Published by Institute of Physics Publishing for SISSA/ISAS

Received: *May 3, 2002*
Revised: *October 27, 2003*
Accepted: *October 27, 2003*

A tentative theory of large distance physics

Daniel Friedan

Department of Physics and Astronomy, Rutgers, The State University of New Jersey
Piscataway, New Jersey 08854-8019 U.S.A., and
Raunvísindastofnum Háskólans Íslands, Reykjavík, Ísland
The Natural Science Institute of the University of Iceland, Reykjavik, Iceland
Email: friedan@physics.rutgers.edu

Abstract: A theoretical mechanism is devised to determine the large distance physics of spacetime. It is a two dimensional nonlinear model, the lambda model, set to govern the string worldsurface in an attempt to remedy the failure of string theory, as it stands. The lambda model is formulated to cancel the infrared divergent effects of handles at short distance on the worldsurface. The target manifold is the manifold of background spacetimes. The coupling strength is the spacetime coupling constant. The lambda model operates at 2d distance Λ^{-1}, very much shorter than the 2d distance μ^{-1} where the worldsurface is seen. A large characteristic spacetime distance L is given by $L^2 = \ln(\Lambda/\mu)$. Spacetime fields of wave number up to $1/L$ are the local coordinates for the manifold of spacetimes. The distribution of fluctuations at 2d distances shorter than Λ^{-1} gives the *a priori* measure on the target manifold, the manifold of spacetimes. If this measure concentrates at a macroscopic spacetime, then, nearby, it is a measure on the spacetime fields. The lambda model thereby constructs a spacetime quantum field theory, cutoff at ultraviolet distance L, describing physics at distances larger than L. The lambda model also constructs an effective string theory with infrared cutoff L, describing physics at distances smaller than L. The lambda model evolves outward from zero 2d distance, $\Lambda^{-1} = 0$, building spacetime physics starting from $L = \infty$ and proceeding downward in L. L can be taken smaller than any distance practical for experiments, so the lambda model, if right, gives all actually observable physics. The harmonic surfaces in the manifold of spacetimes are expected to have novel nonperturbative effects at large distances.

http://jhep.sissa.it/archive/papers/jhep102003063/jhep102003063.pdf

JHEP10(2003)063

Contents

JHEP10(2003)063

1. Introduction

I propose here a systematic, mechanical theory of large distance physics. The mechanism is a two dimensional nonlinear model, the *lambda model*, whose target manifold is a manifold of spacetimes. Each spacetime is characterized by its riemannian metric and certain other spacetime fields. In the lambda model, spacetime as a whole fluctuates locally in two dimensions. The distribution of the fluctuations at short two dimensional distance is a measure on the manifold of spacetimes. If this measure concentrates at a macroscopic spacetime, then, nearby, it is a measure on the spacetime fields in that macroscopic spacetime. Spacetime quantum field theory is thereby constructed as the effective description of large distance physics. But the dynamics that governs the large distance physics is the local dynamics of the two dimensional nonlinear model, the lambda model.

I only formulate the theory here. I describe its structure and speculate about its prospects. I do no calculations in the theory. The arguments are based on abstract general principles. Most of the technical details are left to be filled in. I concentrate on the task of formulating a well-defined theoretical structure that is capable of providing a comprehensive and useful theory of the large distance physics of the real world. The theory that I am proposing does appear capable of selecting a specific discrete set of macroscopic spacetimes, producing a specific spacetime quantum field theory in each. In particular, the theory appears capable of producing specific, calculable, nonperturbatively small mass parameters in the effective spacetime quantum field theories. But calculations are needed to check whether the theory actually does accomplish this. If the theory does work as envisioned, it will be a comprehensive, definitive, predictive theory of large distance physics, whose reliability can be checked by detailed comparison with existing experimental knowledge of the real world.

1.1 Renormalization of the general nonlinear model

This work began with the renormalization of the two dimensional general nonlinear model [1, 2, 3]. The general nonlinear model is a two dimensional quantum field theory. It is defined as a functional integral

$$\int \mathrm{D}x\ \mathrm{e}^{-A(x)} \tag{1.1}$$

over all maps $x(z, \bar{z})$ from the plane to a fixed compact riemannian manifold, called the target manifold. The couplings of the general nonlinear model are comprised in a riemannian metric $h_{\mu\nu}(x)$ on the target manifold, called the target metric or the metric coupling. The classical action is

$$\int \mathrm{d}^2 z\, \frac{1}{2\pi}\, h_{\mu\nu}(x)\, \partial x^\mu\, \bar{\partial} x^\nu\,. \tag{1.2}$$

Each wave mode $\delta h_{\mu\nu}(x)$ of the reimannian metric on the target manifold is a coupling constant λ^i in the two dimensional quantum field theory, parametrizing a perturbation

$$\delta A(x) = \int \mathrm{d}^2 z\, \frac{1}{2\pi}\, \delta h_{\mu\nu}(x)\, \partial x^\mu\, \bar{\partial} x^\nu \tag{1.3}$$

of the action. The general nonlinear model is 'general' in the sense that the riemannian metric on the target manifold is not assumed to have any special symmetries.

The general nonlinear model was shown to be renormalizable [1, 2, 3]. The renormalized couplings of the model were shown to comprise an effective riemannian metric on the target manifold, at every two dimensional distance. The renormalization group was shown to act as a flow on the manifold of riemannian metrics. The infinitesimal renormalization group generator $\beta^i(\lambda)$ became a vector field on the manifold of riemannian metrics. The renormalization group fixed point equation $\beta = 0$, expressing two dimensional scale invariance, became, at large distance in the target manifold, the equation $R_{\mu\nu} = 0$. This was recognized as Einstein's equation for the spacetime metric in general relativity, without matter.

Renormalization is based on an extremely large ratio between two distances. The quantum field theory is constructed at a short distance Λ^{-1}. The theory is used, its properties calculated, at a long distance μ^{-1}. Inverse powers of the extremely large ratio Λ/μ act to suppress the effects of all coupling constants having negative scaling dimensions. In the general nonlinear model, the distances Λ^{-1} and μ^{-1} are two dimensional distances.

