PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Strings at Superplanckian Energies: In Search of the String Symmetry [and Discussion]

D. J. Gross and J. R. Ellis

Phil. Trans. R. Soc. Lond. A 1989 **329**, 401-413
doi: 10.1098/rsta.1989.0086

Phil. Trans. R. Soc. Lond. A **329**, 401–413 (1989)
Printed in Great Britain

Strings at superplanckian energies: in search of the string symmetry

By D. J. Gross

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, *U.S.A.*

The characteristic energy scale of superstring theory, which attempts to unify all the interactions of matter with gravity, is the Planck energy of 10^{28} eV. Although this energy is 16 orders of magnitude higher than currently accessible energies, it is important to consider the nature of string physics in this region since it could shed light on the non-perturbative physics at the Planck scale, which determines the structure of the vacuum.

In this paper I review some recent attempts to explore this domain. In particular, I discuss string scattering at very high energies, the indications of the existence of a large symmetry that is restored at short distances and the possible breakdown of our concepts of space-time at these energies.

1. Introduction

We have heard much in this meeting about the technology of two-dimensional conformal field theory, which is useful for the construction of classical vacua for string theory. However, the real problem in string theory is not to find more and more classical solutions, but to find the dynamics that picks the unique vacuum, and thereby to make contact with the real world.

It is not unusual to have more than one classical vacuum in a quantum mechanical system, especially in supersymmetric theories where there are often many 'flat directions' in which one can vary the expectation value of some field without changing the energy of the vacuum. In such a situation one can develop a perturbation theory about each vacuum separately. However, unless perturbative instabilities arise, the correct vacuum involves non-perturbative physics, e.g. tunnelling. The trouble is that in string theory we have a similar situation, but we do not know the lagrangian, we do not even know what the coordinates are. All we have are the classical solutions and perturbation theory about them. All of our understanding of string theory is based on the following two ingredients: a rule for finding the classical vacua (equivalent to classical solutions) of some theory whose lagrangian we do not know, and rules for calculating S-matrix elements in perturbation theory.

This is not good enough. One of the reasons it is not good enough is that perturbation theory diverges very badly. This is a recent result of Gross & Periwal (1988) (our proof is for the bosonic string theory, mainly for technical reasons, but I believe it is true for all string theories). That is, if you construct string scattering amplitudes perturbatively in powers of the coupling (which in principle is determined by the dynamics but in practice, namely in perturbation theory, is a free parameter), then you discover that this perturbation expansion has zero radius of convergence. The coefficients grow faster than $n!$, so not only is it divergent it is also not Borel summable. This is bad for perturbation theory but very good for physics. If it were not the case that perturbation theory diverged then you could sum it, ending up with billions and billions of classical vacua, and the full quantum mechanical perturbations about them. Because these are all consistent theories, order by order in the coupling, you would end up with billions of consistent theories. All of them would contradict observation as all of them have unbroken supersymmetry and all of them have a massless dilaton. So it is good that we are not allowed to sum the series.

The meaning of a perturbation theory that diverges in this way is that the vacuum is unstable. One indication of this is that if you try to sum such a pertubative expansion (say by the Borel technique), you would find that the vacuum has complex energy, which indicates vacuum instability. Such series occur often in quantum mechanics, for example in the case of the double well potential. They also occur in totally consistent physical field theories, such as quantum chromodynamics (QCD). In both of these cases the non-Borel summable divergences are related to the existence of instantons, which are signals of quantum tunnelling that mixes different vacua. Tunnelling effects behave as $\exp[-1/g^2]$, which cannot be reproduced by a perturbative expansion. We regard this divergence as an indication that all of the string solutions that people have found are unstable at the quantum level as a result of non-perturbative effects. We must develop a formalism that allows us to go beyond perturbation theory, find the non-perturbative effects that will pick out a unique vacuum, break supersymmetry, generate a mass for the dilaton and all other good things.

The main trouble in going beyond perturbation theory is that we do not have the analogue of the lagrangian, or a second-quantized hamiltonian. We do not even know what the natural variables of string theory are, nor what the natural description of configuration space is. All we have are these rules for constructing S-matrix amplitudes in perturbation theory. There have been many attempts to go beyond perturbation theory to find the correct basic formulation of the theory. I think it honest to say that none of these suggestions has got very far. A sign of how bad things are is the difficulty in finding non-stationary classical solutions of string theory. An appropriate *euclidean* solution of this variety could be used to explore non-perturbative phenomena semi-classically. Although we have developed much understanding of stationary classical solutions, which are given by conformal field theories, not one non-stationary solution has been constructed!

I think it is very important to keep an open mind about these issues. The most obvious approach such as the construction of *string field theory* might not be correct. It is very instructive to recall the other instance in which strings appear in physics, namely as a description of the large-scale behaviour of QCD, where they describe the approximate structure of hadronic electric flux tubes. The large-distance behaviour of SU_N gauge theory, in the limit of large N, might very well be described by an effective string theory. To leading order in $1/N$ we would have a non-interacting system of strings, whose spectrum would consist of stable hadrons. The coupling of the hadrons would be proportional to $1/\surd N$, and one could develop a perturbative expansion of hadronic scattering amplitudes, much as one does in string theory.

