seen, is equivalent perturbatively to the first quantized theory. When quantizing the point particle in the Gupta–Bleuler formalism, we were led to the equations of motion:

$$[\Box - m^2]\phi = 0 \tag{1.9.1}$$

which can be derived from the second quantized action:

$$L = \tfrac{1}{2}[\partial_\mu \phi \partial^\mu \phi + m^2\phi^2]. \tag{1.9.2}$$

One of the most powerful techniques we used to explore the first quantized theory was symmetry. We would now like to study the question of symmetries within the second quantized formalism.

First, let us calculate the equations of motion by making a small variation in the field and requiring that the action be invariant under this variation:

$$\delta S = 0 = \int d^D x \left(\frac{\delta L}{\delta \phi}\delta\phi + \frac{\delta L}{\delta \partial_\mu \phi}\delta\partial_\mu \phi \right). \tag{1.9.3}$$

Let us now integrate by parts, using $\delta\partial_\mu\phi = \partial_\mu\delta\phi$:

$$\delta S = \int d^D x \left(\frac{\delta L}{\delta\phi} - \partial_\mu \frac{\delta L}{\delta \partial_\mu \phi} \right) \delta\phi + \int d^D x \partial_\mu \left(\frac{\partial L}{\delta \partial_\mu \phi}\delta\phi \right). \tag{1.9.4}$$

If we temporarily ignore the surface term, the action is stationary if we have the following equation of motion:

$$\partial_\mu \frac{\delta L}{\delta \partial_\mu \phi} - \frac{\delta L}{\delta \phi} = 0 \tag{1.9.5}$$

If we insert the Lagrangian into this equation, we obtain the equations of motion, which reproduce the constraint found earlier in the first quantized formalism.

Let us now make a small change in the fields, parametrized by a small, as yet unspecified, number ε^α:

$$\delta\phi = \frac{\delta\phi}{\delta\varepsilon^\alpha}\delta\varepsilon^\alpha. \tag{1.9.6}$$

If we insert this into the previous equation for the variation of the action, keep the surface term intact, and assume that the equations of motion are satisfied, then we have the following:

$$\delta S = \int d^D x \partial_\mu \left(\frac{\partial L}{\delta\partial_\mu\phi}\frac{\delta\phi}{\delta\varepsilon^\alpha} \right) \delta\varepsilon^\alpha. \tag{1.9.7}$$

Let us define the tensor in the parentheses as the *current*:

$$J^\mu_\alpha = \frac{\delta L}{\delta\partial_\mu\phi}\frac{\delta\phi}{\delta\varepsilon^\alpha}. \tag{1.9.8}$$

Then we have the important equation

$$\delta S = \int d^D x \partial_\mu J^\mu_\alpha \delta\varepsilon^\alpha. \tag{1.9.9}$$

Thus, if the action S is stationary under this variation, we have a conserved current $J^{\mu\alpha}$:

$$\partial_\mu J^{\mu\alpha} = 0. \tag{1.9.10}$$

We will use this equation over and over again in the discussion of strings when we want to extract the current for supersymmetry and conformal invariance. Finally, we note that the integrated charge Q^α associated with the current is constant in time:

$$\int d^D x \partial_\mu J^{\mu\alpha} = \int d^{D-1} x \partial_0 J^{0\alpha} + \text{surface term}. \tag{1.9.11}$$

Thus,

$$\begin{aligned} Q^a &= \int d^{D-1} x J^\alpha_0, \\ \partial_\mu J^{\mu\alpha} = 0 \quad &\rightarrow \quad \frac{dQ^\alpha}{dt} = 0. \end{aligned} \tag{1.9.12}$$

Finally, we wish to construct yet another conserved current associated with the action. Let us make a small variation in the space–time variable:

$$\delta x^\mu = \varepsilon^\mu. \tag{1.9.13}$$

Under this charge, the volume element of the integral changes as

$$\delta d^D x = d^D x \partial_\mu \delta x^\mu. \tag{1.9.14}$$

Coulomb Quantization

The gauge invariance permits us to take the gauge

$$\nabla_i A_i^a = 0.$$

We can integrate over the A_0 component because it has no time derivatives, so the Coulomb formulation is explicity ghost-free. (The price we pay for this, of course, is the loss of manifest Lorentz invariance, which must be checked by hand.) In this gauge, the action becomes

$$L = +\tfrac{1}{2}(\partial_0 A_i^a)^2 - \tfrac{1}{4}(F_{ij}^a)^2 + \cdots, \tag{1.9.25}$$

where all fields are transverse. This is the canonical form for the Lagrangian.

Gupta–Bleuler Quantization

The advantage of the Gupta–Bleuler formulation is that we can keep manifest Lorentz symmetry without violating unitarity. For example, let us take the gauge

$$\partial_\mu A^{\mu a} = 0. \tag{1.9.26}$$

In this gauge, the propagator for massless vector particles becomes

$$\frac{\eta_{\mu\nu}}{p^2}. \tag{1.9.27}$$

Notice that the propagator explicitly contains a ghost. The timelike excitation has a coefficient of -1 in the propagator, which represents a ghost. However, we are free to quantize in this covariant approach because we will impose the ghost-killing constraint on the Hilbert space:

$$\langle \phi |\, \partial_\mu A^{\mu a} \,| \psi \rangle = 0. \tag{1.9.28}$$

This constraint allows us to solve for and hence eliminate the ghost modes. Thus, although the free propagator will allow ghosts to propagate, the Hilbert space is ghost-free, so the theory itself is both Lorentz invariant and ghost-free.

BRST Quantization

The BRST approach begins by calculating the Faddeev–Popov determinant (1.6.10). Let us calculate the determinant of the matrix:

$$\begin{aligned} M^{ab}(x, y) &= \frac{\delta(\partial_\mu A^{\mu a}(x))}{\delta \Lambda^b(y)} \\ &= \partial_\mu D^\mu \frac{\delta \Lambda^a(x)}{\delta \Lambda^b(y)} \\ &= \partial_\mu D^\mu (\delta^4(x - y)\delta^{ab}). \end{aligned} \tag{1.9.29}$$

Unfortunately, string theory evolved historically as a first quantized theory. Thus, string theory has been evolving backward, with the second quantized geometric theory still in its infancy. For pedagogical reasons, we have introduced string theory from a semihistorical point of view, beginning with the first quantized theory and later developing the second quantized theory and M-theory. We hope that future accounts of string theory will reverse this sequence.

To reduce the level of arbitrariness in the first quantized theory as much as possible, in this chapter we have tried to lay the groundwork for string theory in the formalism of path integrals. This functional formalism has the great advantage that we can express the first and second quantized gauge theories with equal ease. We find, in fact, that large portions of the path integral formulation of point particles can be incorporated wholesale into string theory.

The path integral method postulates two fundamental principles that express the essence of quantum mechanics:

(1) The probability $P(a, b)$ of a particle going from point a to point b is given by the absolute value squared of a transition function $K(a, b)$,

$$P(a, b) = |K(a, b)|^2.$$

(2) The transition function is given by the sum of a phase factor e^{iS}, where S is the action, taken over all possible paths from a to b,

$$K(a, b) = \sum_{\text{paths}} k e^{iS}.$$

In the limit of continuous paths, we have

$$K(a, b) = \int_a^b Dx e^{iS},$$

where

$$Dx = \lim_{N\to\infty} \prod_{i=1}^{3} \prod_{n=1}^{N} dx_{i,n}.$$

The action S of the first quantized point particle is given by the length of the path that the particle sweeps out in space–time. We can represent the Lagrangian for the point particle in three ways:

$$\begin{aligned} &\text{first-order (Hamiltonian) form:} \quad L = p_\mu \dot{x}^\mu - \tfrac{1}{2} e(p_\mu^2 + m^2), \\ &\text{second-order form:} \quad L = \tfrac{1}{2}(e^{-1}\dot{x}_\mu^2 - em^2), \qquad (1.10.1) \\ &\text{nonlinear form:} \quad L = -m\sqrt{-\dot{x}_\mu^2}. \end{aligned}$$

Unfortunately, because all three forms of the action are parametrization-invariant, the path integral diverges. Thus, the quantization procedure must break this gauge symmetry and yield the correct measures in the functional.

These actions can be quantized in three ways, each with its own advantages and disadvantages:

(1) Coulomb Quantization
By explicitly fixing the value of some of the fields, such as

$$x_0 = t = \tau$$

we can eliminate the troublesome negative metric states and the Lagrangian becomes $\frac{1}{2}mv_i^2$. The Coulomb quantization method is therefore manifestly ghost-free. However, the disadvantage of this method is that it is very awkward because manifest Lorentz symmetry is broken and must be checked at every level.

(2) Gupta–Bleuler Quantization
The advantage of the Gupta–Bleuler quantization method is that we have a manifestly covariant quantization program. Of course, negative metric ghosts are now allowed to circulate in the theory, but they are eventually eliminated by imposing the gauge constraints directly onto the Hilbert space:

$$[p_\mu^2 + m^2]\,|\phi\rangle = 0.$$

Thus, the S-matrix is ultimately ghost-free. The disadvantage of this approach, however, is that the imposition of these gauge constraints, especially at the interacting level, is frequently quite difficult.

(3) BRST Quantization
This method of quantization keeps the good features of both approaches. The theory is manifestly covariant, but the S-matrix is still unitary because the addition of ghost fields in the theory cancels precisely against the negative metric states. The BRST method imposes the gauge $e = 1$ in the first-order form and then inserts the Faddeev–Popov term Δ_{FP} into the functional to get the correct measure. We can exponentiate this determinant into the action by using Grassmann variables:

$$\Delta_{\text{FP}} = \det|\partial_\tau| = \int d\theta\, d\bar{\theta} e^{i\int d\tau \bar{\theta}\partial_\tau \theta}.$$

The resulting gauge-fixed action has a residual symmetry, called the BRST symmetry, which is generated by Q, the BRST charge.