The metric coupling of the general nonlinear model is naively dimensionless. The fluctuations in the model give each coupling constant λ^i an anomalous scaling dimension $-\gamma(i)$. It was shown that the anomalous scaling dimensions $-\gamma(i)$ are the eigenvalues of a covariant second order differential operator on the target manifold, acting on the wave modes of the riemannian target metric. Each coupling constant λ^i is an eigenmode with eigenvalue $-\gamma(i)$. The numbers $\gamma(i)$ take the form $\gamma(i) = p(i)^2$ up to corrections for the curvature of the target manifold, where $p(i)$ is the spacetime wave number of the wave mode λ^i.

In a renormalized quantum field theory, the coupling constants λ^i having $\gamma(i) > 0$ are irrelevant. Their effects are suppressed by factors $(\Lambda/\mu)^{-\gamma(i)}$. The quantum field theory depends only on the λ^i having $\gamma(i) = 0$, which are the marginal coupling constants, and the λ^i having $\gamma(i) < 0$, which are the relevant coupling constants. Thus the small distance modes of the target riemannian metric, the wave modes of high wave number, became irrelevant coupling constants λ^i in the renormalized general nonlinear model. The large distance wave modes of the riemannian metric became the marginal and relevant coupling constants in the general nonlinear model.

The target manifold of the general nonlinear model was taken to be compact and riemannian so that the model would be well-defined as a two dimensional quantum field theory. Assuming a riemannian target manifold ensured that the action $A(x)$ would be bounded below. Assuming a compact target manifold ensured a discrete spectrum of anomalous scaling dimensions $-\gamma(i)$. It followed from these assumptions that only a finite number of marginal and relevant coupling constants λ^i could occur in the general nonlinear model. The marginal and relevant coupling constants λ^i are the parameters for variations of the quantum field theory. So the space of general nonlinear models was shown to be a finite dimensional manifold.

When spacetime geometry was translated into the language of the renormalization of the general nonlinear model, it became possible to imagine that the physics of real spacetime might be found encoded within that abstract machinery. It became possible to imagine that real spacetime might in fact be the target manifold of the general nonlinear model, and that Einstein's equation on the physical metric of spacetime might in fact be the fixed point equation $\beta = 0$ of the renormalization group acting on the metric coupling of the general nonlinear model.

Renormalized two dimensional quantum field theory offers a small set of abstract basic principles which are distinguished, definitive and tractable in comparison with the possible principles of spacetime physics that they would replace. The wave modes of the spacetime metric become the coupling constants λ^i which parametrize the two dimensional quantum field theory. The equation of motion $R_{\mu\nu} = 0$ on the spacetime metric becomes the renormalization group fixed point equation $\beta^i(\lambda) = 0$, expressing scale invariance of the two dimensional quantum field theory. Positivity of the spacetime metric becomes unitarity of the two dimensional quantum field theory. The compactness of spacetime becomes the discreteness of the spectrum of two dimensional scaling fields. Small distance in spacetime becomes irrelevance in the two dimensional quantum field theory. Geometric conditions on spacetime became natural regularity conditions on two dimensional quantum field theories. The renormalization of the coupling constants λ^i of the general nonlinear model is a systematic and reliable calculus. A construction can be formulated in the language of the renormalized general nonlinear model with confidence in its coherence, although explicit calculations might remain technically difficult.

When the general nonlinear model was shown to be renormalizable, it was pointed out [1, 2, 3] that a manifold of nontrivial compact riemannian solutions to the one loop fixed point equation $R_{\mu\nu} = 0$ were already known to exist, namely the Calabi-Yau spaces [4, 5]. But the two dimensional scale invariance of a general nonlinear model with a Calabi-Yau target manifold was violated by the two loop contribution to the beta function. Nontrivial two dimensional scale invariance was discovered in the supersymmetric version of the general nonlinear model with Calabi-Yau target manifold, when the remarkable cancellation among the two loop contributions to the beta function was discovered [6].

1.2 Application in string theory

The nontrivial scale invariant general nonlinear models found a role in perturbative string theory [7]. The general nonlinear model constructs the string worldsurface in a curved background spacetime. The target manifold of the nonlinear model is the background spacetime in which strings scatter. The two dimensional plane gives the local two dimensional patches out of which the string worldsurface is made. Consistency of the string theory requires the string worldsurface to be scale invariant, so the coupling constants λ^i in the general nonlinear model of the worldsurface must satisfy the fixed point equation $\beta^i(\lambda) = 0$. The manifold of scale invariant general nonlinear models forms the manifold of possible background spacetimes.

1.3 The failure of string theory, as it stands[1]

String theory, as it stands, has failed as a theory of physics because of the existence of a manifold of possible background spacetimes. All potentially observable properties of string theory depend on the geometry and topology of the background spacetime in which the strings scatter. In string theory, a specific background spacetime has to be selected by hand, or by "initial conditions", from among the manifold of possibilities. Many continuously adjustable parameters must be dialed arbitrarily to specify the background spacetime. The existence of a manifold of possible background spacetimes renders string theory, as it stands, powerless to say anything definite that can be checked.

1.4 Physics is reliable knowledge

Physics is reliable knowledge of the real world, based on experiment. A new theory of physics must establish its reliability first by explaining existing knowledge of the real world. New theories of physics build on existing reliable theories. A candidate theory of physics obtains credibility first by giving definite explanations of established theories. A new theory inherits the reliability of the theories it explains. For example, special relativity explained newtonian mechanics. General relativity explained special relativity and newtonian gravity. Quantum mechanics explained classical mechanics. Bohr's correspondence principle, which guided the formation of quantum mechanics, was an explicit statement that a candidate theory of physics must explain the existing reliable theory.