If one had such a formulation of QCD it would be very difficult to extend it to a non-perturbative framework. Much of the underlying simplicity, especially the conceptual simplicity, of the theory would not be evident. An analogue of the problems we now face would be the attempt to discover quarks, gluons and non-abelian gauge theory starting from an effective string theory of QCD. The reason is that the formulation of the theory appropriate for large-scale physics does not reveal the essence of QCD. In QCD, as in most theories, the ultimate simplicity lies at short distances. In this régime the variables in terms of which one describes the system are fewer and simpler and the symmetries of the theory are manifest and evident.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

If this is true for strings as well, then to make progress we must try to understand the high-energy behaviour, or the high-energy phase, of string theory. We must also be ready for the possibility that a totally different kind of structure will be appropriate for the description of this domain; but we may hope that it will prove to be conceptually simpler and more elegant.

Thus we must explore the physics of strings at energies above their characteristic energy scale. In the case of superstring theory, which attempts to unify all the interactions of matter with gravity, this is the Planck energy of 10^{28} eV. Although this energy is 16 orders of magnitude higher than currently accessible energies, it is important to study this domain. The idea is to do what an experimental high-energy physicist would if she had unlimited funds available, namely build an accelerator with the highest possible energy and scatter protons (or gravitons) and look at what comes out. We don't have unlimited funds but we have a theory, or at least a small handle on the theory. So we can do thought (gedanken) experiments, pushing the theory to its limits where it might reveal new structures, or at least new problems. In particular, if the theory possesses a much larger symmetry than the symmetry that is evident in low-energy scattering amplitudes (i.e. supergauge + general coordinate invariance) than it might become evident in the structure of high-energy scattering amplitudes. In this paper I review some recent attempts to explore this domain. In particular I discuss string scattering at very high energies, the indications of the existence of a large symmetry that is restored at short distances and the possible breakdown of our concepts of space-time at these energies.

2. The high-energy behaviour of string scattering amplitudes

I shall describe, briefly, the study that I carried out, with Paul Mende (Gross & Mende 1987, 1988) of the high-energy behaviour of string scattering. From these calculations a few very interesting features emerge. First, the limit of very high energies or, because the only mass scale in the theory is the Planck mass, the zero Planck mass limit, is the *semi-classical* limit of first-quantized string theory. Namely, order by order in the perturbative expansion of string scattering amplitudes, the functional integrals over the space-time surfaces mapped out by the propagating strings are dominated by a saddle point, which corresponds to a classical solution of the string equations of motion on a particular Riemann surface.

The fact that the zero Planck mass limit is a semi-classical limit has interesting implications that I shall discuss. One of them is that it is easy to calculate the high-energy behaviour order by order in perturbation theory. Also, the result has a large amount of universality and is very stringy and unusual. Finally, there appears in the structure of the scattering amplitudes hints of a large and totally mysterious symmetry of string theory that is restored as we take the Planck mass to zero (Gross 1988).

Let us first examine the behaviour of string scattering at large energy in the Born approximation. This is straightforward because the amplitude is given as a ratio of gamma functions and the limit is easily deduced with the aid of Stirling's formula. Consider the elastic scattering of four tachyons, the ground state of the bosonic string. (For simplicity I shall discuss mostly the bosonic string, although most of my remarks hold for the super and heterotic strings as well.) The Virasoro–Shapiro amplitude is given by

$$A_{\text{tree}} = g^2 \frac{\Gamma(-1-\frac{1}{8}s)\,\Gamma(-1-\frac{1}{8}t)\,\Gamma(-1-\frac{1}{8}u)}{\Gamma(2+\frac{1}{8}s)\,\Gamma(2+\frac{1}{8}t)\,\Gamma(2+\frac{1}{8}u)}, \tag{2.1}$$

where s, t, u are the Mandelstam variables $s = -(p_1+p_2)^2$, $t = -(p_1+p_3)^2$, $u = -(p_1+p_4)^2$; $p_i^2 = 8$ and $s+t+u = -32$. For large energy, $s \to \infty$, and fixed s-channel scattering angle ϕ, $\sin^2\frac{1}{2}\phi = -t/(s+32) \approx -t/s$, $\cos^2\frac{1}{2}\phi = -u/(s+32) \approx -u/s$, we have,

$$A_{\text{tree}} = 8\mathrm{i}g^2\,\mathrm{e}^{-8}(stu)^{-3}\exp\left[-\tfrac{1}{4}(s\ln s + t\ln t + u\ln u)\right]$$
$$= \mathrm{i}g^2 2^9 s^{-1}(\sin\phi)^{-6}\exp\{\tfrac{1}{4}(s+32)\,[\sin^2\tfrac{1}{2}\phi\ln\sin^2\tfrac{1}{2}\phi + \cos^2\tfrac{1}{2}\phi\ln\cos^2\tfrac{1}{2}\phi]\}. \qquad (2.2)$$

This exponential fall-off is quite different from that of the power behaviour characteristic of field theory. It is clearly an essential feature of string theory, surely related to the finiteness of this theory, yet its deeper meaning is still unclear. It contradicts the rigorous lower bound of quantum field theory, which states that $|A(s,\cos\phi)| \geqslant \exp[-\sqrt{s}\ln sf(\phi)]$.

The proof of this bound uses unitarity, the existence of a finite mass gap and, most importantly, the assumption of polynomial boundedness in the energy, for fixed t, of the scattering amplitudes. Of course, perturbative string theory amplitudes violate all of these assumptions so there is no contradiction. In quantum field theory polynomial boundedness is a consequence of locality, and the power fall-off at high energies is a consequence of power singularities in the operator product of local fields at light-like separations. In the days of axiomatic field theory people asked themselves what is the worst kind of singularity that one could tolerate in the product of field operators, $\phi(x)\,\phi(y)$, and still have a local theory. For example, if you have any finite sum of delta functions, $(\partial_x^n\,\delta(x-y))$, than you get a local distribution, which vanishes when $x \neq y$. But you can even have an infinite sum as long as the coefficients do not grow too rapidly. When you work out what that means in momentum space it means that you can't have more raid fall-off than $\exp(-\sqrt{s})$. If you violate the lower bound you violate locality.