When we generalize these methods to the interacting case, the path integral formulation begins with the fundamental formula for the transition function for N-particle scattering:

$$\begin{aligned} A(k_1, k_2, \ldots, k_N) &= \sum_{\text{topologies}} g^n \int Dx\, \Delta_{\text{FP}} \\ &\quad \times \exp\left\{ i\int dt\, L(t) + i\sum_{i=1}^{N} k_\mu x_i^\mu \right\} \\ &= \sum_{\text{topologies}} g^n \langle e^{i\sum_{i=1}^{n} k_\mu x_i^\mu} \rangle. \end{aligned} \tag{1.10.2}$$

The first quantized description for the N-particle scattering amplitude is clumsy because we must explicitly sum over certain topologies, which must be put in by hand. This means that unitarity is not obvious in the first quantized formalism. Later, we will see that this problem in the first quantized point particle theory carries over directly into the first quantized string theory. In the second quantized description, however, all topologies can be derived explicitly from a single action.

The transition from a first to a second quantized description is straightforward in the path integral formulation. For example, the propagator can be written in either first or second quantized language:

$$\begin{aligned}\Delta_{ab} &= \int_{x_a}^{x_b} Dx\, e^{i\int_{t_a}^{t_b} dt\, L(t)} \\ &= \int D\psi\, D\psi^* \psi(x_a)\psi^*(x_b) e^{i\int Dx L(\psi)},\end{aligned} \tag{1.10.3}$$

where

$$\begin{aligned}\langle x|\psi\rangle &= \psi(x), \\ L(t) &= \tfrac{1}{2}m\dot{x}_i^2, \\ L(\psi) &= \psi^*(i\partial_t - H)\psi.\end{aligned} \tag{1.10.4}$$

The last equation is the Lagrangian for the Schrödinger wave equation, which can be derived beginning with the postulates of path integrals and $L = \frac{1}{2}mv_i^2$.

For the interactions, it is also possible to extract the second quantized vertices from the first quantized theory in exactly the same fashion. We simply write down the functional integral over a world sheet where the point particle splits into other point particles, and then write this Green's function as a functional integral over second quantized fields.

We will shortly see the advantage of carefully working out the details of point particle path integrals. We will find that almost all of this formalism carried over directly into the string formalism!

References

[1] For reviews of string theory, see J. H. Schwarz, ed., *Superstrings*, Vols. 1 and 2, World Scientific, Singapore, 1985; J. H. Schwarz, *Phys. Rep.* **89**, 223 (1982).
[2] M. B. Green, *Surv. High Energy Phys.* **3**, 127 (1983).
[3] V. Alessandrini, D. Amati, M. LeBellac, and D. I. Olive, *Phys. Rep.* **1C**, 170 (1971).
[4] S. Mandelstam, *Phys. Rep.* **13C**, 259 (1974).
[5] C. Rebbi, *Phys. Rep.* **12C**, 1 (1974).
[6] P. Frampton, *Dual Resonance Models*, Benjamin, New York, 1974.
[7] J. Scherk, *Rev. Mod. Phys.* **47**, 1213 (1975).
[8] G. Veneziano, *Phys. Rep.* **9C**, 199 (1974).
[9] M. Jacob, ed., *Dual Theory*, North-Holland, Amsterdam, 1974.

[10] J. H. Schwarz, *Phys. Rep.* **8C**, 269 (1973).
[11] M. B. Green and D. Gross, eds., *Unified String Theories*, World Scientific, Singapore, 1986.
[12] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory*, Vols. 1 and 2, Cambridge University Press, Cambridge, 1986.
[13] H. Georgi and S. L. Glashow, *Phys. Rev. Lett.* **32**, 438 (1974).
[14] D. Z. Freedman, P. Van Nieuwenhuizen, and S. Ferrara, *Phys. Rev.* **D13**, 32 (1976).
[15] S. Deser and B. Zumino, *Phys. Lett.* **62B**, 335 (1976).
[16] T. Regge, *Nuovo Cimento* **19** 558 (1961).
[17] N. H. Christ, R. Friedberg, and T. D. Lee, *Nucl. Phys.* **B202**, 89 (1982).
[18] G. C. Wick and T. D. Lee, *Nucl. Phys.* **B9**, 209 (1969).
[19] M. Kaku, *Nucl. Phys.* **B203**, 285 (1982).
[20] M. Kaku, *Phys. Rev. Lett.* **50**, 1895 (1983); *Phys. Rev.* **27**, 2809, 2819 (1983).
[21] G. Veneziano, *Nuovo Cimento* **57A**, 190 (1968).
[22] M. Suzuki, unpublished.
[23] P. Ramond, *Phys. Rev.* **D3**, 2415 (1971).
[24] A. Neveu and J. H. Schwarz, *Nucl. Phys.* **B31**, 819 (1971).
[25] A. Neveu and J. H. Schwarz, *Phys. Rev.* **D4**, 1109 (1971).
[26] K. Kikkawa, B. Sakita, and M. A. Virasoro, *Phys. Rev.* **184**, 1701 (1969).
[27] M. Kaku and L. P. Yu, *Phys. Lett.* **33B**, 166 (1970).
[28] M. Kaku and L. P. Yu, *Phys. Rev.* **D3**, 2992, 3007, 3020 (1971).
[29] M. Kaku and J. Scherk, *Phys. Rev.* **D3**, 430 (1971).
[30] M. Kaku and J. Scherk, *Phys. Rev.* **D3**, 2000 (1971).
[31] V. Alessandrini, *Nuovo Cimento* **2A**, 321 (1971).
[32] C. Lovelace, *Phys. Lett.* **32B**, 703 (1970).
[33] C. Lovelace, *Phys. Lett.* **34B**, 500 (1971).
[34] Y. Nambu, Lectures at the Copenhagen Summer Symposium (1970).
[35] T. Goto, *Progr. Theoret. Phys.* **46**, 1560 (1971).
[36] P. Goddard, J. Goldstone, C. Rebbi, and C. B. Thorn, *Nucl. Phys.* **B56**, 109 (1973).
[37] S. Mandelstam, *Nucl. Phys.* **B64**, 205 (1973); **B69**, 77 (1974).
[38] M. Kaku and K. Kikkawa, *Phys. Rev.* **D10**, 1110, 1823 (1974).
[39] J. Scherk and J. H. Schwarz, *Nucl. Phys.* **B81**, 118 (1974). See also T. Yonea, *Progr. Theoret. Phys.* **51**, 1907 (1974).
[40] M. B. Green and J. H. Schwarz, *Phys. Rev.* **149B**, 117 (1984); **151B**, 21 (1985).
[41] D. J. Gross, J. A. Harvey, E. Martinec, and R. Rohm, *Nucl. Phys.* **B256**, 253 (1986); **B267**, 75 (1986).
[42] C. S. Hsue, B. Sakita, and M. A. Virasoro, *Phys. Rev.* **D2**, 2857 (1970).
[43] J. L. Gervais and B. Sakita, *Nucl. Phys.* **B34**, 632 (1971); *Phys. Rev.* **D4**, 2291 (1971); *Phys Rev. Lett.* **30**, 716 (1973); see also D. B. Fairlie and H. B. Nielsen, *Nucl. Phys.* **B20**, 637 (1970).
[44] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.
[45] C. Becchi, A. Rouet, and R. Stora, *Ann. Phys.* **98**, 287 (1976).
[46] I. V. Tyupin, Lebedev preprint FIAN No. 39 (1975), unpublished.
[47] L. D. Faddeev and V. N. Popov, *Phys. Lett.* **25B**, 29 (1967).

$$= \langle x_1 | \int_0^\infty d\tau e^{-t(\Box - m^2)} | x_2 \rangle$$
$$= \int_0^\infty d\tau \int_{x_1}^{x_2} Dx e^{-(1/2)\int_0^\tau d\bar{\tau}(\dot{x}_\mu^2 - m^2)}, \tag{2.1.1}$$

where we integrate over all possible trajectories of a particle located at $x_\mu(\tau)$ which start at x_1 and end at point x_2. The interactions, we saw, were introduced by hand into the theory by postulating a particular set of topologies over which this particle can roam. The scattering amplitude, for example, is

$$A(k_1, k_2, \ldots, k_n) = \sum_{\text{topologies}} \int Dx \ \Delta_{\text{FP}} e^{-\int L d\tau + i \sum_{i=1}^N k_i \cdot x_i}$$
$$= \sum_{\text{topologies}} \left\langle \prod_{i=1}^N e^{i k_i \cdot x^i} \right\rangle, \tag{2.1.2}$$

where we integrate over topologies that form the familiar Feynman diagrams for ϕ^3 or ϕ^4 theory.

It is important to notice that the resulting Feynman diagram is a graph, *not* a manifold. At the interaction point, the local topology is not $\mathbf{R}^n$, so it cannot be a manifold. There is no correlation between the internal lines and the interaction points. This means that we can introduce arbitrarily high spins at the interaction point of the first quantized relativistic point particle. Thus, the first quantized point particle theory has an infinite degree of arbitrariness, corresponding to the different spins and masses we can place at the interaction point. Furthermore, the ultraviolet singularities of each Feynman diagram correspond to the number of ways we can "pinch" the diagram by shrinking an internal line to zero, thus deforming the local topology.

This picture, however, totally changes with the string. Although the path integral formalism looks almost identical, there are profoundly important differences. In particular, the sum over histories becomes a sum over all possible

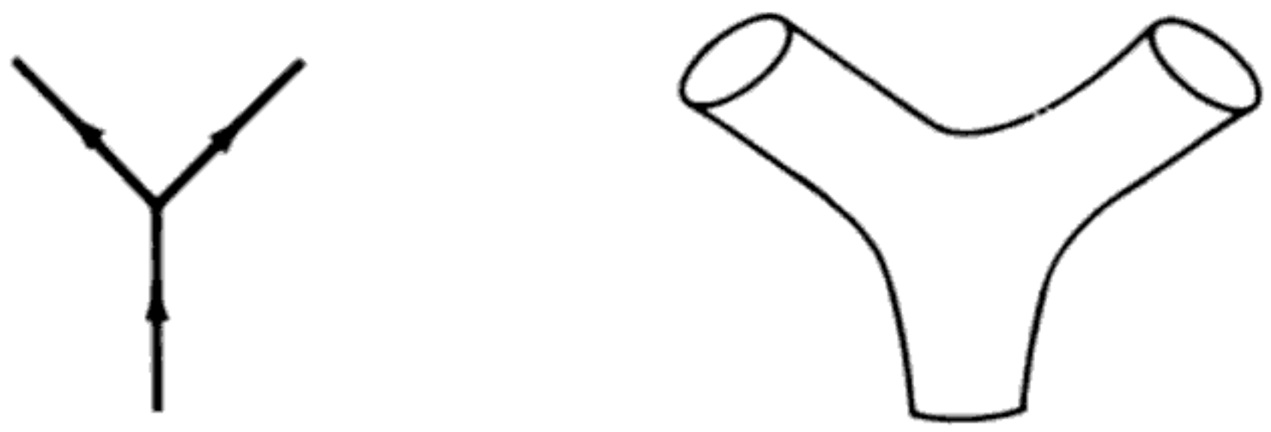

FIGURE 2.1. Vertex functions for point particles and strings. A large number of point particle theories are possible, based on different spins and isospins, because the Feynman diagrams are graphs. Only a few string theories are known, however, because the interactions are restricted to be manifolds, not graphs. Conformal symmetry, modular invariance, and supersymmetry place enormous restrictions on the manifolds we may use to construct superstring theories that have no counterpart in point particle theory.

