The Strength of Nonperturbative Effects in String Theory

Stephen H. Shenker
Department of Physics and Astronomy
Rutgers University
Piscataway, NJ 08855-0849

We argue that the leading weak coupling nonperturbative effects in closed string theories should be of order $e^{\exp(-C/\alpha)}$ where $\alpha$ is the closed string coupling constant. This is the case in the exactly soluble matrix models. These effects are in principle much larger than the $e^{\exp(-C/\alpha^2)}$ effects typical of the low energy field theory. We argue that this behavior should be generic in string theory because string perturbation theory generically behaves like $(2\pi)^g$ at genus $g$.

August 1990

* Presented at the Cargèse Workshop on Random Surfaces, Quantum Gravity and Strings, May 28-June 1, 1990.

Nonperturbative effects are crucial ingredients in any attempt to describe the real world by string theory. Vacuum selection, supersymmetry breaking, and the vanishing of the cosmological constant are all examples of issues that must be addressed in an intrinsically nonperturbative way. As of yet there has been no real progress in understanding such phenomena in the critical superstring from a fully string theoretic point of view. The theory has not even been formulated nonperturbatively.

In the last year, though, substantial progress has been made in understanding nonperturbative phenomena in simple models of string theory corresponding to string propagation in less than [1] [2] or equal to one [3] dimension. These systems are formulated in a nonperturbative way as matrix models [4] and have been shown to be related to topological field theory [5].

Such simple models have played an important role in theoretical physics. One need only remember how much was learned about scaling, universality, and the operator product expansion from the two dimensional Ising model, and about confinement and the $U(1)$ problem from the Schwinger model. In order to generalize from these models it was crucial to identify which of their properties were generic, and not a consequence of the special simplicity that allowed them to be solved. The purpose of this note is to point out one such property in these simple string theories.

These simple string theories all have leading weak coupling nonperturbative effects of magnitude $e^{\exp(-C/\alpha)}$ where $C$ is a numerical constant and $\alpha$ is the closed string coupling constant, i.e., genus $g$ amplitudes carry a factor of $\alpha^{2g-2}$. It is this property, we claim, that is generic in string theory. Note that the size of these effects is in principle much larger than one would expect from a low energy field theoretic approximation where leading nonperturbative effects would have the characteristic $e^{\exp(-C/\alpha^2)}$ form.

We begin by reviewing nonperturbative phenomena in the matrix models. In the original exact solution of the one matrix model [1] the specific heat in the properly scaled continuum limit was expressed as a solution of a nonlinear ordinary differential equation—the string equation. Every derivative in the string equation is accompanied a factor of

It may well be more correct to interpret the Liouville field as a dimension and think of these theories as living in less than or equal to two dimensions.
\( \kappa \), the scaled version of \( 1/N \). For small \( \kappa \) the leading nonperturbative effects can be found by linearizing the equation around a reference \( \kappa = 0 \) algebraic solution. Because each derivative carries a \( \kappa \), a WKB solution to the linear problem will be of the form
\[ \exp(-f(x)/\kappa) \], displaying the \( \kappa \) dependence described above.

For example, the specific heat \( q(x = 1) \) of the \( m = 2 \) pure gravity one matrix model with even potential is described by the Painlevé I equation
\[ u'' - \frac{\kappa^2}{3} u = z. \]  
(1)

Linearizing around the algebraic behavior \( u = x^{1/3} \) for \( x > 0 \), we find a WKB solution to the homogeneous equation whose exponential dependence is of the form
\[ u_{1m} \sim \exp\left(-\frac{4\sqrt{\kappa}}{5\kappa} x^{3/5}\right). \]  
(2)

The small imaginary part of the "triplly truncated" solution to (1) that David [6] [7] showed describes the analytically continued \( m = 2 \) integral will be of this form. This exponentially small imaginary part reflects the nonperturbative instability of the model due to its unbounded potential.