In string theory we find this crazy exponential behaviour, crazy from the point of view of field theory, and therefore we should be interested because it is a sign that something new is happening at short distances in string theory. There certainly is no way that one could represent this physics by an effective field theory that is truly local. This is one of the most interesting lessons that we learn from the calculation of high-energy string scattering.

Let us now consider the general G-loop amplitude, $A_G(P_i)$, given by the path integral

$$A_G(P_i) = \int \frac{\mathscr{D}g_{\alpha\beta}}{\mathscr{N}}\mathscr{D}X^\mu \exp\left\{-\frac{1}{2\alpha'}\int \mathrm{d}^2\xi\sqrt{g}\,g^{\alpha\beta}\,\partial_\alpha X^\mu\,\partial_\beta X_\mu\right\}\prod V_i(P_i), \qquad (2.3)$$

where the string tension α' is proportional to M_{Planck}^{-2}, V_i is the vertex operator for the emission of the ith particle and the integral is over all compact surfaces of genus G. In this formula it looks like α' plays the role of Planck's constant and $\alpha' \to 0$ is the classical limit. However, we must remember that to discuss the classical limit of a theory one must specify the observables. In our case the observables are the vertex operators of physical states, all of which depend on the external momenta as

$$V(P_i) = \int \mathrm{d}^2\xi_i\sqrt{g}\,\mathrm{e}^{\mathrm{i}P_i X(\xi_i)}\mathscr{V}(P_i, X(\xi_i)),$$

where $\mathscr{V}(P_i, X(\xi_i))$ depends on the quantum numbers of the particle. In the limit of high energy, or $\alpha' \to \infty$, or $M_{\text{Planck}} \to 0$, it is then clear that $X^\mu \approx \alpha'(1/\partial^2)\,P^\mu$. In other words X^μ should be scaled by α', in which case the exponent appears with α' upstairs. Thus $\alpha' \to \infty$, or $M_{\text{Planck}} \to 0$, is the semi-classical limit and in this limit the integral will be dominated by a particular two-dimensional surface, X^μ and $g_{\alpha\beta}$. In the critical dimension the integral is invariant under the group of diffeomorphisms and Weyl rescalings of the metric, which is why we have divided by $\mathcal{N}$, the volume of this group. It can then be reduced to an integral over a slice of the finite-dimensional moduli space $\mathcal{M}_G$ of Riemann surface of genus G (D'Hoker & Phong 1989). Consider, for simplicity, external tachyons whose vertex operators are $\mathcal{V}(P_i) = 1$. The X^μ integral is easily done yielding

$$A_G = g^{2G+2} \int_{\mathcal{M}_G} [\mathrm{d}\boldsymbol{m}] \prod_i \mathrm{d}^2\xi_i \sqrt{g(\xi_i)}\, \Omega(\boldsymbol{m}, \xi_i) \exp\left[-\tfrac{1}{2}\textstyle\sum P_i P_j G_m(\xi_i, \xi_j)\right], \tag{2.4}$$

where $\boldsymbol{m}$ are a set of coordinates for moduli space, $\Omega(\hat{\boldsymbol{m}}, \hat{\xi}_i)$ is the measure on the moduli space and $G_m(\xi_i, \xi_j)$ is the scalar Green function, $\Delta_0 G_m(\xi_i, \xi_j) = -2\pi\delta^2(\xi_i, \xi_j)$, on the surface.

The only place that the momenta enter into this expression is in the exponential. Therefore, in the limit that *all* $P_i P_j$ become large, we can perform the integral by saddle-point techniques. The exponent,

$$\mathcal{E}(P_i, \xi_i, \boldsymbol{m}) \equiv \tfrac{1}{2} \sum_{i<j} P_i P_j G_m(\xi_i, \xi_j), \tag{2.5}$$

can be regarded as the electrostatic energy of two-dimensional Minkowski charges P_i, placed at positions ξ_i on a Riemann surface of genus G with moduli $\boldsymbol{m}$. Thus to determine the amplitude, we must find the values of ξ_i and $\boldsymbol{m}$ for which the action $\mathcal{E}(P_i, \xi_i, \boldsymbol{m})$ is extremal. This is equivalent to the problem of finding the points of electrostatic equilibrium (in general unstable) on an arbitrary two-dimensional surface which is allowed to change its shape at no energy cost! This is a beautiful variational problem. It does not appear to have been investigated, perhaps because it is only for Minkowski charges (where we can set $P_i^2 \approx 0$ compared to $P_i P_j$) that the problem is strictly conformally invariant.

Because this exponential term is common to all string theories, as well as to amplitudes involving other external states, *the dominant large-energy behaviour of string amplitudes will be independent both of the theory and the particular quantum numbers of the scattered particles*. It will also be independent of the nature of the compactification of some of the spatial dimensions, as long as we consider scattering particles that are not wrapped many times around some internal dimension or whose internal momenta are not large. Of course the prefactor will depend to some degree on the theory and on the external particles. This is an important result. It is an indication of the universality of high-energy behaviour, analogous to the universality of the singularities of the operator product expansion in field theory. It opens the way for the systematic study of the high-energy behaviour of amplitudes with any number of external particles of any type. Such an investigation could throw light on what replaces the operator product expansion in string theory.