Present knowledge of the laws of physics is summarized in the combination of classical general relativity and the standard model of elementary particles, to the extent that the standard model has been confirmed by experiment. A candidate theory of physics must establish its reliability by explaining this currently successful theory. It must explain

[1]**Note added for publication.** The introduction and abstract have been revised to address a concern of the referee. Phrases such as 'the failure of string theory' and 'string theory cannot . . . ' were interpreted by the referee as absolute assertions of impossibility. These phrases have now been revised by systematically replacing 'string theory' with 'string theory, as it stands'. The revision is meant to make more explicit the diagnostic nature of my discussion. My purpose in noting 'the failure of string theory, as it stands' is to diagnose the failure in order to attempt a remedy. I see string theory, as it stands, to be an S-matrix theory, unable to give a mechanical description of large distance physics. I see the need to choose from among a multiplicity of possible background spacetimes as a failure, suggesting to me that there is something wrong with the specification of what is a background spacetime, in string theory as it stands. I propose to provide string theory with a mechanism that actually produces the background spacetime. As it happens, this new machinery, if it should succeed, would end up supplanting string theory. It would generate a mechanical description of all of large distance physics, relegating the string S-matrix to a formal role as a formally consistent description of presently unobservable small distance physics. Nevertheless, the design is to fix string theory, as it stands. I take a very conservative view of 'string theory, as it stands'. The part of string theory that stands solidly, that does not rest on conjectures and assumptions, is the perturbative string S-matrix. I do not exclude the possibility that some other more or less conjectural versions of string theory might some day succeed. But I take the position that conjecture, even when strongly supported by internal evidence, remains conjecture. I worry that some of the underlying assumptions might turn out to be unreliable, spacetime supersymmetry in particular. Perhaps some version of string theory will eventually succeed as physics, perhaps it will be a version that realizes current conjectures. Meanwhile, for diagnostic purposes, what string theory can do is what it actually manages to do, as it stands, not what it is conjectured to be able to do, or what it might some day do.

JHEP10(2003)063

classical general relativity and the standard model in detail. At the very least, a candidate theory of physics must contain fewer adjustable parameters than does the standard model, and must give reliable methods to calculate precise numerical values for the masses and coupling constants in the standard model. A candidate theory of physics that is not capable of explaining the standard model and classical general relativity cannot obtain reliability, because it cannot be checked against the existing knowledge of the real world.

The standard model of elementary particles is a quantum field theory. General relativity is a classical field theory, but can be regarded equally well as a quantum field theory which is accurately approximated by its classical field theoretic limit at the spacetime distances where gravity is observed. A theory of physics should explain quantum field theory. It should explain why quantum field theory in spacetime has been so successful at the spacetime distances accessible to observation.

A candidate theory of physics must be capable of *producing* spacetime quantum field theory. More, it must be capable of producing one specific quantum field theory, containing specific, nontrivial, calculable mass parameters and coupling constants. One specific quantum field theory, the standard model, has been successful in physics, not quantum field theory in general. Quantum field theory in general has too many free parameters to be a useful search space in which to find a definite explanation of the standard model. A mechanism is needed that is capable of producing a specific spacetime quantum field theory, the one that is actually seen in the real world. The unnaturally small value of the cosmological constant suggests that, if such a mechanism is at work in the real world, it does not work generically, but rather in a very specific fashion, to produce a very specific spacetime quantum field theory.

I stress *capability.* The first step in forming a theory of physics is to find a well-defined theoretical structure *capable* of producing a specific spacetime quantum field theory. Only then is there a chance of explaining the standard model and general relativity, and of making definite predictions. When such a theoretical structure is found, it becomes worthwhile to perform calculations to determine whether the capabilities are realized. Of course, success in physics requires actually giving definite explanations of existing knowledge and actually making definite predictions that are verified. But a first prerequisite in a candidate theory of physics is a structure *capable* of providing definite, unequivocal explanations and predictions. A theory of physics must be capable of making definite statements that can be checked. It must be capable of doing useful work in the real physical world.

1.5 Only large distance physics is observable

In units of the Planck length, $l_{\mathrm{P}} = (1 \times 10^{19}\,\mathrm{Ge}V)^{-1}$, the smallest distance probed by feasible experiments is a very large dimensionless number, on the order of $1 \times 10^{16} = (1 \times 10^{3}\,\mathrm{Ge}V\, l_{\mathrm{P}})^{-1}$, or perhaps $1 \times 10^{14} = (1 \times 10^{5}\,\mathrm{Ge}V\, l_{\mathrm{P}})^{-1}$. In any theory of physics in which spacetime distances are dimensionless numbers and in which the unit of distance lies within a few orders of magnitude of the Planck length, the only theoretical explanations and predictions that can be checked against experiment are those made in the large distance limit of the theory.

1.6 The long-standing crisis of string theory

The long-standing crisis of string theory is its complete failure to explain or predict any large distance physics. String theory, as it stands, cannot say anything definite about large distance physics. String theory, as it stands, is incapable of determining the dimension, geometry, particle spectrum and coupling constants of macroscopic spacetime. String theory, as it stands, cannot give any definite explanations of existing knowledge of the real world and cannot make any definite predictions. The reliability of string theory cannot be evaluated, much less established. String theory, as it stands, has no credibility as a candidate theory of physics.

Recognizing failure is a useful part of the scientific strategy. Only when failure is recognized can dead ends be abandoned and useable pieces of failed programs be recycled. Aside from possible utility, there is a responsibility to recognize failure. Recognizing failure is an essential part of the scientific ethos. Complete scientific failure must be recognized eventually.

String theory, as it stands, fails to explain even the existence of a macroscopic spacetime, much less its dimension, geometry and particle physics. The size of the generic possible background spacetime is of order 1 in dimensionless units. Large distances occur only in macroscopic spacetimes, which are found near the boundary of the manifold of background spacetimes. String theory, as it stands, cannot explain the existence of a macroscopic spacetime, being incapable of selecting from among the manifold of possible background spacetimes.

Even if some particular macroscopic background spacetime is chosen arbitrarily, by hand or by "initial conditions", string theory, as it stands, still fails to be realistic at large distance. The large distance limit of string theory, as it stands, consists of the perturbative scattering amplitudes of the low energy string modes, which are particle-like. But the particle masses are exactly zero, and the low energy scattering amplitudes are exactly supersymmetric. String theory, as it stands, fails to provide any mechanism to generate the very small nonzero masses that are observed in nature, or to remove the exact spacetime supersymmetry, which is not observed in nature. More broadly, string theory, as it stands, is incapable of generating the variety of large characteristic spacetime distances seen in the real world. At best, for each macroscopic background spacetime in the manifold of possibilities, string theory, as it stands, gives large distance scattering amplitudes that form a caricature of the scattering amplitudes of the standard model of particle physics.