Another example of leading nonperturbative effects in the one matrix models occurs in the flow [8] from the well defined \( m = 3 \) theory [9] to the \( m = 2 \) theory. The string equation here is
\[ u'' - \kappa^2 uu'' - \frac{\kappa^2}{2} (u')^2 + \frac{\kappa^4}{10} u''' + T_2 [u'' - \frac{\kappa^2}{3} u'] = z. \]  
(3)

\( T_2 \) is the scaling field describing the flow to the \( m = 2 \) theory. The \( \kappa = 0 \) algebraic equation is
\[ u'' + T_2 u'' = z. \]  
(4)

The matrix integral tells us that the correct solution to expand around is the purely real solution that has a jump discontinuity at \( z = \frac{4}{27} T_2^2 = z_0 \). When \( \kappa \) is finite this jump gets smoothed into a nonperturbatively sharp boundary layer that is another signature of the instability of the \( m = 2 \) theory. The leading exponential precursor to the boundary layer (\( z > z_0 \)) is given by linearizing in the neighborhood of the jump and is of the form
\[ \exp\left(-\frac{c(z - z_0)}{\kappa}\right), \quad c = \left(10 - \sqrt{10} T_2/3\right)^\frac{1}{\kappa}. \]  
(5)

The matrix integral provides us with a natural explanation for leading nonperturbative effects in terms of auxiliary saddle points. In terms of eigenvalues the one matrix integral is
\[ \int d\lambda_1 \ldots d\lambda_N \exp(-S) \]  
(6)

where \( S \) is given by
\[ S = -\frac{1}{2} \sum_{ij} \log(\lambda_i - \lambda_j)^2 + \sum_{i} N V(\lambda_i/\sqrt{N}). \]  
(7)

The \( m = 2 \) critical point which we will discuss first can be realized with the simple potential \( V(\lambda) = \lambda^3/2 - a \lambda^4 \). Each sum in (7) is of order \( N^2 \) and so at large \( N \) a saddle point exists [10]. The lowest action and so perturbative saddle corresponds to all the eigenvalues in the well around \( \lambda = 0 \). This saddle is only locally stable; the eigenvalues would rather be at \( \infty \), the instability discussed above.

The lowest action saddle that describes eigenvalues leaving the well is made by moving just one eigenvalue to the local maximum in the effective potential formed from \( V \) and the coulomb repulsion of the remaining eigenvalues (whose positions are essentially unchanged for large \( N \)). The key point is that this saddle corresponds to the movement of just one eigenvalue out of \( N \) so its change in action is proportional to \( N \), not \( N^3 \). This means that the imaginary part of the integral is proportional to \( e^{-CN} \) and not the \( e^{-CN^2} \) we might have expected. In the scaled continuum limit \( N \) becomes \( 1/\kappa \) and the action of this saddle should become universal. The numerical value of this action should just match the value in (2) with \( z = 1 \). David [7] has recently verified that this is the case.

In the flow from \( m = 3 \) to \( m = 2 \), the effective potential for the last eigenvalue develops a secondary minimum for \( \kappa \) near \( z_0 \) as \( T_2 \) is turned on. This minimum is above the eigenvalue filling level as long as \( \kappa > z_0 \). The leading nonperturbative correction in this region (5) should correspond to one eigenvalue occupying the secondary minimum.

These one eigenvalue saddle points are simple examples of string instantons. Because they involve motion of only one out of \( N \) eigenvalues they have anomalously low action and so produce anomalously large effects.

The \( D = 1 \) model [3] also has \( \exp(-C/\kappa) \) nonperturbative effects. This model is formulated as one dimensional quantum mechanics of \( N \) noninteracting fermions with \( \Lambda \)
equal to $1/N$ in a potential $V$ like the one discussed above. The leading nonperturbative effects here correspond, as noted in the original papers, to tunneling out of the metastable well. Tunneling effects go like $\exp(-C/\hbar) = \exp(-CN)$. Here $C$ is the barrier penetration factor or instanton action. Again, the crucial point is that leading effects come from just one eigenvalue tunneling so the instanton action is not proportional to $N$. In the properly scaled continuum limit the tunneling effects for the eigenvalues at the top of the fermi sea become of order $\exp(-C/\kappa)$.

A particularly interesting one dimensional model is the one formulated by Marinari and Parisi [11] that describes a string propagating in one super dimension. Its matrix model realization is one dimensional supersymmetric matrix quantum mechanics. For a cubic superpotential the Hamiltonian in the $0$ fermion number sector becomes a standard $N$ decoupled eigenvalue problem (where the eigenvalues are to be viewed as fermions of a different kind) in a fourth order potential. At the critical point ($\alpha = \alpha_c = 0$) around which the model scales the potential has a cubic inflection point. The region near the inflection point dominates the scaling limit so we can model the potential as $V(\alpha) = \lambda^3 - \alpha$. A small secondary minimum exists for $\alpha > 0$. The leading nonperturbative effects in this model, including supersymmetry breaking, are presumably driven by instantons. The fermi level at $N = \infty$ is at the secondary minimum, and so the appropriate instanton describing the behavior of the top eigenvalue (whose effect will be leading) is the one beginning from the fermi level in the main well and ending at the secondary minimum [12]. Its action is (for $H = \hat{p}^2 + V$)

$$S_{\text{instanton}} = \frac{2}{5} - \frac{3}{4}N \kappa = \frac{2}{5} - \frac{3}{4} \hat{\kappa}$$

where $\kappa = \alpha^{1/4}N = \kappa$ is the $m = 2$ scaled coupling constant appropriate to this model. Instanton effects will be $\exp(-C/\kappa)$ here.