This behaviour also means that the determination of the asymptotic behaviour of fixed angle, high-energy scattering in string theory is a much simpler problem than in field theory. If we were to try to evaluate the behaviour of field theoretic amplitudes by similar methods (i.e. introduce Schwinger or Feynman parameters, perform the integrals over the momenta and search for a saddle point in the remaining integrals), we would fail. The integrals over the Feynman or Schwinger parameters, the analogue of the moduli for field theory, are dominated by end-point contributions. One can use power counting to estimate the fall-off of field theory graphs in perturbation theory, but the precise coefficients are not easily obtained. In string theory it turns out, as we shall see below, that the saddle points lie in the middle of moduli space, far from the boundaries. For this reason we can ignore the infrared divergences of the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

bosonic string. They arise from points on the boundary of moduli space which do not dominate the high-energy behaviour. In other words, if we were to cut off the infrared instabilities of the bosonic theory by means of a cut-off, the high-energy behaviour would be independent of the cut-off.

In string theory, we will be able to derive a precise form for the asymptotic behaviour of the G-loop amplitude. Not only can we easily find the dominant saddle point, $(\hat{m}, \hat{\xi}_t)$, but we will be able to calculate the measure on moduli space at these points, $\Omega(\hat{m}, \hat{\xi}_t)$, as well as calculate the gaussian fluctuations about the saddle!

Let us redo the tree approximation by considering electrostatics on a sphere. Take a sphere, put four Minkowski charges (whose sum is zero) on the sphere and find the positions of equilibrium. The sphere is conformally the same as the complex plane (including the point at infinity), as can be seen by a stereographic projection that is a conformal transformation. The Green's function on the plane is simply $G(z, z') = \ln|z-z'|$. So the electrostatic energy is very simple. It is just

$$\mathscr{E} = s[\ln|z_1-z_2|+\ln|z_3-z_4|]+t[\ln|z_1-z_3|+\ln|z_2-z_4|]+u[\ln|z_1-z_4|+\ln|z_2-z_3|].$$

Because all spheres are equivalent to the complex plane there are no moduli. So the only thing we have to vary are the positions of the charges. Because of the conformal invariance this problem has an SL$(2, C)$ (Moibus) invariance. The action will be invariant under a transformation $z \rightarrow (az+b)/(cz+d)$, that takes the plane into the plane. The only SL$(2, C)$ invariant combination of four zs is their cross ratio,

$$\lambda = (z_1-z_3)(z_2-z_4)/(z_1-z_4)(z_2-z_3).$$

By using momentum conservation it is trivial to see that, because $s+t+u = 0$,

$$\mathscr{E} = t\ln|\lambda|+u\ln|1-\lambda|,$$

which is easy to extremize. Differentiating with respect to λ we find $t/\lambda = u/(1-\lambda)$. Thus $\lambda = -t/s = \sin^2\frac{1}{2}\phi$. Plugging this back into the action we find that the minimal action is just the expression we derived before by using Stirling's formula. So we reproduce the result for the sphere and we have understood that this expression comes from a saddle point. Similar methods can then be used to discuss an arbitrary Riemann surface with any number of handles.

It turns out that to study the problem of a Riemann surface of large genus it is useful to think of a Riemann surface as a branched cover of the complex plane, or as an *algebraic curve*, $y = f(z)$, where y and z are complex variables. This describes a curve in CP^2, or equivalently y is a meromorphic function of z and describes a branched cover of the complex plane, a Riemann surface.

For example the torus is a two-sheeted cover of CP^1, the complex plane including the point at infinity, branched over four points; i.e. the algebraic curve

$$y^2 = (z-a_1)(z-a_2)/(z-a_3)(z-a_4). \tag{2.6}$$

To construct the torus in this way, take two copies of the complex plane, one over the other. Next, make square root branch cuts on each sheet between a_1 and a_3, and between a_2 and a_4. Then construct 'crossed bridges', identifying the left edge of a cut on one sheet with the right edge of the corresponding cut on the other, and vice versa. The surface is clearly closed and connected, and a simple triangulation shows it has genus 1. What we have done is connect two spheres by bridges to obtain a torus.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

We are interested in an algebraic curve of the same type, an N sheeted cover of the Riemann surface. This is a curve with branch points at a_i, each one of which has N sheets, so that when you go around each of these branch points N times you come back to the original sheet.

The most general curve of this form is

$$y^N = \prod_{i=1}^{L} (z-a_i)^{L_i}. \tag{2.7}$$

As long as the L_is are relatively prime to N then each branch point a_i is of order $N-1$ and is a point common to all N sheets. (Note that a square root branch point is of order one!) The number of branch points is precisely L if we impose that $\sum L_i = 0 \pmod N$, so that the point at infinity is regular. The curve has a Z_N automorphism, the largest generic automorphism of a Riemann surface. It is not difficult to show that the genus of such a curve is given by

$$G = \tfrac{1}{2}(N-1)(L-2). \tag{2.8}$$

I believe that the dominant saddle points, the elastic scattering at high energy, are just such Z_N curves, with four branch points at which we put the four external charges (momenta). A remarkable feature of these curves is that we can easily do electrostatics on them. On a general Riemann surface electrostatics, i.e. calculating the inverse of the laplacian, is a very complicated business. But *electrostatics is trivial on a branched cover of the sphere if the charges are all placed at the branch points*! This is because each sheet is simply a copy of the complex plane, and each sheet sees the same sources because the points where the charges are located are common to all of the sheets. It is easy to see that in this circumstance the electrostatic field will be proportional to the electrostatic field of the four charges on the complex plane and can immediately be written down, no matter how many sheets we have. Let us consider the most general situation of an N-sheeted Riemann surface, where we place charges P_i at the branch points a_i. (There may be more than four branch points and some of the P_i might vanish. We suppress the Minkowski indices.) We claim that the electric field produced by these charges, $\boldsymbol{E}(z) = E_x(z) + \mathrm{i}E_y(z)$, take the same value on all N sheets, and is given by

$$\boldsymbol{E}_N(z) = \frac{1}{N}\boldsymbol{E}_1(z) = \frac{1}{N}\sum_i \frac{P_i}{z-a_i}, \tag{2.9}$$

where $\boldsymbol{E}_1(z)$ is the electric field of the trivial genus 0 case, the sphere.