The massless string modes are the manifestations, locally in the macroscopic spacetime, of the continuous degeneracy of the manifold of background spacetimes. The failure of string theory, as it stands, to generate nonzero small particle masses is a consequence of its failure to resolve the continuous degeneracy of the manifold of spacetimes. The continuous degeneracy of the manifold of background spacetimes makes string theory, as it stands, unacceptable as a candidate theory of physics. If the continuous degeneracy were accepted, then, by assumption, it would be impossible to determine the dimension and geometry of macroscopic spacetime or the masses and coupling constants of the elementary particles.

JHEP10(2003)063

String theory, as it stands, fails to produce spacetime quantum field theory at large distance. String theory, as it stands, gives only scattering amplitudes. String theory, as it stands, cannot explain the standard model, or general relativity, because it cannot *produce* a spacetime quantum field theory as an effective description of large distance physics. The practice in string theory, as it stands, is to *assume* that spacetime quantum field theory describes the large distance physics. First, a macroscopic background spacetime is chosen by hand, arbitrarily, from among the manifold of possibilities. Then string theory scattering amplitudes are calculated perturbatively in the chosen background spacetime. The perturbative string theory is invariant under some spacetime supersymmetries. The massless particle-like states and their perturbative large distance scattering amplitudes are identical to the perturbative large distance scattering amplitudes derived from a supersymmetric field theory lagrangian in the arbitrarily chosen macroscopic spacetime. It is then *assumed* that the large distance physics in the chosen macroscopic spacetime is given by some quantized version of the supersymmetric spacetime field theory. The continuous degeneracy of the manifold of background spacetimes appears as a continuous degeneracy of the manifold of ground states of the spacetime field theory, as a continuous degeneracy of the manifold of possible vacuum expectation values of the spacetime fields. The supersymmetric spacetime field theory is then examined for possible nonperturbative effects that might break the degeneracy of the manifold of ground states.

The assumption that spacetime quantum field theory governs the large distance physics is not justified. There is no derivation of spacetime quantum field theory from string theory, as it stands. There is no construction from string theory, as it stands, of any effective spacetime quantum field theory governing the large distance physics, even given an arbitrary choice of background spacetime. String theory, as it stands, is incapable of explaining any spacetime field theory, classical or quantum mechanical. String theory, as it stands, provides nothing at large distance but perturbative scattering amplitudes for gravitons and other massless particles. It is true that the same perturbative scattering amplitudes for massless particles can be derived from massless supersymmetric quantum field theories, but this formal coincidence does not justify the claim that string theory, as it stands, explains quantum field theory, or the claim that string theory, as it stands, implies quantum field theory at large distances.

In particular, there is no justification for the claim that string theory, as it stands, explains or predicts gravity. String theory, as it stands, gives perturbative scattering amplitudes of gravitons. Gravitons have never been observed. Gravity in the real world is accurately described by general relativity, which is a classical field theory. There is no derivation of general relativity from string theory, as it stands. General relativity can be regarded as the large distance classical limit of quantum general relativity, if an ultraviolet cutoff is imposed to make sense of quantum general relativity. A cutoff quantum general relativity would give the same formal perturbative low energy scattering amplitudes for massless gravitons as does string theory. But it is illogical to claim, from this formal coincidence between two technical methods of calculating unobserved graviton scattering amplitudes, that string theory, as it stands, explains classical general relativity, or that string theory, as it stands, explains gravity, or that string theory, as it stands, is a quantum

JHEP10(2003)063

theory of gravity. String theory, as it stands, does not produce any mechanical theory of gravity, much less a quantum mechanical theory.

In any case, a quantum theory of gravity is unnecessary at present. No physical effects of quantum gravity have been observed, and there is no credible possibility of observing any. What *is* needed is a theory which produces general relativity as an effective classical field theory at large distance. It might produce classical general relativity by producing at large distance an effective quantized general relativity that is deep in its classical regime. But what is essential to produce is the classical, mechanical spacetime field theory of gravity.

String theory, as it stands, is only a perturbative theory. The widespread practice is to assume that there exists a nonperturbative formulation of string theory, and that this hypothetical nonperturbative formulation would be a quantum mechanical theory, microscopic in spacetime, invariant under some exact, fundamental spacetime supersymmetries. If such a nonperturbative formulation of string theory did exist, then it might well follow that the large distance physics in that hypothetical theory would be governed by supersymmetric spacetime quantum field theory, and that the fate of the degeneracy of the manifold of background spacetimes would be determined by nonperturbative field theoretic effects at large distance in spacetime in that supersymmetric quantum field theory. But it is only an assumption that there exists such a nonperturbative, microscopic, quantum mechanical formulation of string theory. Any reasoning about a hypothetical nonperturbative version of string theory is unreliable if it rests on the assumption of spacetime quantum field theory at large distance, without any way to derive spacetime quantum field theory from string theory.

The assumption of fundamental, exact, quantum mechanical spacetime supersymmetry is a very strong extrapolation from perturbative string theory, where spacetime supersymmetry is only a perturbative symmetry of the scattering amplitudes in individual background spacetimes. Adopting this assumption requires accepting as inevitable the continuous degeneracy of the manifold of background spacetimes. An assumption as strong as fundamental spacetime supersymmetry loses credibility as a guide in searching for a theory of physics if it cannot lead to definite explanations of existing knowledge and definite predictions. There is certainly no physical evidence to support the assumption that spacetime supersymmetry is a fundamental property of nature. At most, it is possible that indications of approximate spacetime supersymmetry might be found experimentally in the not so distant future. Contrast the radical assumptions of the old quantum theory, which obtained credibility by giving definite explanations of the black body spectrum, the photoelectric effect, the Balmer series, the Rydberg constant, and much more of atomic physics, before eventually leading to quantum mechanics.

Any reasoning about a hypothetical nonperturbative version of string theory is unreliable if it assumes fundamental spacetime supersymmetry but is unable to make definite, unequivocal explanations or predictions that could be used to check that assumption. The search for a theory of physics should not be based on dogma. Certain symmetries are observed in the real world, to a certain accuracy, in a certain range of spacetime distances. This does not justify a dogma of fundamental symmetry in theoretical physics, much less a dogma of fundamental spacetime supersymmetry. It has rarely proved fruitful in physics

JHEP10(2003)063

to cling indefinitely to assumptions that are incapable of producing definite explanations of existing knowledge or definite predictions.