The contraction of two of these tangent vectors yields a metric:

$$g_{ab} = \partial_a X_\mu \partial_b X^\mu, \tag{2.1.5}$$

where we have now replaced the two variables (τ, σ) with the set (a, b), where a, b can equal either 0 or 1. The infinitesimal area on this surface can be written simply as

$$d\,\text{Area} \sim \sqrt{\det |g_{ab}|}\, d\sigma\, d\tau. \tag{2.1.6}$$

In analogy to the point particle case, where the action is the length swept out by the point, we now define our action to be the surface area of this world sheet. Our Lagrangian is therefore [1–4]:

$$L = \frac{1}{2\pi\alpha'}\sqrt{\dot{X}_\mu^2 X'^{\mu 2} - (\dot{X}_\mu X'^\mu)^2}, \tag{2.1.7}$$

where the prime represents σ differentiation and the dot represents τ differentiation. (When we discuss M-theory, we will introduce the theory of membranes in arbitrary dimensions in the same fashion. But instead of two coordinates on a world sheet, we now have $(p + 1)$-coordinates on a world volume. The action for the p-brane is still the same: the square root of the determinant of the metric tensor. The action is just the volume of the membrane.) The action is just the Lagrangian integrated over the world sheet, which is the total area of the two-dimensional surface:

$$S = \int d\sigma\, d\tau L(\sigma, \tau). \tag{2.1.8}$$

The Green's function for the propagation of a string from configuration X_a at "time" τ_a to configuration X_b at "time" τ_b, as well as the path integral over a surface that expresses the topology of several interacting strings, can be represented as

$$K(X_a, X_b) = \int_{X_a}^{X_b} DX e^{-\int_{\tau_a}^{\tau_b} d\tau \int_0^\pi d\sigma L},$$

$$Z = \sum_{\text{topologies}} \int d\mu DX e^{-\text{area}}, \tag{2.1.9}$$

where $DX = \prod_{\mu,\sigma,\tau} dX_\mu(\sigma, \tau), d\mu$ represents the measure of integration over the location of the external legs, and where we have made a Wick rotation in the τ variable $(\tau \to -i\tau)$ so the integral converges.

The correspondence between the point particle path integral formalism that we carefully developed in the previous chapter and the string formalism is quite remarkable. We find that almost the entire point particle formalism can

be imported into the string formalism:

$$\left\{ \begin{array}{c} x_\mu(\tau) \\ \text{length} \\ \prod_{\mu,\tau} dx_\mu(\tau) \end{array} \right\} \rightarrow \left\{ \begin{array}{c} X_\mu(\sigma,\tau) \\ \text{area} \\ \prod_{\mu,\sigma,\tau} dX_\mu(\sigma,\tau) \end{array} \right\}.$$

Similarly, the path integral for the point particle and the string theory have surprising similarities. The N-point function for the N-string scattering amplitude can also be written as a Fourier transform, similar to point particle path integrals:

$$\left\{ \begin{array}{c} \int_{x_i}^{x_j} Dx e^{-\int_{\tau_1}^{\tau_2} L(x)d\tau} \\ \left\langle \prod_{i=1}^{N} e^{ik_i x^i} \right\rangle \end{array} \right\} \rightarrow \left\{ \begin{array}{c} \int_{X_i}^{X_j} DX e^{-\int_0^\pi d\sigma \int_{\tau_1}^{\tau_2} L(X)d\tau} \\ \left\langle \prod_{i=1}^{N} e^{ik_i X^i} \right\rangle \end{array} \right\}.$$

Although there are remarkable similarities between point particle and string theories when expressed in the language of path integrals, the crucial difference between them emerges when we analyze the topologies over which the objects can move. For the point particle case, the topologies are *graphs*, as in Feynman graphs, whereas the topologies for string theories are *manifolds*:

$$\text{Graphs} \rightarrow \text{Manifolds}.$$

One of the crucial reasons why there are so many point particle actions (and so few string actions) is the difference between graphs and manifolds. The nontrivial restrictions placed on manifolds severely restrict the number of consistent string theories.

As in the case of the point particle, the choice of parametrization was totally arbitrary. Thus, our actions must be reparametrization invariant. To see this, let us make an arbitrary change of variables:

$$\begin{aligned} \tilde{\sigma} &= \tilde{\sigma}(\sigma,\tau), \\ \tilde{\tau} &= \tilde{\tau}(\sigma,\tau). \end{aligned} \tag{2.1.10}$$

Under this reparametrization, the string variable changes as

$$\delta X^\mu = X'^\mu \delta\sigma + \dot{X}^\mu \delta\tau. \tag{2.1.11}$$

Because the area of a surface is independent of the parametrization, the action is manifestly reparametrization invariant, which is easily checked.

As before, let us now write down the canonical conjugates of the theory:

$$P_\mu = \frac{\delta L}{\delta \dot{X}^\mu} = \frac{1}{2\pi\alpha'} \frac{X'^2 \dot{X}_\mu - (\dot{X}_\nu X'^\nu) X'_\mu}{\sqrt{\det |\partial_a X^\nu \partial_b X_\nu|}}.$$

As in the point particle case, these momenta are not all independent. In fact, we find two identities that are satisfied by the canonical momenta:

$$\text{Constraints:} \quad \begin{cases} P_\mu^2 + \dfrac{1}{(2\pi\alpha')^2} X_\mu'^2 \equiv 0, \\ P_\mu X'^\mu \equiv 0. \end{cases} \tag{2.1.12}$$

Thus, the canonical momenta are constrained by these two conditions. If we calculate the Hamiltonian of the system, we find that it vanishes identically as in (1.4.9):

$$H = P_\mu \dot{X}^\mu - L \equiv 0. \tag{2.1.13}$$

The vanishing of the Hamiltonian and the presence of constraints among the various momenta are indications that the system is a gauge system with an infinite redundancy. The reparametrization invariance of the system is the origin of this redundancy. We can therefore write down the close correspondence between the constraints of the point particle theory and the string theory:

$$\{p^2 + m^2 = 0\} \quad \rightarrow \quad \left\{ \begin{array}{c} P_\mu^2 + \dfrac{1}{(2\pi\alpha')^2} X_\mu'^2 = 0, \\ P_\mu X'^\mu = 0. \end{array} \right\}$$

Before we begin a detailed discussion of the quantization of the string, it is instructive to investigate the purely classical motions of this string. Let us first classically set the parameter τ equal to time, so that

$$\begin{aligned} \dot{X}_\mu &= (1, v_i), \\ X'_\mu &= (0, X'_i). \end{aligned} \tag{2.1.14}$$

Then let us factor out X'^2 from the action (2.1.7):

$$L = -\frac{1}{2\pi\alpha'} |X_i'^2|^{1/2} (1 - \tilde{v}_i^2)^{1/2}, \tag{2.1.15}$$

where $\tilde{v}$ is the velocity component perpendicular to the string:

$$\tilde{v}_i = v_i - \frac{v_k X'_k}{X_j'^2} X'_i. \tag{2.1.16}$$

The boundary conditions that we derive from this gauge-fixed action include

$$\tilde{v}_i^2 = 1. \tag{2.1.17}$$

This means that the ends of the classical string travel at the speed of light.

We can also calculate the energy of the classical string. Let us assume that the string is in a configuration that maximizes its angular momentum, i.e., it is a rigid rod that rotates with angular velocity ω around an axis labeled by the unit vector $\mathbf{r}$. The string can be parametrized as

$$\mathbf{X} = \sigma \mathbf{r}, \tag{2.1.18}$$

where

$$\dot{\mathbf{r}} = \boldsymbol{\omega} \times \mathbf{r},$$
$$\mathbf{r} \cdot \boldsymbol{\omega} = 0, \tag{2.1.19}$$

and $-l \leq \sigma \leq 1$.

To calculate the energy and the angular momentum of the system, we must first write down the Lorentz generators associated with the string:

$$M^{\mu\nu} = \int d\sigma (P^\nu X^\mu - P^\mu X^\nu). \tag{2.1.20}$$

Notice that this generates the algebra of the Lorentz group if we impose the Poisson brackets:

$$[X_\mu(\sigma), P_\nu(\sigma')] = \eta_{\mu\nu}\delta(\sigma - \sigma'). \tag{2.1.21}$$

We can now calculate the energy and the angular momentum from the components of the Lorentz generator [5]:

$$\begin{aligned} E &= \frac{1}{2\pi\alpha'} \int_{-l}^{l} d\sigma (1 - \omega^2\sigma^2)^{-1/2} \sim \frac{l}{2\alpha'}, \\ \mathbf{J} &= \frac{1}{2\pi\alpha'} \int_{-l}^{l} d\sigma \frac{\sigma^2}{(1 - \omega^2\sigma^2)^{1/2}} [\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})] \\ &= \boldsymbol{\omega}(4\alpha'\omega^3)^{-1} \\ &\sim \boldsymbol{\omega} E^2 \alpha'. \end{aligned} \tag{2.1.22}$$

Thus, the angular momentum of the rotating string is proportional to the square of the energy of the system:

$$|\mathbf{J}| \sim E^2. \tag{2.1.23}$$

If we plot the energy squared on the x-axis and the angular momentum on the y-axis, then we obtain a curve called the *Regge trajectory*. The slope of the Regge trajectory is given by α' and the curve is linear. Thus, we have obtained the leading Regge trajectory for the classical motion of a rigid rotator. We will, throughout this book, take the normalization $\alpha' = \frac{1}{2}$. This is an arbitrary convention. However, we will see later that the intercept a_0 of the leading trajectory must be equal to one, which is fixed by conformal invariance once we quantize the theory. Thus, we set

$$\begin{aligned} \alpha' &= \tfrac{1}{2}, \\ a_0 &= 1. \end{aligned} \tag{2.1.24}$$

When we quantize the system, we will find that there is an infinite number of such parallel Regge trajectories, but with increasingly negative y-intercepts.