In a similar fashion we can construct the saddle-point expression for the space-time coordinate X^μ, which is equal in our analogue electrostatic problem to the electrostatic potential. It is therefore given by

$$X^\mu_{cl}(z) = \frac{\mathrm{i}}{N}\sum_{i=1}^{4} P_i^\mu \ln|z-a_i|. \tag{2.10}$$

It is now trivial to calculate the action for the N-sheeted curve considered above. Because

$$\mathscr{E} \propto \int |\boldsymbol{E}_N|^2 = N\int_{\text{sphere}} \left|\frac{1}{N}\boldsymbol{E}_1\right|^2,$$

it follows that

$$\mathscr{E}_N = \mathscr{E}_1/N = -\frac{1}{2N}\sum_{i<j} P_i P_j \ln|a_i - a_j|. \tag{2.11}$$

In the case of four branch points, and four charges located at these branch points, $\mathscr{E}_N$ is proportional to the energy for the sphere. Extremizing with respect to the modulus $\lambda = (a_1-a_3)(a_2-a_4)/(a_1-a_2)(a_3-a_4)$, we find the same result as before, $\lambda = -t/s$, so

$$\mathscr{E}_N = (1/4N)(s\ln s + t\ln t + u\ln y). \tag{2.12}$$

One problem that remains open is the extension of these results to the many-particle case. I do not even know how to solve the problem of multiparticle high-energy scattering in the Born approximation. The corresponding *electrostatic problem* is to find the equilibrium positions of M Minkowski null charges, P_i, which sit on the sphere. This problem is very straightforward to formulate, is a very pretty problem, but the general solution is unknown. I present it here as a challenge.

The equations one must solve are the following:

$$\sum_{j\neq i} \frac{P_i P_j}{z_j - z_i} = 0, \quad i = 1, \ldots, M, \tag{2.13}$$

which express the vanishing of the force on the ith particle, located at z_i. This equation can be written in many different ways. For example an equivalent condition is that the electric field $E^\mu(z) = \sum_i P_i^\mu/(z-z_i)$ satisfy

$$E^\mu E_\mu = 0. \tag{2.14}$$

This expresses the fact that the induced metric on the string world-sheet is conformal

$$g_{zz} = \partial_z X_\mu \partial_z X^\mu = 0, \quad g_{\bar{z}\bar{z}} = \partial_{\bar{z}} X_\mu \partial_{\bar{z}} X^\mu = 0. \tag{2.15}$$

I know of one solution to the multicharge problem, which exists whenever all M momenta lie within a four-dimensional subspace. In this case the position of the ith charge is independent of the value of all the other charges and is given by

$$z_i = \frac{P_i^1 + \mathrm{i}P_i^2}{P_i^0 + P_i^3} = \frac{P_i^0 - P_i^3}{P_i^1 - \mathrm{i}P_i^2}. \tag{2.16}$$

This solution has a very nice geometric interpretation, which I will discuss elsewhere (Gross 1989)†. The action for this solution is proportional to $\sum_{i\neq j} P_i P_j \ln(P_i P_j)$. The general solution for M particles, for $M > 5$ when generically the particles lie outside of four dimensions, is unknown.

Calculating the classical action is only the beginning of a semi-classical calculation. One must also calculate the measure on moduli space at these points, $\Omega(\hat{m}, \hat{\xi}_i)$, as well as calculate the gaussian fluctuations about the saddle. Remarkably, using many techniques of conformal field theory, it turns out to be possible to explicitly do these calculations (Gross & Mende 1987). The result is

$$A_G = \mathrm{i}^G 2^{23G+8}\pi^{9G+3}(G+1)^{9G+3} \times g^{2G+2}\,\mathrm{e}^{-(s\ln s + t\ln t + u\ln u)/4(G+1)}$$
$$\times (stu)^{-3G^2+4G+9/3G+3} |\lambda(1-\lambda)|^{-6G^2+14/3G+3}$$
$$\times \prod_{k=1}^{G} [J_k(\lambda)\, J_k(1-\lambda)]^{-\frac{1}{2}} \left[\left(\sin\frac{\pi k}{N}\right)^{-1} F_k(\lambda)\, F_k(1-\lambda)\right]^{-13}. \tag{2.17}$$

† I have been informed by David Fairlie that this solution was known to him since 1972.

Here, $G = N-1$, $\lambda = -s/t$, $F_k(\lambda) = F(k/N, 1-k/N; 1; \lambda)$,

$J_k(\lambda) = C_k[(C_k - \frac{1}{2}\lambda)(C_k - \frac{1}{2}(1-\lambda) - (\frac{1}{2}-k/N)^2\lambda(1-\lambda)]$, and $C_k = \lambda(1-\lambda)(\mathrm{d}/\mathrm{d}\lambda) F_k(\lambda)$.

Note that this series is very divergent. This makes it unreasonable to try to sum it. Instead we should try, as I shall below, to deduce properties of the series that are true to all orders.