At this workshop Parisi [14] has presented a calculation of supersymmetry breaking. The nonzero vacuum energy is just the escape rate in the Langevin evolution of eigenvalues.

In a recent interesting preprint Karliner and Migdal [13] have given an extensive analysis of the Marinari-Parisi model. They have shown by a rescaling of $H$ with cubic plus linear potential that the model scales for arbitrary coupling and have demonstrated how to use the Gelfand-Diki differential equation for the resolvent to determine its behavior.

in the $m = 2$ one matrix potential given by (7). The escape rate is dominated by the action of the one eigenvalue saddlepoint discussed above and so the exponential dependence of the vacuum energy is just given by (2) with $\kappa = 1$. The original universality arguments for the $m = 2$ string equation establish the universality of this effect. Normalizing coupling constants, we find that $\hat{\kappa}$ in (8) is related to $\kappa$ in (2) by $2\hat{\kappa} = 3\kappa$. Comparing (2) and (8) we see that the nonvanishing of the vacuum energy is a two instanton effect, as is standard in supersymmetric theories.

Another signature of the $\exp(-C/\kappa)$ effects in these models is the large order behavior of their perturbation theory. Writing the perturbation expansion of, say, the free energy $F$ as $e^{F} = \sum_{k} e^{2k} A_{2k}$, these models all have asymptotic behavior

$$A_{2k} \sim (2\kappa)^{2k}$$

as $g \to \infty$. The relation of this to nonperturbative effects is perhaps most simply described by the Borel transform

$$B(t) = \sum_{k} t^{k} (\frac{A_{2k}}{2\kappa})^{k}$$

and its inverse

$$e^{F}(\kappa) = \int_{0}^{\infty} dt e^{-t} B(t).$$

Singualarities in the Borel $t$ plane on the positive real axis create ambiguities in the reconstruction of $F$ from $B$. Singularities at $t_{0}$ are related to nonperturbative effects in the physical quantity $F$ of magnitude $e^{-t_{0}}$. The large order behavior (9) implies that the nearest singularity to the origin is at $|t_{0}| = \kappa$ giving nonperturbative effects of size $\exp(-C/\kappa)$. This presumably sets the scale for other singularities in $B$ at $|t| \sim 1/\kappa$. (Of course there can be other singularities much further away from the origin giving rise to much weaker nonperturbative effects).
This situation should be contrasted with that in field theory where in the loop expansion with $\alpha^2$ as the loop counting parameter the perturbation series $\sum_n \alpha^{2n} A_n$ typically has asymptotics $A_n \sim C^{\alpha^2} \ell$ (not $2\ell!$). The appropriate Borel transform is $B(n) = \sum_n \alpha^{2n} A_n / \ell^n$. It has a leading singularity at $|\alpha| = C/\alpha^2$ producing the usual $\exp(-C/\alpha^2)$ nonperturbative effects of field theory.

It is well known that large order behavior can be described by instanton techniques. The one eigenvalue instanton discussed above are the source of this $(2\pi)^t$. The connection between supersymmetry breaking in the Marinari-Parisi model and the one eigenvalue saddle point that controls large order behavior in the $m=2$ one matrix model (and hence in certain quantities in the Marinari-Parisi model) shows that the same instanton effects control supersymmetry breaking and large orders in perturbation theory in the Marinari-Parisi model.

In the exactly soluble models there are far more useful techniques available to study nonperturbative effects than the asymptotics of perturbation theory. This is not the case in more complicated string theories like the critical strings where complete nonperturbative formulations do not yet exist. In such systems large order behavior can provide the first glimpse into their nonperturbative structure.

The point of this note is to argue that the asymptotic behavior of perturbation theory in all closed string theories should be as in (9) and so leading nonperturbative effects should be of order $\exp(-C/\alpha)$.