As we have stressed, there is a crucial difference between the point particle case and the string, which is that the string system has a larger set of constraints that generate the gauge group of reparametrizations. For example, if

where

$$f \times g = fg' - gf'. \tag{2.1.30}$$

It is also possible to show that this algebra satisfies the Jacobi identities:

$$[L_{[f}, [L_g, L_{h]}] = 0, \tag{2.1.31}$$

where the brackets represent all possible cyclic symmetrizations. This algebra is called the *Virasoro algebra* [6], which will turn out to be one of the most powerful tools we have in constructing the string theory.

As in (1.4.11), we can elevate the constraints into the action with Lagrange multipliers $\lambda(\sigma, \tau)$ and $\rho(\sigma, \tau)$:

$$L = P_\mu \dot{X}^\mu + \pi\alpha'\lambda\left[P_\mu^2 + \frac{X_\mu'^2}{(2\pi\alpha')^2}\right] + \rho P_\mu X'^\mu. \tag{2.1.32}$$

By functionally integrating out over these Lagrange multipliers, we arrive at the previous set of constraints. Not surprisingly, this new action has its own reparametrization group parametrized by η and ε:

$$\begin{aligned}
\delta X_\mu &= 2\pi\alpha'\varepsilon P_\mu + \eta X_\mu', \\
\delta P_\mu &= \left[\frac{\varepsilon X_\mu'}{2\pi\alpha'} + \eta P_\mu\right]', \\
\delta\lambda &= -\dot{\varepsilon} + \lambda'\eta - \eta'\lambda + \rho'\varepsilon - \rho\varepsilon', \\
\delta\rho &= -\dot{\eta} + \lambda'\varepsilon - \lambda\varepsilon' + \rho'\eta - \eta'\rho.
\end{aligned} \tag{2.1.33}$$

The advantage of this form for the action is that it is first-order and does not have the bothersome square roots of the original action. As in the point particle case, this indicates that there exists yet one more form for the action, expressed in terms of an auxiliary field. To find this third formulation of the action, let us introduce a new independent field

$$g_{ab}(\sigma, \tau) \tag{2.1.34}$$

which represents a metric on a two-dimensional surface. Unlike our previous discussion, this metric is now totally independent of the string variable. Let us write down the Polyakov form of the action [7] ($g = |\det g_{ab}|$):

$$L = -\frac{1}{4\pi\alpha'}\sqrt{g}g^{ab}\partial_a X_\mu \partial_b X^\mu. \tag{2.1.35}$$

This is a generalization of the second-order point particle action (1.4.14). Notice that the Polyakov action resembles an action with scalar fields interacting with an external two-dimensional gravitational field. This action, too, possesses manifest reparametrization invariance:

$$\begin{aligned}
\delta X^\mu &= \varepsilon^a \partial_a X^\mu \\
\delta g^{ab} &= \varepsilon^c \partial_c g^{ab} - g^{ac}\partial_c \varepsilon^b - g^{bc}\partial_c \varepsilon^a, \\
\delta\sqrt{g} &= \partial_a(\varepsilon^a\sqrt{g}).
\end{aligned} \tag{2.1.36}$$

At first, we suspect that these actions are totally equivalent, so that we can choose one and drop the others. This is apparently not so, for two subtle reasons:

(1) Because we are dealing with a first quantized theory, we have to take the sum over all interacting topologies that are swept out by the string. For the Nambu–Goto string, the precise nature of these topologies is ambiguous and must be specified by hand. However, for the Polyakov form of the action, which contains an independent metric tensor, we can eliminate most of this ambiguity by specifying that we sum over all *conformally and modular inequivalent configurations*. (These terms will be defined later.) This will become a powerful constraint once we start to derive loops and will determine the function measure uniquely. The measure and the topologies in the Nambu–Goto action, however, are not well defined. (We must point out, however, that this rule of integrating over inequivalent surfaces does not automatically satisfy unitarity. This still must be checked by hand.)
(2) The gauge fixing of Weyl invariance for the Polyakov action, although trivial classically, poses problems when we make the transition to quantum mechanics. An anomaly appears when we carefully begin the quantization process. In fact, this conformal anomaly will disappear only in 26 dimensions!

Let us now discuss the quantization of the string action. The strategy we will take in quantizing the free theory to obtain the physical Hilbert space will be first to extract the symmetry of the action, then the currents, and then the algebra formed by the generators of this symmetry. (For the string, the symmetry will be reparametrization invariance and the algebra will be the Virasoro algebra.) Then we must apply the constraints onto the Hilbert space, which eliminates the ghosts and creates a unitary theory. It is important to keep this strategy in mind as we begin the quantization of the string.

$$\text{Action} \rightarrow \text{Symmetry} \rightarrow \text{Current} \rightarrow \text{Algebra} \rightarrow \text{Constraints} \rightarrow \text{Unitarity}.$$

As in the point particle case, we can begin the quantization program in several ways. There are three formalisms in which to fix the gauge of the theory: (1) Gupta–Bleuler (conformal gauge), (2) light cone gauge, and (3) BRST formalism. The advantages and disadvantages of each area as follows:

(1) The Gupta–Bleuler is perhaps the simplest of the three formalisms. We allow ghosts to appear in the action, which permits us to maintain manifest Lorentz invariance. The price we must pay, however, is that we must impose ghost-killing constraints on the Hilbert space. Projection operators must be inserted in all propagators. For trees, this is trivial. For higher loops, however, this is exceedingly difficult.
(2) The advantage of the light cone gauge formalism is that it is explicitly ghost-free in the action as well as the Hilbert space. There are no com-

plications when going to loops. However, the formalism is very awkward and Lorentz invariance must be checked at each step of the way.

(3) The BRST formalism combines the best features of the previous two formalisms. It is manifestly covariant, like the Gupta–Bleuler formalism, and it is unitary, like the light cone formalism, because the negative metric ghosts cancel against the Faddeev–Popov ghosts.

Let us now discuss each quantization scheme separately.

2.2 Gupta–Bleuler Quantization

The Gupta–Bleuler formalism will maintain Lorentz invariance by imposing the Virasoro constraints on the state vectors of the theory:

$$\langle\phi| L_f |\psi\rangle = 0, \tag{2.2.1}$$

where $\langle\phi|$ and $|\psi\rangle$ represent states of the theory. This constraint will eliminate ghosts in the state vectors, allowing us to keep nonphysical negative metric ghosts intact in the action.

Classically, the metric tensor has three degrees of freedom that we can gauge away, two arising from reparametrization invariance and one from Weyl invariance. Since the metric tensor has three degrees of freedom, we can gauge all of its components away:

$$g_{ab} = \delta_{ab} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.2.2}$$

which we call the *conformal gauge*. (There are complications, as we have said, in taking the conformal gauge for the quantum theory and for higher loops.) Our action reduces to

$$S = \frac{1}{4\pi\alpha'} \int_0^\pi d\sigma \int d\tau (\dot{X}_\mu^2 - X_\mu'^2). \tag{2.2.3}$$

This is exceptionally simple because the action now corresponds to an uncoupled free string. This action yields the free equations of motion:

$$\left(\frac{\partial^2}{\partial\sigma^2} - \frac{\partial^2}{\partial\tau^2} \right) X_\mu(\sigma, \tau) = 0 \tag{2.2.4}$$

with the boundary condition:

$$X_\mu'(0, \tau) = X_\mu'(\pi, \tau) = 0 \tag{2.2.5}$$

which we need to enforce when we integrate by parts and eliminate the surface term. The solutions of the equations of motion are arbitrary functions of $\sigma + \tau$ and $\sigma - \tau$:

$$X^\mu(\sigma, \tau) = X_1^\mu(\sigma + \tau) + X_2^\mu(\sigma - \tau). \tag{2.2.6}$$

The canonical commutation relations are now

$$[P_\mu(\sigma), X_\nu(\sigma')] = -i\eta_{\mu\nu}\delta(\sigma - \sigma'), \tag{2.2.7}$$

where

$$\delta(\sigma - \sigma') = \frac{1}{\pi}\left(1 + 2\sum_{n=1}^{\infty}\cos n\sigma \cos n\sigma'\right). \tag{2.2.8}$$

There is, of course, an infinite number of possible representations of the path integral. However, as in the point particle case, we can always choose the simplest one, the harmonic oscillator basis [8], where the Hamiltonian becomes diagonal. Unlike the point particle case, however, we now have an infinite number of oscillators, one set for each normal mode:

$$\begin{aligned} X_n^\mu &= \tfrac{1}{2} i\sqrt{2\alpha'}(a_n^\mu - a_{-n}^\mu), \\ P_n^\mu &= \frac{1}{\sqrt{2\alpha'}}(a_n^\mu + a_{-n}^\mu), \end{aligned} \tag{2.2.9}$$

where we can satisfy the canonical commutation relations if we set

$$[a_{n\mu}, a_{m\nu}^\dagger] = \delta_{nm}\eta_{\mu\nu}. \tag{2.2.10}$$

It is also conventional to introduce an equivalent set of oscillators:

$$\begin{aligned} \alpha_m^\mu &= \sqrt{m}a_m^\mu, \qquad m > 0, \\ \alpha_{-m}^\mu &= \sqrt{m}a_m^{\dagger\mu}, \qquad m > 0, \end{aligned}$$

$$X^\mu(\sigma, \tau) = x^\mu + 2\alpha' p^\mu \tau + i\sqrt{2\alpha'}\sum_{n\neq 0}\frac{\alpha_n^\mu}{n}e^{-in\tau}\cos n\sigma. \tag{2.2.11}$$

Written in this basis, the Hamiltonian takes on an especially simple form (see (1.3.37)):

$$\begin{aligned} H &= \int_0^\pi d\sigma(P_\mu \dot{X}^\mu - L) \\ &= \pi\alpha'\int_0^\pi \left(P_\mu^2 + \frac{1}{(2\alpha'\pi)^2}X_\mu'^2\right)d\sigma \\ &= \sum_{n=1}^{\infty} n a_{n\mu}^\dagger a_n^\mu + \alpha' p_\mu^2, \qquad \alpha_{0\mu} = \sqrt{2\alpha'}p_\mu, \end{aligned} \tag{2.2.12}$$

where we have made an infinite shift in the zero point energy. At this point, the mass of the lowest-order particle is not well defined because we made this infinite shift, but we will later show that this lowest particle is actually a tachyon. We will show that the intercept of the model is fixed at 1.