3. High-energy symmetries

It is often the case that spontaneously broken symmetries of a physical theory are hard to recognize at low energy, but become evident in the high-energy behaviour of the theory. Thus the broken $SU_2 \times U_1$ symmetry of the electroweak interactions can be seen by examining weak scattering amplitudes at energies high enough that the W and Z_0 masses can be neglected. String theory surely possesses a very rich symmetry, as suggested by its incredible degree of uniqueness; however, this symmetry is little understood. Presumably this is because most of the string symmetry is spontaneously broken in the known ground states, leaving only the familiar gauge symmetries unbroken and manifest. Perhaps all the string states are gauge particles, but most are massive because of spontaneous symmetry breaking. Perhaps at very high energies, so high that the Planck mass (which in string theory is proportional to the string tension $T = 1/\pi\alpha'$) can be neglected, the full symmetry of string theory is restored. Perhaps we can discover this symmetry (if it exists) by studying string theory in the high energy (or $\alpha' \to \infty$) limit. In particular, if at high energies a larger symmetry is restored there should then exist linear relations between the scattering amplitudes that should be valid order by order in perturbation theory. If so these might be discoverable by analysing the high-energy behaviour of the theory perturbatively.

We can use the results described above to search, and to find, such relations. We are now interested in the dependence of the high-energy scattering amplitudes on the quantum numbers of the external particles. Recall the expression, (2.3), for the amplitude where *the only factor that depends on the nature of the particles being scattered is* $\prod_i \mathscr{V}_i$! Therefore, if we can determine the saddle-point surface, in other words the $\hat{m}_i$, ξ_i and X^{μ}_{cl}, which only depend on the number of particles participating in the scattering but not on their quantum numbers, then we can immediately deduce a linear relation between any two scattering amplitudes involving the same number of particles with the same momenta. For example, in the case of the scattering of particles a_i and particles b_i

$$A^{a_i}_G(P_i) = \prod_i \mathscr{V}_{a_i}(X^{\mu}_{cl}, P_i) / \prod_i \mathscr{V}_{b_i}(X^{\mu}_{cl}, P_i)\, A^{b_i}_G(P_i)\,(1 + O(\alpha')). \tag{3.1}$$

This is an amusing relation, but one of little use by itself as it only relates amplitudes at a given order of perturbation theory. Fortunately, for the dominant saddle points discussed above the dependence of X^{μ}_{cl} on the order of perturbation theory, G, is trivial. All of these surfaces (including those with different L_i) give rise to the same exponential factor in (2.3), all of them have the same Green's function and for all of them X^{μ} is given by (2.10).

It is a remarkable fact that the saddle-point surfaces are all identical, except for the scale (which goes as $1/N$), for all the saddle points in each order of perturbation theory. Consequently, the only genus dependence present in equation (3.1) comes from the $(1/N)$s in the X^{μ}_{cl}s present in the $\mathscr{V}_i$s, namely $\mathscr{V}(X^{\mu N}_{cl}, P) = \mathscr{V}(X^{\mu 1}_{cl}/N, P)$. If all the $\mathscr{V}_{a_i}$s were homogeneous functions of the X^μs of the same degree, then (3.1) would be independent of G. Unfortunately, this is not the case, except for the vertex operators of the simplest string states. However, the factors of $1/N$ can be replaced by derivatives with respect to the momenta. This is because the operator $\mathscr{D} \equiv 2/\alpha' S \sum_i P_i \cdot \partial/\partial P_i$ brings down a factor of N, when acting on $e^{-\frac{1}{4}\alpha' S}$, where $S = s \ln s + t \ln t + u \ln u$. When acting on the other polynomial terms in the momenta $\mathscr{D}$ is of order $1/\alpha'$. Therefore we can replace the $1/N$ by $\mathscr{D}$. Thus $\mathscr{V}(X_{cl}^{\mu N}, P) = \mathscr{V}(X_{cl}^{\mu 1} \mathscr{D}, P)$, when acting on $A_{N-1}(P_i)$, to leading order in $1/\alpha'$.

Using this trick we can write down linear relations between any two four-particle scattering amplitudes;

$$\prod_i \mathscr{V}_{b_i}(X_{cl}^\mu \mathscr{D}, P_i)\, A_G^{a_i}(P_i) = \prod_i \mathscr{V}_{a_i}(X_{cl}^\mu \mathscr{D}, P_i)\, A_G^{b_i}(P_i)\,(1 + O(\alpha')), \tag{3.2}$$

which are independent of G! We can now argue that this relation should hold for the complete amplitude, in the $\alpha' \to \infty$ limit, because it holds order by order in perturbation theory. This is a traditional method of proving properties of quantum field theory (such as the operator product expansion, symmetries, the renormalization group, etc.) by establishing their validity order by order in perturbation theory. It is not, however, without dangers, especially in proving asymptotic theorems. None the less we shall make this assumption. Then, because as $\alpha' \to \infty$ all the particles become massless, we can relate the scattering amplitudes for any set of particles. Indeed all the four-particle scattering amplitudes can be expressed in terms of, say, the four-tachyon amplitude $A_{\text{tachyon}}(P_i)$:

$$A_{a_i}(P_i) = \prod_i \mathscr{V}_{a_i}(X_{cl}^\mu \mathscr{D}, P_i)\, A_{\text{tachyon}}(P_i)\,(1 + O(\alpha')). \tag{3.3}$$

This is a remarkable result, which hints at a very large and unusual symmetry of string theory that might be restored at high energies. Similar relations could surely be derived for the superstring and could also be extended to multiparticle amplitudes as most of the above story goes through for these. If this is the case then the full S-matrix of the $\alpha' \to \infty$ limit of string theory (say for the heterotic string) could be expressed in terms of the dilaton S-matrix. One could even contemplate constructing explicitly the $\alpha' \to \infty$ limit of the theory by plugging these relations into the unitarity equations (which according to the arguments of Gross & Mende (1988) might be valid for the $\alpha' \to \infty$ limit because the high-energy behaviour is dominated by amplitudes in which all internal momentum transfers and energies are large), and using these to solve for the dilaton amplitudes and thereby the full $\alpha' = \infty$ theory. This approach is made quite complicated by the accumulation of an infinite number of particles at zero mass as $\alpha' \to \infty$.