The basic reason for the behavior (9) is the large volume of the moduli space of closed Riemann surfaces of genus $g$, $M_g$, that one integrates over to calculate genus $g$ perturbative amplitudes. This volume can be estimated by dividing $M_g$ into cells each with volume depending at most exponentially with the genus. The natural way to do this for moduli spaces of surfaces with at least one puncture is to use the Feynman diagrams of Witten’s open string field theory [17] that produce a triangulation of the moduli space

---

[18]. Diagrams (cells) for moduli space of genus $g$ surfaces with $n$ punctures, $M^g_n$, are made up of $v = 4g + 2n - 4 - 2\chi$ cubic vertices. The number of such diagrams can be counted by large $N$ matrix techniques [18] and is just the coefficient of $N^n$ in the large $N$ matrix gaussian expectation value

\[ \frac{1}{v!} \left< \frac{\text{tr} M^2}{3} \frac{\text{tr} M^3}{3} \cdots \frac{\text{tr} M^n}{3} \right> \]

where there are $v$ vertices, the $v!$ makes the vertices indistinguishable and the $v!$’s account for a cyclic symmetry. Note that this enumeration of surfaces is dual to the one usually considered in matrix models [4]. The total number of diagrams in an open string field theory at order $\alpha^{-\chi}$ is given by (12) with $N = 1$ (to count all numbers of punctures equally). This is just zero dimensional ordinary $\phi^4$ field theory and the result is clearly $\sim C^\chi(-\chi)!$. This leads to nonperturbative effects of order $\exp(-C/\alpha)$. We expect such effects in open string theories because $\alpha$ is the open string coupling constant and the theory can be formulated as a simple string field theory.

For closed string theories we want to enumerate the number of diagrams at genus $g$ with a given number of punctures $n$, i.e., the term of order $N^n$ in (12). The techniques of [19] [18] [20] allow the direct evaluation of this number. For simplicity we specialize to the case $n = 1$. The answer is, asymptotically for large $g$,

\[ \text{Number of Cells}(M^g_1) \sim C^{-2}(2\pi)!. \]

This is not surprising in light of the above results for open strings ($-\chi \sim 2\pi$). The only thing that needs to be checked is that diagrams with just one puncture are not complicated, e.g., each cell is contractible, not

---

\footnote{This has been discussed in the $D = 1$ model by Ginsparg and Zinn-Justin [3].}

\footnote{The first work on the large order behavior of string perturbation theory was done by Gross and Periwal [18]. They showed that the series for the $D = 26$ bosonic string diverges and is not Borel summable by giving a $g!$ lower bound with positive coefficients.

\footnote{This is just the leading term in $a$ at a given order in $1/N$ in a one matrix ($M^d$) model and was evaluated for this reason (for an $M^d$ theory) in [19]. This demonstrates (2\pi)'' behavior in the rescaled one matrix theory.}
should it be unusually small. Therefore we expect the integral over each cell to be \( \sim C^{-2g} \).

Combining with (13) we have the estimate for a genus \( g \) string amplitude \( A_g \) coming from

\[ A_g \sim C^{-2g(2g)!} \] (14)

as claimed above.

We now try to sharpen these arguments in the specific case of the \( D = 26 \) critical bosonic string in flat space. The vacuum amplitude \( V_g \) for the closed string can be written

\[ V_g = \int_{M_g} \mu WP Z(2)Z'(1)^{-13} \] (15)

where \( Z(x) \) is the Selberg zeta function that describes regulated functional determinants

and \( \mu WP \) is the Weil-Peterson volume form. The integrand is manifestly positive everywhere and so \( V_g > 0 \). Penner [22] has shown that the Weil-Peterson volume of each cell in the triangulation of \( M_g \) is bounded below by \( C^{-2g} \). This result combined with (13) gives the rigorous bound

\[ \int_{M_g} \mu WP > C^{-2g(2g)!} \] (16)

It is very plausible, based on degeneration arguments for example, that this bound is true for the no puncture case, \( M_g \), as well, although this has not yet been shown rigorously.

The \( Z \) function part of the integrand has been bounded by Gross and Periwal [18] in their original work on large order behavior. They showed that away from the compactification divisor \( Z(2)Z'(1)^{-13} > C^{-2g} \). Near the divisor the integrand is dominated by the tachyon double pole divergence. We cut off the integral by, say, replacing \( Z(2)Z'(1)^{-13} \) by a genus independent constant whenever it exceeds that constant. Combining this with our information on the Weil-Peterson volume we arrive at the plausible bound

\[ V_g > C^{-2g(2g)!} \] (17)

Since away from the compactification divisor there is no reason for the integrand in (16) to become large we expect (17) to be a reliable estimate rather than a bound. The superstring case is more subtle to analyze because of potential cancellations. Nonetheless the expectation is that for certain quantities the \( (2g)! \) will continue to hold. This is the case in the Minami-Parisi supersymmetric model. We also expect this behavior to hold in non-critical strings, both ordinary and supersymmetric [25]. One effect of the Liouville functional integral will be to provide the power of the cosmological constant that is absorbed in defining the continuum string coupling constant in these theories.