Spacetime supersymmetry does give beneficial formal effects in spacetime quantum field theories of particle physics, but these benefits could as well be provided by accidental, approximate spacetime supersymmetry. In a weakly coupled theory, perturbative spacetime supersymmetry would be enough to protect mass parameters that are perturbatively zero, so that very small masses could be produced by nonperturbative, supersymmetry violating effects. Spacetime supersymmetry provides benefits for formal calculation, giving powerful analytic control over quantum mechanical theories and especially over spacetime quantum field theories. But the price of control is the supersymmetry itself. Supersymmetry is not observed in nature, and the theoretical control is lost with the loss of supersymmetry. Useful theoretical control must come from some other source.

The assumption that physics has a microscopic quantum mechanical formulation is of course supported by an enormous body of physical evidence. Microscopic quantum mechanics has had triumphant success, culminating in the local quantum field theory that is the standard model of elementary particles. But the evidence for microscopic quantum mechanics is entirely at very large distance in spacetime. However strong is the evidence for quantum mechanics at large distance, that evidence does not require that microscopic quantum mechanics in spacetime must be the fundamental language of physics. The evidence only requires that microscopic quantum mechanics be produced at large distance in any theory of physics. Quantum mechanics in spacetime is the language in which large distance physics is to be read out, but it is not necessarily the language in which physics is to be written.

Likewise, the fact that certain beautiful mathematical forms were used in the period 1905-1974 to make the presently successful theory of physics does not imply that any particular standard of mathematical beauty is fundamental to nature. The evidence is for certain specific mathematical forms, of group theory, differential geometry and operator theory. The evidence comes from a limited range of spacetime distances. That range of distances grew so large by historical standards, and the successes of certain specific mathematical forms were so impressive, that there has been an understandable psychological impulse in physicists responsible for the triumph, and in their successors, to believe in a certain standard of mathematical beauty. But history suggests that it is unwise to extrapolate to fundamental principles of nature from the mathematical forms used by theoretical physics in any particular epoch of its history, no matter how impressive their success. Mathematical beauty in physics cannot be separated from usefulness in the real world. The historical exemplars of mathematical beauty in physics, the theory of general relativity and the Dirac equation, obtained their credibility first by explaining prior knowledge. General relativity explained newtonian gravity and special relativity. The Dirac equation explained the non-relativistic, quantum mechanical spinning electron. Both theories then made definite predictions that could be checked. Mathematical beauty in physics cannot be appreciated until after it has proved useful. Past programs in theoretical physics that have attempted to follow a particular standard of mathematical beauty, detached from the requirement of correspondence with existing knowledge, have failed. The evidence for

JHEP10(2003)063

beautiful mathematical forms in nature requires only that a candidate theory of physics explain those specific mathematical forms that have actually been found, within the range of distances where they have been seen, to an approximation consistent with the accuracy of their observation.

1.7 Formal clues

The search for an explanation of the standard model and of general relativity has become a speculative enterprise, because the guiding theoretical principles of local spacetime field theory, quantum mechanics and symmetry have proved inadequate. A strategy must be chosen. It is necessary to decide what formal clues might be useful. No particular choice of strategy is inherently valid. Only the outcome of the search can give validity.

My primary formal clue to a possible theory of large distance physics has been the expression of spacetime geometry in the renormalization of the general nonlinear model. The appearance of the field equation $R_{\mu\nu} = 0$ of general relativity as the fixed point equation $\beta = 0$ of the general nonlinear model suggested that spacetime field theory might somehow be derived from the general nonlinear model. The renormalization of the general nonlinear model isolates the large distance wave modes of its target manifold, decoupling the irrelevant small distance wave modes. The renormalization uses the extreme shortness of the two dimensional distance Λ^{-1} where the model is constructed, compared to the two dimensional distance μ^{-1} where the properties of the model are seen. The large distance physics of spacetime is to be found in the short distance structure of the renormalized general nonlinear model.

String theory was the second clue. String theory gave a specific technical context, the string worldsurface, in which to place the general nonlinear model. The fixed point equation $\beta = 0$ for the general nonlinear model of the string worldsurface is the condition of two dimensional scale invariance, which is needed for string theory to be consistent. The possible background spacetimes for string theory are determined by imposing the equation $\beta = 0$ on the general nonlinear model of the worldsurface. Thus the Einstein equation $R_{\mu\nu} = 0$ arises as a consistency condition in string theory.

But a consistency condition is not an equation of motion. A consistency condition is not the mechanical dynamics of a field theory in spacetime. String theory, as it stands, does not have a dynamical mechanism that constructs an effective quantum field theory at large distance in spacetime, whose equation of motion is $\beta = 0$.

Nor does string theory, as it stands, have a dynamical mechanism that selects the background spacetimes to be those in which the worldsurface is scale invariant. Without such a mechanism, there cannot be a reliable characterization of the possible background spacetimes for string theory.

The third clue was the failure of string theory, as it stands, at large distance. This failure provides a formal task for a theory of large distance physics to accomplish, the task of determining dynamically the possible background spacetimes for string theory. String theory, as it stands, although useless at large distance, is formally successful in the ultraviolet as a technical perturbative algorithm for calculating ultraviolet scattering amplitudes in a given background spacetime. But scattering amplitudes are not sufficient for physics.

A theory of physics must produce an effective mechanical theory at large distance, if it is to explain existing knowledge. The evidence for the reliability of theoretical physics includes all the evidence for newtonian mechanics, newtonian gravity, classical electromagnetism, special relativity, general relativity, non-relativistic quantum mechanics and the standard model. Any candidate theory of physics must be capable of producing each of those mechanical theories as an approximation in the appropriate regime. A theory that gives only scattering amplitudes is not capable of this. Scattering amplitudes can be derived from a mechanical theory, but mechanics cannot be derived from a theory of scattering amplitudes. Scattering amplitudes intrinsically represent small distance physics as observed by a relatively large experimentalist. String theory, as it stands, might well serve adequately as a technical perturbative algorithm for calculating ultraviolet scattering amplitudes. It might serve as a formal representation of unobservable small distance physics. But a reliable and effective mechanism outside string theory, as it stands, is needed to determine the large distance physics of spacetime.