The above relations connect amplitudes involving particles of different and arbitrarily high spin. If they are generated by a symmetry transformation of the $\alpha' \to \infty$ theory it must be one whose conserved charges have arbitrarily high spins. This would contradict the Coleman–Mandula theorem (Coleman & Mandula 1967), which limits the maximal spin of a conserved charge to be one. Perhaps the Coleman–Mandula theorem is invalid for the $\alpha' \to \infty$ limit of string theory. One of the assumptions of the theorem is the *particle-finiteness* assumption (Coleman & Mandula 1967), which states that for any finite M there are only a finite number of particles with mass less than M. For $\alpha' = \infty$ there are an infinite number of massless particles, in which case the theorem need not apply.

It might also be the case that the Coleman–Mandula theorem is valid. The theorem states that if there exist higher spin conserved charges than the S-matrix equals the identity and the theory is trivial. This is because a higher spin conserved charge, $Q_{\mu_1, \mu_2, \mu_3, \ldots}$, would take values equal to $\sum_i P^i_{\mu_1} P^i_{\mu_2} P^i_{\mu_3} \ldots$ on asymptotic states of spinless particles of momenta $P^1, P^2, P^3 \ldots$. The only way this can be conserved is if all the individual momenta are conserved, i.e. only forward and backward scattering is allowed. Then, if one accepts the usual analyticity of scattering amplitudes, no scattering at all is allowed (in more than two space-time dimensions). Perhaps this theorem is valid, the higher spin symmetries do exist and consequently the scattering amplitudes do all vanish as $\alpha' \rightarrow \infty$. This is certainly suggested by the exponential fall-off of the scattering amplitudes, order by order in perturbation theory, as in (2.3), as $\alpha' \rightarrow \infty$.

What does it mean to say that the high-energy phase of the theory has vanishing S-matrix? It clearly implies the existence of an enormous symmetry of the theory, since the individual momenta of each string mode are separately conserved. From this point of view the maximal symmetry of the theory is so powerful as to render the theory trivial. *The non-trivial dynamics of the string is then a consequence of symmetry breaking.*

This idea bears some resemblance to Witten's ideas concerning the high-energy phase of gravity, wherein the metric vanishes, physics becomes ultralocal and only topological observables are measurable.

In this case the relations (3.2), although of the form 0/0, would still have much content. They might be expressions of the remaining symmetry in the presence of a non-vanishing M_{Planck}. Or they might be consequences of the simple nature of the symmetry breaking interactions.

In any case these relations hint at a very unfamiliar and new kind of symmetry which, to date, we do not understand. This is perhaps not too surprising as it is very likely that the high-energy phase of string theory is a very unfamiliar phase. Indeed I shall argue that at high energies our standard picture of space and time must break down.

4. Constructing a string microscope

The traditional method of exploring short distances is to scatter particles at ever increasing energies. Our most energetic accelerators function as our most sensitive microscopes. This is because the obstacle to perfect spatial resolution is quantum mechanical fuzziness, as expressed by the Heisenberg uncertainty principle; $\Delta X \approx \hbar/P$. As the momentum of our beams increases we can resolve shorter and shorter distances and probe the very concept of the space-time continuum. With the aid of existing high-energy accelerators we have explored distances down to 10^{-16} cm with no sign of any breakdown in our concepts of space and time. If we were to probe down to distances of order the Planck length, however, it is very likely that some modification of current concepts will be necessary. Many people, in considering the notion of the space-time continuum in the context of quantum gravity where the metric itself fluctuates, have speculated about notions of 'space-time foam', which might alter the character of space-time at the Planck length, or advanced the idea that there is a cut-off (such as a lattice cut-off) below which the continuum loses its meaning.

We now possess a consistent theory of quantum gravity, string theory, in which this question can be addressed in a quantitative fashion. With the aid of the semi-classical methods developed here I shall attempt to explore this issue, to construct a string microscope and explore whether space and time, as defined by operational stringy methods, remains a valid concept to arbitrarily 'short distances'. String theory possesses no cut-off at short distances, Poincaré invariance does not break down and we therefore may contemplate experiments at arbitrarily high energies. None the less, I shall argue that, because of the non-local nature of stringy probes, there is a minimal distance, of order the Planck length, below which we cannot actually resolve distances.

In scattering experiments we do not measure distances directly. We measure only momenta from which we indirectly infer the spatial structure of the scattering event. Thus, tests of the nature of space-time inevitably involve much theoretical analysis. The claim that we have tested the point-like nature of quarks and leptons to distances of order $l \approx 10^{-16}$ cm is based on arguments that if the *theory* is modified at momenta of order Λ then experiment requires that Λ be greater than $\hbar/l$. I do not know of a direct way to tell whether string theory will truly require a modification of the notions of space and time at short distances.