Notice that each oscillator mode is basically uncoupled from the other oscillator modes. In fact, the Hamiltonian is diagonal in the Fock space of harmonic oscillator excitations. Taking this specific representation of the string function

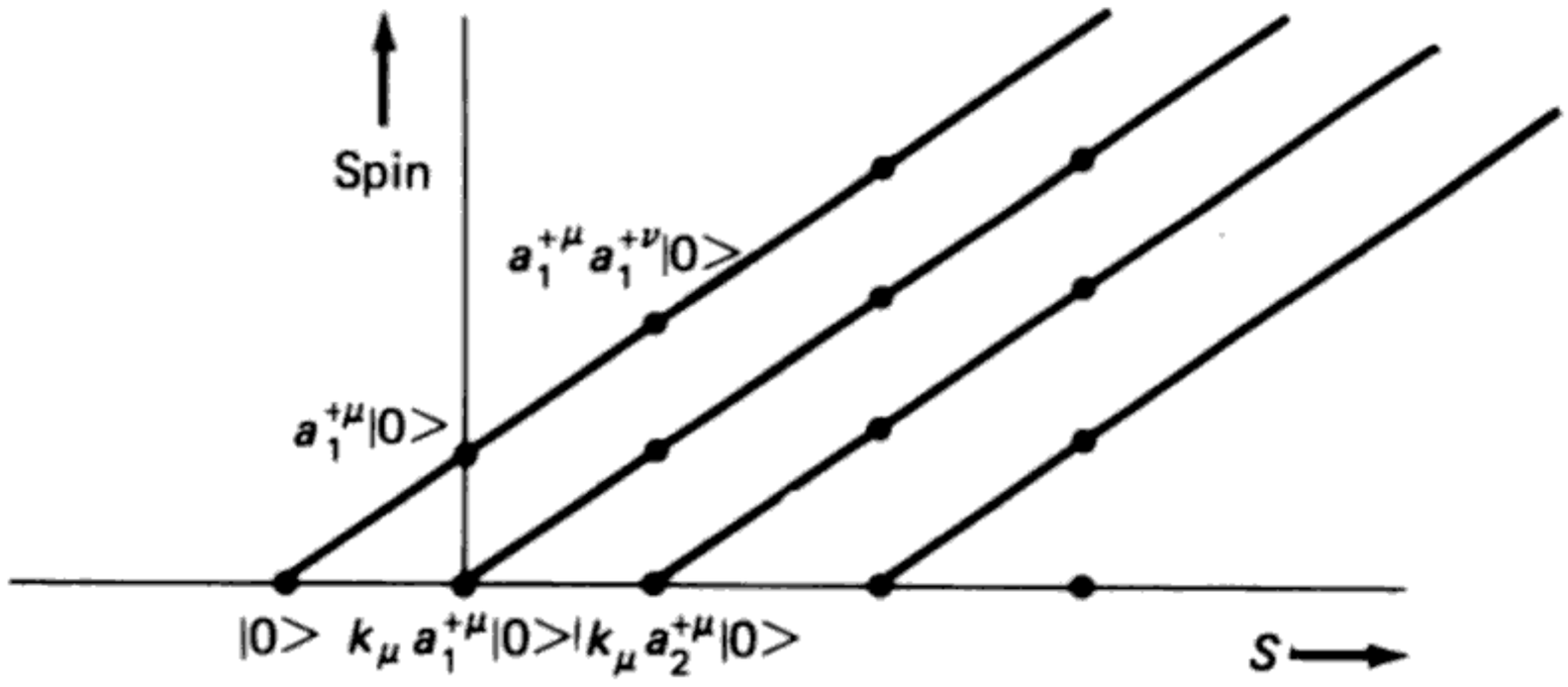

FIGURE 2.3. Regge trajectories for the open string. The x-axis corresponds to the energy squared and the y-axis to the spin. The particle farthest to the left is the tachyon, which corresponds to the vacuum of the Fock space. The massless spin-1 particle is the Maxwell or Yang–Mills field, which corresponds to a single creation operator acting on the vacuum. There is an infinite number of Regge trajectories, corresponding to the infinite excitations of a relativistic string or the infinite number of states in the Fock space.

from the infinite number of possibilities is a great advantage because the allowed eigenstates of our Hamiltonian are now simply the products of the Fock spaces of all possible harmonic oscillators:

$$\text{eigenstates:} \quad \prod_{n,\mu} \{a_{n,\mu}^\dagger\} \, |0\rangle \,, \tag{2.2.13}$$

where the vacuum is defined as

$$a_{n\mu} \, |0\rangle = 0, \qquad n \geq 0. \tag{2.2.14}$$

The spectrum of the lower lying states can be categorized as (see Fig. 2.3):

$$\begin{aligned} \text{tachyon} &\rightarrow |0\rangle \,, \\ \text{massless vector} &\rightarrow a_1^{\dagger\mu} \, |0\rangle \,, \\ \text{massless scalar} &\rightarrow k_\mu a_1^{\dagger\mu} \, |0\rangle \,, \\ \text{massive spin-2} &\rightarrow a_1^{\dagger\mu} a_1^{\dagger\nu} \, |0\rangle \,, \\ \text{massive vector} &\rightarrow a_2^{\dagger\mu} \, |0\rangle \,. \end{aligned} \tag{2.2.15}$$

(The fact that the string theory is so simple to quantize can be traced to the fundamental fact that the Hamiltonian is quadratic in the string variables; this means that it decomposes into an infinite series of free particles. A vast collection of vacuum solutions to string theory can be constructed because, in essence, it is a free theory. However, this simplicity breaks down completely when we analyze membranes in M-theory. We will find that the Hamiltonian is now quartic, making the quantization intractable. In contrast to the simplicity of string theory, there is still no satisfactory method for quantizing free membranes.)

As expected, we recover the leading Regge trajectory that we obtained earlier from classical arguments, and also an infinite number of daughter trajectories

The oscillator decomposition is given by

$$X_\mu(\sigma) = x_\mu + \left(\tfrac{1}{2}\alpha'\right)^{1/2} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} (a_n e^{-in\sigma} + \tilde{a}_n e^{in\sigma} + a_n^\dagger a^{in\sigma} + \tilde{a}_n^\dagger e^{-in\sigma})_\mu \tag{2.2.21}$$

$$P_\mu(\sigma) = \frac{p_\mu}{2\pi} + \frac{1}{2\pi\sqrt{2\alpha'}} \sum_{n=1}^{\infty} \sqrt{n}(-ia_n e^{-in\sigma} - i\tilde{a}_n e^{in\sigma} + ia_n^\dagger e^{in\sigma} + i\tilde{a}_n^\dagger e_n^{-in\sigma}).\mu$$

The Hamiltonian for the closed string is

$$H = \pi \int_0^{2\pi} d\sigma \left(\alpha' P_\mu^2 + \frac{X_\mu'^2}{4\pi^2\alpha'}\right)$$

$$= \sum_{n=1}^{\infty} (n a_n^\dagger a_n + n\tilde{a}_n^\dagger \tilde{a}_n) + \alpha' p_\mu^2. \tag{2.2.22}$$

Again, the Fock space consists of all elements created out of harmonic oscillators, but this time there is an extra constraint that is not found for the open string:

$$(L_0 - \tilde{L}_0)\,|\phi\rangle = 0. \tag{2.2.23}$$

(The interpretation of this constraint is that the closed string must be independent of the origin of the σ-coordinate. For example, the operator $\int d\sigma\, e^{i\sigma(L_0 - L_0)}$ can be interpreted in two ways. First, it generates rotations in σ-space, so we average over a rotation of 2π in σ-space. Second, if we perform the integral, we have $\delta(L_0 - \tilde{L}_0)$, which is constraint (2.2.23) when applied to the Hilbert space. We will return to this constraint later.)

The Fock space consists of (see Fig. 2.4):

$$\begin{aligned} &\text{tachyon} \rightarrow |0\rangle\,, \\ &\text{massless spin-2} \rightarrow a_1^{\dagger\mu} \tilde{a}_1^{\dagger\nu}\,|0\rangle\,, \end{aligned} \tag{2.2.24}$$