1.8 The lambda model

My strategy has been to analyze the failure of string theory at large distance in an arbitrarily fixed background spacetime. The technical symptom of failure is a short distance pathology in the string worldsurface, a logarithmic divergence at short two dimensional distance, bi-local in form, produced by degenerating handles attached locally to the worldsurface. The divergence is due to the existence of marginal coupling constants in the general nonlinear model of the worldsurface. Marginal coupling constants express the continuous degeneracy of the manifold of possible spacetimes. The divergence is an infrared problem in spacetime, because the marginal coupling constants are the large distance wave modes of spacetime.

A theoretical mechanism is then devised to cancel the bi-local divergence. The mechanism is a two dimensional nonlinear model, the *lambda model*. The target space of the lambda model is the manifold of spacetimes, which is the manifold of renormalized general nonlinear models of the string worldsurface. In the lambda model, spacetime as a whole fluctuates in two dimensions.

The fields $\lambda^i(z,\bar{z})$ of the lambda model are local sources in the general nonlinear model. They are coupled to the marginal and slightly irrelevant two dimensional quantum fields $\phi_i(z,\bar{z})$ of the general nonlinear model. The lambda fields $\lambda^i(z,\bar{z})$ fluctuate with a propagator designed so that, acting as a bi-local source, it cancels the bi-local effects of a handle attached locally on the worldsurface. The couplings of the lambda model are completely determined by the cancellation requirement. In particular, the coupling strength of the lambda model is equal to the spacetime coupling constant g_s of the perturbative string theory.

There are two widely separated two dimensional distances. The coupling constants λ^i of the renormalized general nonlinear model are normalized at μ^{-1}, the long two dimensional distance. The lambda fields $\lambda^i(z,\bar{z})$ fluctuate at short two dimensional distances, up to a sliding characteristic two dimensional distance Λ^{-1} which stays much shorter than μ^{-1}. The renormalization of the general nonlinear model suppresses the effects of the short

JHEP10(2003)063

distance fluctuations of the coupling constant λ^i by a factor $(\Lambda/\mu)^{-\gamma(i)}$, where $-\gamma(i)$ is the anomalous dimension of λ^i. The coupling constants which have $\gamma(i)\ln(\Lambda/\mu) \gg 1$ are irrelevant at the short two dimensional distance Λ^{-1}. There are only a finite number of non-irrelevant coupling constants λ^i in the general nonlinear model, so the target manifold of the lambda model is finite dimensional.

The number L defined by $L^2 = \ln(\Lambda/\mu)$ is a spacetime distance, because the anomalous dimensions $-\gamma(i)$ are the eigenvalues of differential operators in spacetime that are quadratic in spacetime derivatives. The effects of the spacetime wave mode λ^i are suppressed by factors $\mathrm{e}^{-L^2\gamma(i)}$. The irrelevant coupling constants in the general nonlinear model are the λ^i with $\gamma(i)L^2 \gg 1$. These are the spacetime wave modes at spacetime distances $1/p(i)$ smaller than L.

The small distance wave modes, being irrelevant coupling constants, are decoupled in the renormalization of the general nonlinear model. Their fluctuations can be omitted from the lambda model at two dimensional distance Λ^{-1}. Only the non-irrelevant coupling constants fluctuate, the coupling constants λ^i having $\gamma(i)L^2 < 1$. These are the large distance spacetime wave modes, the wave modes at spacetime distances larger than L. Thus, from the renormalization of the general nonlinear model, the lambda model inherits a natural, built-in, sliding ultraviolet spacetime cutoff distance L. The spacetime wave modes at distances smaller than L are decoupled from the large distance wave modes, so this is an ultraviolet cutoff in the strongest sense.

The lambda model is a two dimensional quantum field theory. As such, its construction starts from the short distance limit at $\Lambda^{-1} = 0$, building outward to nonzero values of the sliding characteristic two dimensional distance Λ^{-1}. So the lambda model builds spacetime physics from the limit at $L = \infty$ downward to finite values of the sliding characteristic large spacetime distance L.

The fluctuations in a nonlinear model at distances shorter than the characteristic two dimensional distance Λ^{-1} distribute themselves to form a measure on the target manifold of the model, called the *a priori* measure of the nonlinear model, following the terminology of lattice statistical mechanics, using '*a priori*' with its literal meaning 'from what is before' or 'from the earlier part' [1, 2, 3]. The *a priori* measure summarizes the short distance fluctuations in the nonlinear model. As the characteristic two dimensional distance Λ^{-1} increases from zero, the fluctuations in the nonlinear model generate the *a priori* measure by a diffusion process on the target manifold.

The target manifold of the lambda model is the manifold of spacetimes, so the *a priori* measure of the lambda model is a measure on the manifold of spacetimes. The manifold of spacetimes is the manifold of general nonlinear models. As Λ^{-1} increases, the fluctuations cause the *a priori* measure of the lambda model to diffuse in the manifold of general nonlinear models, while simultaneously the general nonlinear model is flowing under the renormalization group. The *a priori* measure of the lambda model undergoes a driven diffusion process. The generator of the driving flow is the vector field $-\beta^i(\lambda)$ on the manifold of general nonlinear models. The renormalization group flow pushes the *a priori* measure toward the fixed point submanifold where $\beta(\lambda) = 0$. The lambda model dynamically imposes the two dimensional scale invariance condition $\beta(\lambda) = 0$ on the general nonlinear model.

JHEP10(2003)063

The lambda model is background independent, because, even if a particular background spacetime is initially selected by hand, the fluctuations in the lambda model diffuse the distribution of spacetimes away from the arbitrary initial spacetime. Whatever the arbitrary initial choice of background spacetime, the *a priori* measure diffuses to the unique stable measure of the driven diffusion process. The lambda model eliminates the need in string theory to choose a background spacetime by hand.

The lambda model is a nonperturbative two dimensional quantum field theory. Nonperturbative two dimensional effects in the lambda model due to harmonic surfaces in the manifold of spacetimes appear capable of lifting the continuous degeneracy of the manifold of spacetimes, possibly concentrating the *a priori* measure at some macroscopic spacetimes. If so, then the *a priori* measure of the lambda model, near such a macroscopic spacetime, is a measure on the spacetime wave modes λ^i at spacetime distances larger than L. It is a spacetime quantum field theory with ultraviolet cutoff distance L. The lambda model produces a specific quantum field theory, with equation of motion $\beta = 0$, describing the spacetime physics at large spacetime distances L in each possible macroscopic spacetime.