Let us use the semi-classical description of string scattering at high energies to examine whether strings can be used to construct a microscope that would allow us to probe arbitrarily short distances. This is straightforward because we have seen that the scattering in this limit is dominated by a classical trajectory, so we can imagine translating our analysis directly into a space-time picture. Of course the analysis that I reviewed was based on perturbation theory. We therefore must pretend that this perturbation theory is a good approximation. This is tricky because I argued above that the expansion is always divergent. Nevertheless, for the perturbative vacua about which we are expanding, the coupling, g, which is proportional to the gauge or gravitational coupling, is a free parameter. In reality it is given by the expectation value of a dynamical field (the dilaton), but this is not fixed in perturbation theory. Therefore let us take it to be very, very small so that we can trust perturbation theory, at least in the sense of an asymptotic expansion.

The high-energy scattering of the strings was described by a specific surface given by equation (2.10), where z runs over the Z_N surface described previously. It is not too difficult to picture the scattering event described by this formula. For this purpose it is useful to choose a *gauge* where $X^0(\sigma, \tau) = \tau$ and to describe the trajectories $X^i(\sigma)$ as a function of the time τ. For present purposes it is enough to note that the size of the strings at the time of collision, which defines the size of the interaction region or the size of the distances that we can explore using these probes, grows with increasing energy as

$$X_{\text{coll}} \approx \alpha' E/N, \tag{4.1}$$

where E is a characteristic energy of the scattering. This is unfortunate, given that we are trying to use strings as local probes, because we find that as we increase the energy to probe shorter distances, the effective size of the strings increases.

There is one problem with this conclusion, namely the saddlepoint trajectory is imaginary (i.e. note the i in (2.10)). What does this mean? What it appears to mean is that in the high-energy limit of the first-quantized theory the scattering is an exponentially suppressed process that can only be calculated by analytically continuing to imaginary time. This is not a tunnelling process, after all the scattering is certainly allowed as a real process. (For example the velocity flow $\mathrm{d}X^r(\sigma,\tau)/\mathrm{d}X^0(\sigma,\tau)$ is real and physical.) It is analogous to a particle that backscatters, with an exponentially small amplitude, at energies way above a potential barrier.

This remark might be related to our previous conjecture that the symmetry revealed at high energies is the maximal symmetry of no interaction. Indeed, the only way the strings scatter is by a classically unallowed process that we obtain by analytical continuation to imaginary X^μ!

None the less, even if the trajectory described above is not the real Minkowski trajectory, it should describe averages of the real trajectories and thus the average size of the interaction region. At low energies (compared with the Planck mass) we gain in spatial resolution as we increase the energy of our microscope. However, it appears that when we get to the Planck length things change. The fact that our probes themselves are non-local strings becomes relevant. In fact the strings themselves expand with increasing energy. This contributes to the fuzziness with which we can resolve distances, so that the total fuzziness is

$$\Delta X \approx \hbar c/E + G_N E/g^2 c^5. \tag{4.2}$$

Here we have used the fact that (in the heterotic string theory) the string scale (set by α') and the Planck scale are related by g^2. (We can think of g^2 as the fine structure constant and G_N is Newton's constant of gravity.) This is fortunate, as it means that with our assumption of small g^2, the strings never get within their Schwarzchild radius. Thus we do not have to come to grips with strong gravitational effects. The above smearing is not caused by strong gravity, rather it results from the non-local nature of string probes. It implies that the minimal length that can be probed with strings is of order $1/M_{\text{Planck}}$! If this argument is valid then space-time could have no physical meaning below the Planck length.

I expect that in the final formulation of string theory that space and time will emerge only as an approximate concept, valid or useful for certain approximations to the theory. Thus, for example, many people believe that the proper description of the configuration space of string theory is something like the space of all (cut-off) two-dimensional field theories. Conformal field theories are just the classical solutions of the theory. For these we have a space-time interpretation of the perturbative expansion of string scattering amplitudes about these classical vacua. However, for a general two-dimensional field theory there is no such interpretation. (If this is the case then euclidean gravity is probably on shakier grounds than one might otherwise suspect, because it does not make much sense to analytically continue an approximation.) In such a formalism the space-time manifold is not a primary concept, rather it is a useful description of string physics at low energies and weak coupling (where, presumably, perturbation theory is justified).

Further elaboration of these ideas, along with the practical applications of string theory, awaits the development of a non-perturbative formulation of the theory.

This research was supported in part by NSF grant PHY80-19754.

References

Coleman, S. & Mandula, J. 1967 *Phys. Rev.* **159**, 1251.
D'Hoker, E. & Phong, D. H. 1989 Princeton preprint PUPT-1054. *Rev. mod. Phys.* (In the press.)
Gross, D. J. 1988 *Phys. Rev. Lett.* **60**, 1229.
Gross, D. J. 1989 Kniznik Memorial volume. (In the press.)
Gross, D. J. & Mende, P. F. 1987 *Phys. Lett.* B **197**, 129.
Gross, D. J. & Mende, P. F. 1988 *Nucl. Phys.* B **303**, 407.
Gross, D. J. & Periwal, V. 1988 *Phys. Rev. Lett.* **60**, 2105.

Discussion

J. R. Ellis, F.R.S. (*Theory Division, CERN, Geneva, Switzerland*). The contributions to fixed-angle scattering that Professor Gross discussed fall off more slowly with higher energy when one goes to higher genus. Does this mean that the true behaviour of the theory in this limit is controlled by some 'infinite-genus' configurations that we do not understand?

D. J. Gross. Perhaps; it is very difficult to make sense out of a sum of terms that behaves as $g^n n!\, e^{-A/n}$ in nth order. Were it not for the $n!$ the sum could converge (if g is small enough), but the ubiquitous $n!$ renders the sum divergent. For this reason I have tried to extract from the study of the high-energy behaviour consequences that can be derived order by order, and do not require summing all orders of perturbation theory.