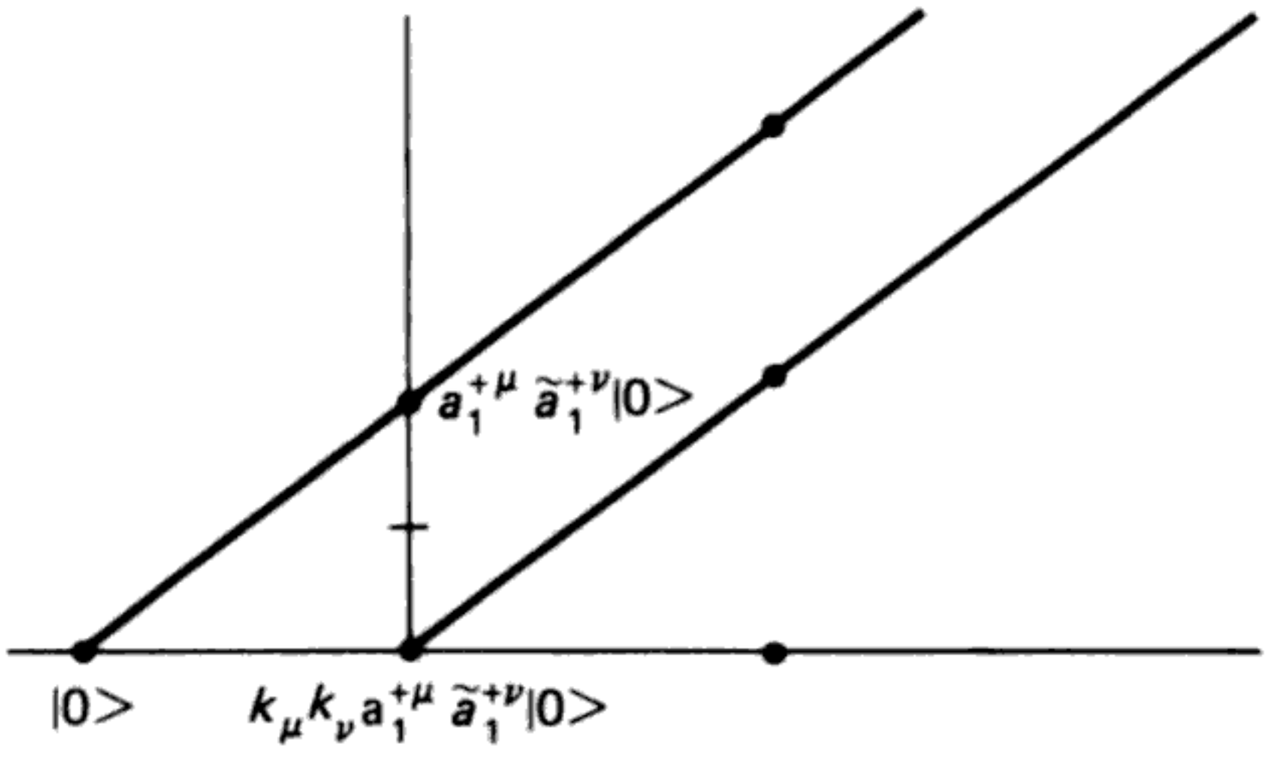

FIGURE 2.4. Regge trajectories for the closed string. The Fork space is built out of two commuting sets of harmonic oscillators. The massless spin-2 particle is the graviton, which corresponds to the product of both types of operators acting on the vacuum.

of quantization necessarily destroys the locality of Virasoro generators in the variable σ, and hence the c-number central term occurs. Thus, the quantization scheme and regularization scheme used to extract finite information from the model are actually inconsistent with conformal symmetry. Fortunately, this inconsistency can be eliminated if we fix the dimension of space–time to be 26.

That there are ghosts in the theory is due to the fact that the zeroth component of the harmonic oscillator has negative metric:

$$\text{ghost} = \{a^{\dagger}_{n,0}\}\,|0\rangle\,. \tag{2.2.29}$$

Thus, the coefficient of the Green's function occurs with a negative sign. In addition, there are zero norm states and negative norm states that must be taken into account.

To analyze the spectrum, let us define a *spurious state* $|S\rangle$ to be one that is orthogonal to all physical real states $|R\rangle$. Spurious states can be written as

$$\begin{aligned} |S\rangle &= L_{-n}\,|\chi\rangle\,, \qquad n > 0, \\ \langle R|S\rangle &= 0, \end{aligned}$$

for some integer n and some state $|\chi\rangle$. (If we take the matrix element of this state with a physical state, the scalar product always vanishes because L_n destroys a physical state.) Now let us construct the spurious state:

$$|\psi\rangle = [L_{-2} + aL^2_{-1}]\,|\phi\rangle\,. \tag{2.2.30}$$

We do not want this state to be part of the physical Hilbert space. However, let us see the conditions under which it might be part of the physical spectrum. Let us set

$$L_1\,|\psi\rangle = L_2\,|\psi\rangle = 0. \tag{2.2.31}$$

This fixes the following:

$$3 - 2a = 0, \qquad \tfrac{1}{2}D - 4 - 6a = 0, \tag{2.2.32}$$

which, in turn, fixes

$$D = 26, \qquad a = \tfrac{3}{2}. \tag{2.2.33}$$

Thus, this spurious state satisfies (2.2.26) and hence is part of the physical Fock space. At first, this seems disastrous. We want our physical Hilbert space to be ghost-free. But notice that in 26 dimensions this state has zero norm (not negative norm). Since $|\phi\rangle$ was arbitrary, we have constructed an infinite class of states $|\psi\rangle$ that are simultaneously spurious and physical. If we take the norm of this higher-order state, we find that it also vanishes in 26 dimensions, making it a null spurious state. This state is still acceptable because the norm of the state is nonnegative. Thus, in 26 dimensions we have an acceptable spectrum for this set of states.

Similar analyses of the state vectors of the theory show that a physical state $|\theta\rangle$ can be constructed which actually has negative norm if D is greater than

26:

$$|\theta\rangle = (a\alpha_{-1}^2 + bk\cdot\alpha_{-2} + c(k\cdot\alpha_{-1})^2)\,|0\rangle\,.$$

Imposing $L_1|\theta\rangle = L_2|\theta\rangle = 0$, we find that $b = a(D-1)/5$ and $c = a(D+4)/10$. We find that the norm of the state is

$$\langle\theta|\theta\rangle = \tfrac{2}{25}a^2(D-1)(26-D).$$

Thus, the dimension of space–time cannot exceed 26 or else negative norm states exist as part of the physical states. In general, we find that the spectrum is ghost-free if the dimension of space–time is less than or equal to 26:

$$\text{ghost-free:} \quad \left\{ \begin{array}{ll} a = 1, & D = 26, \\ a \le 1, & D \le 25. \end{array} \right\} \tag{2.2.34}$$

This exercise was done only for a piece of the Fock space. But can ghosts be eliminated to all orders in the string model? We will return to the difficult problem of ghost elimination in the Gupta–Bleuler formalism at the end of this chapter, when we actually construct the physical Hilbert space and show that it has no negative norm states in 26 dimensions.

2.3 Light Cone Quantization

Choosing the light cone gauge, where all unphysical degrees of freedom are explicitly removed from the very beginning, is possible because we have two gauge degrees of freedom, and hence two gauge-fixing conditions can be inserted into our path integral. One of these gauge-fixing constraints can be the elimination of nonphysical modes from the Hilbert space, as in the Coulomb gauge. Thus the elimination of ghost states, which is quite involved for the Gupta–Bleuler formalism (as we shall see at the end of this chapter), becomes trivial in the light cone gauge.

Let us choose the notation

$$\begin{aligned} X^+ &= \frac{1}{\sqrt{2}}[X^0 + X^{D-1}] \\ X^- &= \frac{1}{\sqrt{2}}[X^0 - X^{D-1}], \end{aligned} \tag{2.3.1}$$

then

$$A_\mu B^\mu = A_i B_i - A^+ B^- - A^- B^+. \tag{2.3.2}$$

Depending on which version of the action in (2.1.41) we use, we will have different gauge constraints. If we start with the original Nambu–Goto action, for example, the gauge conditions in the path integral are

$$Z = \int DXM\Delta_{\mathrm{FP}} \prod_\sigma \delta(X^+(\sigma) - p^+\tau)\delta(\dot{X}_\mu^2 + X_\mu'^2 - 2\dot{X}_\mu X'^\mu)e^{-S}, \tag{2.3.3}$$

where M is a measure term that must be added to have a unitary theory, and the two delta-functions represent the gauge-fixing constraints. The remarkable feature of the second constraint is that the Nambu–Goto action, which is expressed as a highly nonlinear square root, completely linearizes [12]

$$\sqrt{\dot{X}_\mu^2 X_\nu'^2 - (\dot{X}_\mu X'^\mu)^2} \sim \tfrac{1}{2}(\dot{X}_\mu^2 - X_\mu'^2). \tag{2.3.4}$$

(Because the light cone action is no longer a square root, we have a well-behaved action that can be canonically quantized.)

The constraint $X^+ = p^+\tau$ means that the σ dependence within X^+ has completely disappeared and that the "time" τ now beats in synchronism with X^+. We can use the second constraint, in turn, to eliminate the X^- modes, and hence all longitudinal modes have completely disappeared. The action can now be expressed totally in terms of transverse ones.

Next, we will solve the constraints in the first action in (2.1.41). Let us integrate over the Lagrangian multipliers ρ and λ in the Hamiltonian form of the action, and then impose the gauge-fixing constraints:

$$\begin{aligned} Z = \int DXDP \prod_\sigma \delta(X^+(\sigma) - p^+\tau)\delta\left(P^+(\sigma) - \frac{p^+}{\pi}\right) \\ \times \delta\left(P_\mu^2 + \frac{X_\mu^2}{\pi^2}\right)\delta(P_\mu x'^\mu)e^{-S}. \end{aligned} \tag{2.3.5}$$

Because the covariant Hamiltonian (2.1.13) is equal to zero, the only term in the Lagrangian is $P_\mu \dot{X}^\mu$ (the X^- term drops out):

$$L = \int_0^\pi d\sigma\, P_\mu \dot{X}^\mu = \int_0^\pi d\sigma(P_i \dot{X}^i - p^+ P^-(\sigma)). \tag{2.3.6}$$

There are several remarkable features to this formalism. First, we can apply four, not two, constraints onto our Hilbert space, two from gauge fixing and two by integrating over λ and ρ. Second, because the covariant Hamiltonian is equal to zero, the action only consists of $P_\mu \dot{X}^\mu$, but the light cone Hamiltonian emerges out of the decomposition of (2.3.6):

$$H = p^+ \int_0^\pi d\sigma\, P^-(\sigma). \tag{2.3.7}$$

On the other hand, we can solve the constraint for P^-:

$$P^-(\sigma) = \frac{\pi}{2p^+}\left(P_i^2 + \frac{X_i^2}{\pi^2}\right). \tag{2.3.8}$$

Plugging the value for P^- into the definition of the light cone Hamiltonian (2.3.7), we now have

$$H = \frac{\pi}{2}\int_0^\pi \left(P_i^2 + \frac{X_i^2}{\pi^2}\right) d\sigma \tag{2.3.9}$$

which is the just the Hamiltonian (2.2.12) defined over physical transverse states.

Similarly, we can also eliminate the X^- modes by solving another constraint:

$$P_\mu X'^\mu = 0 \tag{2.3.10}$$

which can be solved for X^-, yielding

$$X^-(\sigma) = \int_0^\sigma d\sigma' \frac{\pi}{p^+}[P_i X_i']. \tag{2.3.11}$$

Putting everything together, our functional now becomes

$$Z = \int DX_i DP_i e^{i\int d\tau \int d\sigma (P_i \dot{X}_i - H)}, \tag{2.3.12}$$

where H is the light cone Hamiltonian density. The great advantage of the light cone gauges is that the Virasoro constraints have been explicitly solved, so there is no need to impose them on states. All + modes have been gauged away from the start, and the − modes have been eliminated in terms of transverse states because the Virasoro conditions have been solved exactly through (2.3.8) and (2.3.11). Instead of imposing the Virasoro constraints on the Hilbert space, we simply solve them exactly and eliminate the − modes.