Although the lambda model works from $L = \infty$ downwards in L, it constructs spacetime quantum field theory so that it is local in spacetime, in the sense that the *a priori* measure at a spacetime distance L_1 can be obtained from the *a priori* measure at a smaller distance $L_2 < L_1$ by integrating out the spacetime wave modes at the distances between L_2 and L_1. The lambda model accomplishes this in reverse fashion. The lambda model constructs the *a priori* measure with *increasing* two dimensional distance Λ^{-1}, so with *decreasing* spacetime distance L. The lambda model makes the *a priori* measure at the smaller distance L_2 from the *a priori* measure at the larger distance L_1 by diffusion of the wave modes at the intermediate spacetime distances, between L_2 and L_1. Locality is ensured because integrating out the intermediate wave modes merely undoes the diffusion. The spacetime quantum field theory is constructed so as to be local as a measure on the spacetime wave modes, but there is no guarantee that the effective lagrangian of the spacetime field theory at the smaller distance can be used to determine the effective spacetime quantum field theory at larger distances, except perturbatively. Nonperturbative two dimensional effects in the lambda model might intervene in the construction of the effective spacetime action.

The proposed theory, if successful, will call into question the atomistic assumption that the effective laws of physics at large distance can be deduced from the laws of physics at small distance. It will call into question the atomistic assumption there is a fundamental microscopic formulation of physics. If the proposed theory works, the observed quantum mechanical hamiltonian will be explained, but there will not be a fundamental quantum mechanical hamiltonian.

The lambda model also produces an effective worldsurface at two dimensional distances longer than Λ^{-1}. The effective worldsurface can be used to calculate effective string scattering amplitudes, cut off in the infrared at spacetime distance L. For each large spacetime distance L, the lambda model gives two complementary descriptions of spacetime physics. The spacetime physics at distances larger than L is described by an effective spacetime quantum field theory. The spacetime physics at distances smaller than L is described by effective string scattering amplitudes. The two descriptions of spacetime physics are consistent.

The *a priori* measure on the manifold of spacetimes is the effective spacetime background in which the effective string scattering takes place, at every large spacetime distance L. The relation between the effective string scattering amplitudes and the effective spacetime quantum field theory is not the relation commonly assumed in string theory. The large distance spacetime physics is not derived from any microscopic, small distance physics. In particular, it is not derived from string theory. There is no microscopic quantum mechanical system underlying string theory, that has spacetime quantum field theory as its effective description at large distance. The lambda model provides the spacetime background for the effective string theory, constructing it starting from the limit at spacetime distance $L = \infty$. The fluctuations of the lambda model make quantum corrections to the effective worldsurface in tandem with corrections to the effective metric coupling and *a priori* measure of the lambda model itself, thus ensuring that the effective string scattering amplitudes match, at every large spacetime distance L, the particle scattering amplitudes calculated from the effective quantum field theory.

The lambda model, if right, determines all actually observable physics. The sliding characteristic spacetime distance L can be taken smaller than any distance actually accessible to experiment, while still remaining a large number. As a practical matter, only the large distance physics given by the *a priori* measure can be checked. All calculations of large distance physics can be done in the lambda model. No string calculations are necessary. The fluctuations of the lambda model completely replace the effects of handles at short distance in the worldsurface, so string calculations, and especially string loop calculations, are entirely unnecessary, as far as large distance physics is concerned.

String loop calculations are needed only for perturbative calculations of the effective string scattering amplitudes at unobservably small spacetime distances. There is no practical use for these effective string scattering amplitudes. The only information that the effective string scattering amplitudes give, beyond what is given by the *a priori* measure of the lambda model, is information about physics at unobservably small distances. There is no practical way to make any independent test of the small distance effective string scattering amplitudes.

The lambda model is a nonperturbative theory, while the small distance effective string theory is only perturbative. String theory calculations of scattering amplitudes at spacetime distances smaller than L will not be reliably accurate, because nonperturbative effects in the lambda model at spacetime distances smaller than L will not yet have been taken into account. The only reliable calculations will be the nonperturbative calculations of large distance physics that are made in the lambda model.

String theory is used in three ways in the lambda model. First, the string worldsurface gives a specific technical context in which to place the general nonlinear model. The manifold of spacetimes is the manifold of general nonlinear models of the worldsurface. The detailed specification of the manifold of spacetimes depends on the detailed technical form of the worldsurface in which the general nonlinear model is placed.

Second, string theory gives an algorithm for calculating perturbative corrections to scattering amplitudes in terms of handles in the worldsurface, an algorithm that is formally consistent at small spacetime distance. The lambda model is constructed to cancel the perturbative corrections due to handles attached locally on the worldsurface, so the details of the technical form of the worldsurface determine the detailed definition of the target manifold and the metric coupling of the lambda model.

Once the lambda model is defined, string theory becomes an auxiliary technical apparatus. String theory is used only as a formal representation of unobservable small distance physics in the spacetime constructed by the lambda model. The correspondence between the effective general nonlinear model of the worldsurface and the effective lambda model is used for technical purposes. The correspondence constrains the renormalization of the lambda model. The two dimensional scaling properties of the effective general nonlinear model implies the two dimensional scale invariance of the lambda model.

It appears that, for entirely technical reasons, the heterotic string worldsurface [8] is the only form of the worldsurface that is suitable for the lambda model. The manifold of spacetimes is a graded manifold. Its bosonic and fermionic coordinates are the bosonic and fermionic coupling constants λ^i, which are the wave modes of the bosonic and fermionic spacetime fields. A two dimensional nonlinear model such as the lambda model is well-defined only if its metric coupling is positive definite on the bosonic part of its target manifold. Only for the heterotic string worldsurface is the metric on the manifold of spacetimes positive definite in the bosonic directions. For this technical reason, it seems that only the heterotic string worldsurface is suitable for the lambda model.