We are proposing that this \( (2g)! \) behavior is a basic signature of closed string theories, much as the \( \delta \) behavior in the loop expansion is a signature of particle theories. Roughly speaking it emerges because two open string vertices are required to make a closed string vertex and open strings are described by simple string field theories that have field theoretic large order behavior. The \( (2g)! \) property is an obvious obstruction to building a simple covariant closed string field theory. It seems that any such theory will necessarily have complicated, nonpolynomial interactions to build up the required large order behavior. Aspects of this problem have already been encountered at low genus by a number of workers [26].

As stressed above, the \( (2g)! \) property indicates that the leading nonperturbative effects in closed string theory are of magnitude \( \exp(-C/\kappa) \). These effects are, for small enough \( \kappa \), much larger than those found in a low energy effective field theory analysis of the closed string. The loop counting parameter in those theories is \( \kappa^2 \) so these effects would have the typical field theoretic magnitude \( \exp(-C/\kappa^2) \). Of course in the critical strings the coupling constant is another field in the theory, the dilaton, whose magnitude is conjecturally set by the dynamics of the theory. The question of which effects are larger is then a dynamical one. The main point we want to make here is that there are new, intrinsically stringy, nonperturbative effects that must be understood in any study of the dynamics of string theory. The residue of these effects in the low energy effective field theory is, of course, of particular interest.

We have explained earlier that the \( \exp(-C/\kappa) \) effects in the matrix models can be understood as the signature of one eigenvalue instantons. It will be important to understand
if a similar phenomenon holds in a more general setting. As a first step in this direction we can examine the \( D = 1 \) model which can be reformulated as a kind of field theory [27]. The field here is just the eigenvalue density as a function of \( \lambda \) and time, \( \rho(\lambda, t) \). The one eigenvalue instanton in terms of \( \rho \) is the tree level eigenvalue distribution plus a delta function of strength \( 1/N \) that splits off from its edge, executes the instanton trajectory and then rejoins the tree level distribution. This is a peculiar field configuration, although it is natural in terms of the original eigenvalues. It seems that the field likes to fall into pieces of size \( 1/N \), allowing the existence of anomalously low action instantons.

We must understand what properties of the action for \( \rho \) allow such a singular stationary point. Clearly \( N \) must be involved to set the scale of the delta function. These issues can also be addressed from the point of view discussed by Tom Banks at this workshop [28].

In the question period after this presentation Spenta Wadia made the interesting remark that known results about Yang-Mills theory show the presence of analogous anomalously large nonperturbative phenomena. In the large \( N \) topological expansion, diagrams of genus \( g \) are weighted by \( (1/N^2)^{g-1} \). Instanton effects are of magnitude \( \exp(-C/\epsilon^2) \) where \( \epsilon^2 \) is the Yang-Mills coupling constant. In the large \( N \) limit \( \epsilon^2 = \epsilon^2 N \) is held fixed so instanton effects are order \( \exp(-CN/\epsilon^2) \) [29]. This is just the phenomenon discussed above. We can estimate the large order behavior in the \( 1/N \) expansion by making the usual assumption that it will be independent of the dimension in which the theory is defined.

Two dimensional Yang-Mills theory (with a lattice cutoff) is just the one unitary matrix model. In the scaled continuum limit [30] this model has \( (2\pi)^2 \) behavior, as we have come to expect. Presumably the unscaled limit appropriate for Yang-Mills theory does as well.

The underlying phenomenon in this example is that the Yang-Mills instanton of lowest action occupies just one \( SU(2) \) subgroup of the \( SU(N) \) gauge group and so its action does not scale with \( N \). Again there is an indication of a string field falling into pieces of order \( 1/N \). We hope that further exploration of this phenomenon will cast light on the nonperturbative dynamics of string theory.

Acknowledgements


This note is an extended version of the second part of my presentation at the Cargése Workshop on Random Surfaces, Quantum Gravity and Strings, May 28-June 1, 1990.

I want to thank O. Alvarez, E. Marinari, and P. Windey for organizing an unusually stimulating conference.

This research was supported in part by grant DE-FG05-90ER40559.
References

[28] T. Banks, presentation at the Cargèse workshop.