However, the great disadvantage of the formalism is that we must tediously check for Lorentz invariance at each step of the calculation. Normally, the generators of the Lorentz group are given by

$$\begin{aligned} M^{\mu\nu} &= \int_0^\pi d\sigma [X^\mu P^\nu - X^\nu P^\mu] \\ &= x^\mu p^\nu - x^\nu p^\mu - i \sum_{n=1}^\infty \frac{1}{n} [\alpha_{-n}^\mu \alpha_n^\nu - \alpha_{-n}^\nu \alpha_n^\mu]. \end{aligned} \tag{2.3.13}$$

It is easy to check from (2.2.7) that this satisfies the correct commutation relations for the Lorentz group:

$$[M^{\mu\nu}, M^{\alpha\beta}] = i\eta^{\mu\alpha} M^{\nu\beta} + \cdots. \tag{2.3.14}$$

However, Lorentz invariance has to be checked once again in light of the fact that we have explicitly all ghost modes. Most of the commutators are trivial to check, because they are linear. The troublesome term comes from X^-, which is highly nonlinear and is written as

$$X^- = x^- + p^- \tau + i \sum_{n\neq 0}^\infty \frac{1}{n} \alpha_n^- e^{-in\tau} \cos n\sigma, \tag{2.3.15}$$

where

$$\alpha_n^- = \frac{1}{2p^+} \sum_{i=1}^{D-2} \sum_{m=-\infty}^\infty [:\alpha_{n-m}^i \alpha_m^i: - 2a\delta_{n,0}] \tag{2.3.16}$$

$$\delta b = \varepsilon[c\partial_z b + 2\partial_z c b - \tfrac{1}{2}\partial_z X_\mu \partial_z X^\mu],$$
$$\delta \bar{b} = \varepsilon[\bar{c}\partial_{\bar{z}} \bar{b} + 2\partial_{\bar{z}} \bar{c}\bar{b} - \tfrac{1}{2}\partial_{\bar{z}} X_\mu \partial_{\bar{z}} X^\mu].$$

From this variation and (1.9.8), we can extract out the nilpotent BRST operator Q. However, it is also important to note that in general, given *any* Lie algebra with commutation relations $[\lambda_m, \lambda_n] = f_{mn}^p \lambda_p$, it is possible to construct a nilpotent operator Q [10] out of anticommuting operators c_n and b_m:

$$Q = \sum_{n=-\infty}^{\infty} c_{-n}\left[\lambda_n - \tfrac{1}{2} f_{nm}^p c_{-m} b_p\right], \tag{2.4.6}$$

where

$$\{c_n, b_m\} = \delta_{n,-m}. \tag{2.4.7}$$

Thus, our nilpotent BRST operator can be written in this form:

$$\begin{aligned} Q &= \sum_{n=-\infty}^{\infty} :c_{-n}(L_n^X + \tfrac{1}{2}L_n^{\text{gh}} - a\delta_{n,0}): \\ &= c_0(L_0 - a) + \sum_{n=1}^{\infty}[c_{-n}L_n + L_{-n}c_n] - \frac{1}{2}\sum_{n,m=-\infty}^{\infty} :c_{-m}c_{-n}b_{n+m}:(m-n), \end{aligned} \tag{2.4.8}$$

where L_n^X equals the X-dependent Virasoro generator, and L_n^{gh} is the ghost contribution to the generator. At this point, there are two unspecified parameters in the above equation, the value of the intercept a and the dimension of space–time. Let us calculate the square of Q, which should be zero:

$$Q^2 = \frac{1}{2}\sum_{m=-\infty}^{\infty}\left(\frac{D}{12}(m^3 - m) + \tfrac{1}{6}(m - 13m^3) + 2am\right) c_m c_{-m}. \tag{2.4.9}$$

For this to vanish, we must fix the dimension of space–time to be 26 and the intercept to be equal to 1.

As in the point particle case, we find that the physical states of the theory are given by

$$Q\,|\text{phy}\rangle = 0. \tag{2.4.10}$$

When we separate out the modes, we find that the lowest state satisfies

$$\begin{aligned} (L_0 - 1)\,|\phi\rangle &= 0, \\ L_n\,|\phi\rangle &= 0, \end{aligned} \tag{2.4.11}$$

as before.

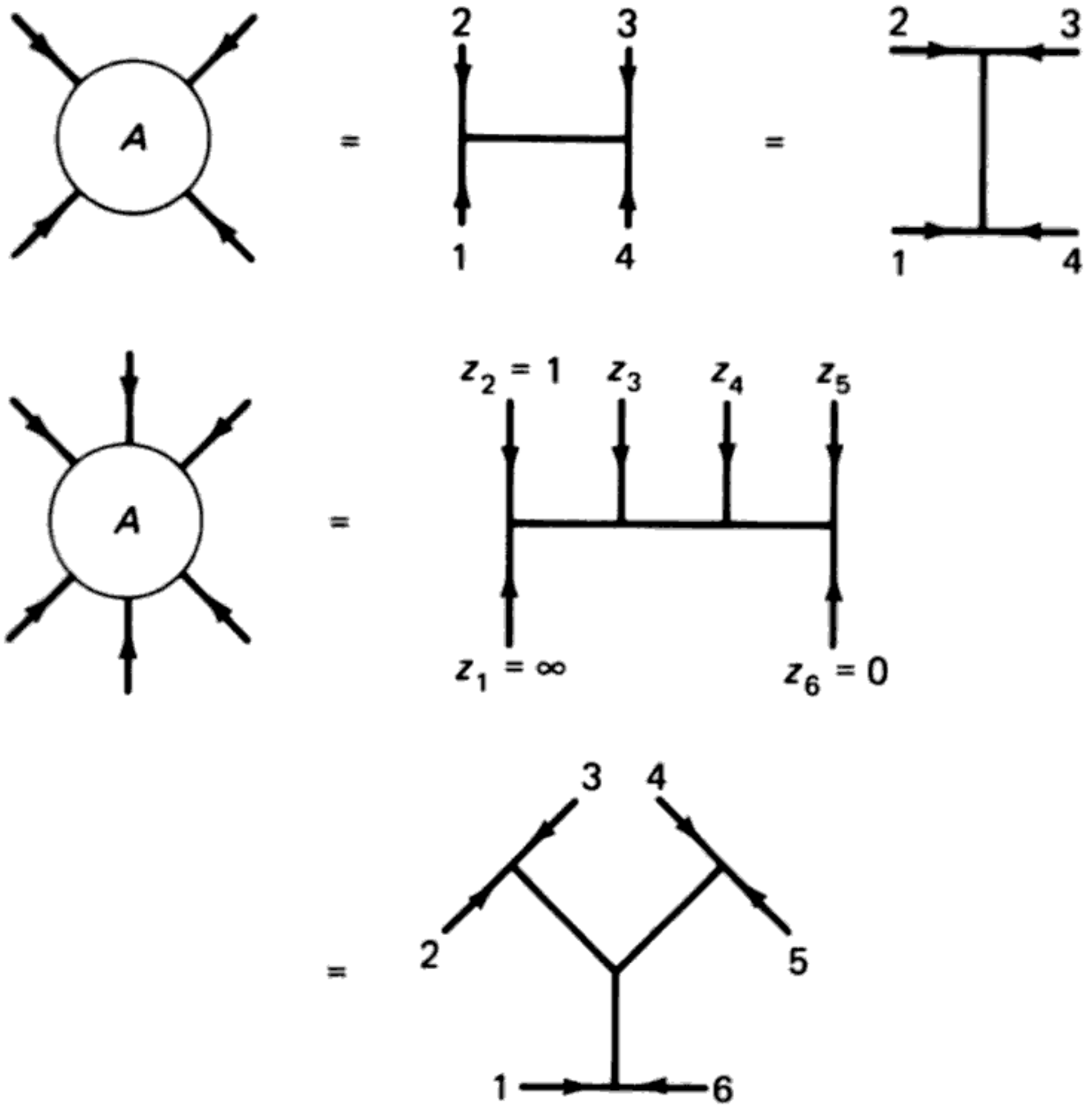

FIGURE 2.5. Duality of the Veneziano model. The four-point amplitude can be decomposed in terms of either s-channel or t-channel poles. This is in contrast to standard point particle field theory, which sums over both s- and t-channel poles. This property of duality extends to the N-point function. For this reason, it was once thought that a field theory of strings was impossible. A field theory of strings would be plagued by overcounting, especially at higher orders.

This expression simplifies considerably if we take the conformal gauge. In particular, we get

$$A_N = \int d\mu \int DX \exp\left\{\frac{-1}{4\pi\alpha'}\int (\partial_z X_\mu \partial_{\bar{z}} X^\mu)\, d^2z + i\sum_{i=1}^{N} k_{i\mu} X^\mu\right\}$$

$$= \sum_{\text{topologies}} \int d\mu \left\langle \prod_{i=1}^{N} e^{ik_i \cdot X} \right\rangle . \tag{2.5.2}$$

In Fig. 2.6 we show how to simplify the string interaction diagram. By letting the string interaction length to go to zero for the external tachyons, we see that the string interaction surface can be reduced to an infinite horizontal strip in the complex plane, extending from $\sigma = 0$ to $\sigma = \pi$ and $\tau = \pm\infty$.

Notice that the action, although it is no longer reparametrization invariant because we have fixed $g_{ab} = \delta_{ab}$, is still conformally invariant. Thus, to avoid overcounting of conformally equivalent surfaces, we will take the set of topologies over which we must sum to be the set of all *conformally inequivalent* two-dimensional complex surfaces.

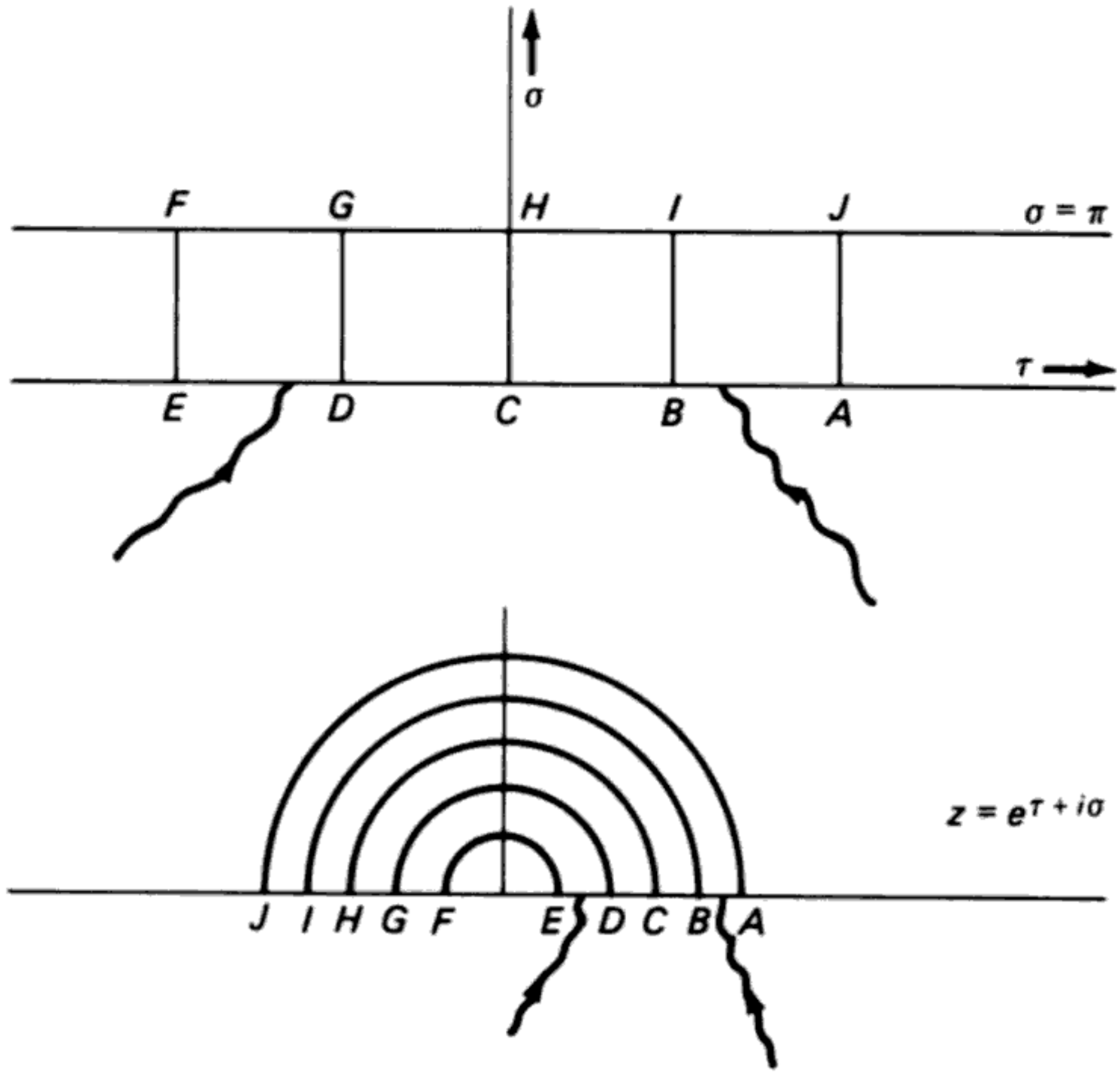

FIGURE 2.6. Conformal surfaces for open string propagation. In the ρ-plane, the surface over which a string propagates is a horizontal strip of width π. The wavy lines at the bottom corresponds to "zero width" strings or external tachyons. In the z-plane, the surface becomes the upper half complex plane. The mapping from one surface to the other is given by the exponential.

Let the world sheet of the N-point tree be a horizontal strip of width π that extends horizontally in the complex plane. The x-axis corresponds to τ and the y-axis corresponds to σ. Let us now change coordinates to complex variables:

$$z = e^{\tau + i\sigma}. \tag{2.5.3}$$

This mapping takes this infinite strip, which describes the world sheet of the interacting string (with zero width tachyons), into the upper half of the complex plane.

Fortunately, the functional integral is a Gaussian that can be evaluated with the identities presented in the previous chapter. Let us define $X_{\text{classical}}$ as the solution to the classical equations of motion:

$$\nabla^2 X_{\mu,\text{cl}} = -2i\pi\alpha' J_\mu, \tag{2.5.4}$$

where

$$J_\mu(z) = \sum_{i=1}^{N} k_{i\mu}\delta(z - z_i), \tag{2.5.5}$$

where z_i are points on the real axis of the complex plane that correspond to the external zero width tachyons interacting with the string. After a Wick rotation

in the τ variable, this is just Poisson's equation for electrostatics. To solve this, we need the Green's function:

$$\nabla^2 G(z, z') = 2\pi\delta(z - z'). \tag{2.5.6}$$

We must calculate the Green's function, with Neumann boundary conditions, for the upper half-plane. The easiest way to calculate this is to borrow a trick from the theory of electrostatics, namely the method of images. Let us place a point charge at the point z' in the upper half-plane. Consider another point charge at the point $\bar{z}'$ that is symmetrically reflected through the x-axis; $\bar{z}'$ is in the lower half-plane. If we are sitting on a point charge z in the upper half-plane, then the potential at that point is proportional to

$$G(z, z') = \ln|z - z'| + \ln|z - \bar{z}'|. \tag{2.5.7}$$

Notice that if we are sitting on the x-axis so that z is real, then the derivative of the Green's function normal to the x-axis is zero. Thus, these boundary conditions are precisely what we want, so (by the uniqueness theorem) this is the Green's function for the upper half-plane.

We can now insert this Green's function back into the integral. The classical value of X that solves (2.5.4) is

$$X_{\text{cl}} = -i\alpha' \int G(z, z')J(z')\,dz.' \tag{2.5.8}$$

Let us now make a shift in the integration variable:

$$X_\mu \to X_{\mu,\text{cl}} + X_\mu. \tag{2.5.9}$$

We find, therefore, that the functional integrals can be performed using (1.7.10):

$$\begin{aligned}
&\int DX \exp\left\{-\frac{1}{4\pi\alpha'}\int \partial_z X_\mu \partial_{\bar{z}} X^\mu\, d^2z + i\int J_\mu X^\mu\, d^2z\right\} \\
&\quad = \exp\left\{\frac{\alpha'}{2}\int J_\mu(z)G(z, z')J^\mu(z')\,dz\,dz'\right\} \\
&\quad = \prod_{i\neq j} \exp\left\{\alpha' k_1 \cdot k_j \ln|z_i - z_j|\right\} \\
&\quad = \prod_{i<j} |z_i - z_j|^{2\alpha' k_i \cdot k_j},
\end{aligned} \tag{2.5.10}$$

where

$$J^\mu(z) = \sum_j \delta(z - z_j)k_j^\mu.$$

Putting everything together in (2.5.11), we find

$$A_N = \int d\mu \prod_{1\le i<j\le N} |z_i - z_j|^{2\alpha' k_i \cdot k_j}. \tag{2.5.11}$$

(Notice that we have explicitly removed the "self-energy" term for $i = j$, which would be divergent. We can truncate the integral and still maintain the conformal properties of the theory. We will find that this truncation will also have to be performed in the harmonic oscillator method.)

Now we must complete the last step, which is to fix the measure $d\mu$.

Our first guess is that, when the amplitude is expressed in terms of z_i, the measure is simply equal to one. This is the correct choice that is compatible with conformal invariance. To prove it, recall that earlier we said we must sum over *all conformally inequivalent surfaces*. Consider the set of conformal transformations that map the upper half-plane into itself, such that the real axis is mapped into itself. In general, the points on the real axis that are mapped into each other transform under a subset of the conformal transformations, the *projective* or Möbius transformations:

$$y' = \frac{ay + b}{cy + d} \tag{2.5.12}$$

for real a, b, c and d such that $ad - bc = 1$. This set of four parameters defines a real matrix with unit determinant:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{2.5.13}$$

In general, the group defined by the set of all real 2×2 matrices that has unit determinant is SL(2, R). Notice that this group of transformations can be generated by making successive transformations:

$$\begin{cases} y \to y + b, \\ y \to ay, \\ y \to \dfrac{1}{y}. \end{cases} \tag{2.5.14}$$

Thus, we wish the amplitude, including the contribution of the measure, to be projectively invariant.

Let us make a projective transformation on the integrand to see how it transforms:

$$\prod_{i<j}(z_i - z_j)^{2\alpha' k_i \cdot k_j} = \prod_{i<j}(z_i' - z_j')^{2\alpha' k_i \cdot k_j} \prod_k (a - cz_k')^2. \tag{2.5.15}$$

We want our measure to cancel out the noninvariant term in the above expression. Let us take our measure to be the number 1 and the integration region to be fixed by $z_i \geq z_{i+1}$. Then there is one last complication. We must still "fix the gauge" for projective transformation or else we will have overcounting. We must integrate once and only once over each projectively distinct configuration of the z_i variables. If the external momentum flows into the upper half-plane at points given by z_i, then we are allowed to fix three of these points at random. This corresponds to "gauge fixing" the projective invariance, which selects out

Superstrings and M-theory are provocative and controversial, but unarguably among the most interesting and active areas of research in current physics. Called by some "the theory of everything," superstrings may solve a problem that has eluded physicists for the past 50 years, the final unification of the two great theories of the twentieth century, general relativity and quantum field theory.

Now, here is a thoroughly revised, second edition of a course-tested comprehensive introductory graduate text on superstrings which stresses the most current areas of interest, not covered in other presentations, including:

- Four-dimensional superstrings
- Kac–Moody algebras
- Teichmüller spaces and Calabi–Yau manifolds
- M-theory membranes and D-branes
- Duality and BPS relations
- Matrix models

The book begins with a simple discussion of point particle theory and uses Feynman path integrals to unify the presentation of superstrings. It has been updated throughout, and three new chapters on M-theory have been added.

Prerequisites are an acquaintance with quantum mechanics and relativity.

ISBN 0-387-98589-1
www.springer-ny.com