Fortunately, the heterotic worldsurface is the only one that is suitable for a formal representation of weakly coupled small distance spacetime physics. The heterotic string theory gives small distance scattering amplitudes of massless chiral fermions, vector bosons, scalar bosons, and gravitons. The perturbative spacetime superymmetry of the heterotic string theory ensures that the general nonlinear model of the heterotic worldsurface contains only marginal and irrelevant coupling constants. There are no relevant coupling constants, which would be the wave modes of spacetime tachyon fields. The perturbative spacetime supersymmetry of the heterotic theory ensures that the perturbative string theory is consistent at small distance. The perturbative spacetime supersymmetry is inherited by the lambda model, where it protects the zero mass spacetime fields against perturbative mass corrections, allowing the possibility of small spacetime masses generated by nonperturbative effects in the lambda model violating the perturbative spacetime supersymmetry.

My most optimistic hope, amounting only to wishful thinking at present, is that nonperturbative weak coupling effects in the lambda model will produce a calculable spectrum of large distances in the spacetime physics generated by the lambda model. Each harmonic surface $\lambda_H(z, \bar{z})$ in the manifold of spacetimes is an instanton in the lambda model, a *lambda instanton*. It is easy to point to the existence of harmonic surfaces in the manifold of spacetimes, but their effects remain speculative until calculated.

Harmonic surfaces in the manifold of spacetimes appear capable of making small contributions to the effective action of the spacetime quantum field theory, giving small masses to the elementary particles. More broadly, they appear capable of eliminating the continuous degeneracies, including spacetime supersymmetries and ordinary gauge symmetries. Spacetime symmetries would then be seen as only accidental and approximate attributes of individual spacetimes.

JHEP10(2003)063

A lambda instanton $\lambda_H(z,\bar{z})$ that is localized in a macroscopic spacetime can be expected to generate spacetime masses of the form $m^2 = \mathrm{e}^{-S(\lambda_H)}$, where $S(\lambda_H)$ is the classical action of the harmonic surface λ_H. The coupling strength of the lambda model is the spacetime coupling constant g_s so $S(\lambda_H)$ is proportional to g_s^{-2}. The actual numerical values of the elementary particle masses might be produced in this way, if g_s^2 is on the order of $1/100$ and if the unit of length is logarithmically close to the Planck length. For example, the mass-squared of the W vector boson is approximately $m_W^2 = \mathrm{e}^{-78}$ in Planck units, and that of the electron is $m_e^2 = \mathrm{e}^{-101}$. If the lambda model does in fact produce such calculable small particle masses, it will become of interest to check whether mass generation by the lambda model can be distinguished experimentally from the quantum field theoretic Higgs mechanism.

Even more fanciful hopes are evoked by writing the inverse square of the Hubble length in Planck units, approximately e^{-281}. It is hard to imagine where such a number might come from, if not a semiclassical, nonperturbative, weak coupling effect. It would be wonderful, though rather much to expect, if semiclassical nonperturbative effects in the lambda model could explain systematically the essential features of the rich spectrum of large characteristic spacetime distances observed in the real world.

If the lambda model does succeed in reducing the continuous degeneracy of the manifold of spacetimes at least to a discrete degeneracy, then the remaining uncertainty would be acceptable, as long as a finite number of experiments could serve to decide which, if any, of the remaining discrete collection of possible spacetimes matches the real world. Definite explanations and predictions could then be made, and tested definitively against existing knowledge and future experiments.

Even if all goes well, even if effects can be found in the lambda model that concentrate the *a priori* measure on a discrete set of macroscopic spacetimes, produce small particle masses, and that fix the spacetime coupling constant g_s^2 in the macroscopic spacetimes at a small value, it will of course still not be guaranteed that one of the resulting spacetime quantum field theories matches the real world. It will be necessary to check that one of the macroscopic quantum field theories produced by the lambda model matches in detail the standard model and the observed cosmology. Formal capabilities do not guarantee that a theory will be successful as physics.

The proposed theory of large distance physics, if it succeeds, will still be only approximate. It will be a weak coupling, semiclassical approximation, unless effective methods can be found to do strongly coupled two dimensional quantum field theory calculations in the lambda model. Moreover, the theory is intrinsically only approximate for $L < \infty$, because the renormalized general nonlinear model is taken away from the strict two dimensional continuum limit at $\Lambda^{-1} = 0$. The renormalization of the general nonlinear model is justified by the divergence of $L^2 = \ln(\Lambda/\mu)$. Renormalization is exact only in the limit of an infinitely wide gulf between the short two dimensional distance Λ^{-1} and the long two dimensional distance μ^{-1}. Renormalization of the general nonlinear model is only approximate when $\ln(\Lambda/\mu) < \infty$.

Feasible experiments in the real world are at values of L^2 larger than some extremely large number, at least 10^{28}, if the unit of distance is within a few orders of magnitude of the Planck length. I suspect that $\ln(\Lambda/\mu) \approx 10^{28}$ is close enough to the two dimensional continuum limit that the theory will be quite precise, intrinsically, at all spacetime distances accessible to experiment.

If successful, this theory of large distance physics will not be a fundamental theory of physics, but it will describe with sufficient accuracy everything that can be checked, unless and until experiments are able to probe physics at spacetime distances approaching the Planck length. If spacetime quantum field theory is explained as merely an epiphenomenon of the lambda model, then quantum field theory, and quantum mechanics, will be seen to be effective descriptions of spacetime physics only at large distance, and will be seen to be inherently approximate except at infinite distance in spacetime. The possibility will then arise that theoretical physics in general might be inherently approximate at finite distance in spacetime. Alternatively, if the lambda model does succeed in giving a very accurate approximate description of large distance physics, it will become the touchstone for candidate exact theories of physics. The challenge will become to find more exact theories that have the lambda model as approximation, and to find experiments capable of distinguishing between the lambda model and any such candidate exact theories, in order to establish the greater reliability of a more exact candidate theory.

1.9 Many questions remain

Many questions about the theory remain. The most immediate questions concern calculations of the local properties of the spacetime quantum field theories produced by the lambda model. Do the lambda instantons that are local in a macroscopic spacetime succeed in removing spacetime supersymmetry, removing local gauge symmetries and generating nonperturbatively small particle masses? Are there effects in the lambda model that fix the spacetime coupling constant at a small numerical value? If the lambda model succeeds in doing these things, then the question becomes, is the verified part of the standard model to be found among the spacetime quantum field theories constructed by the lambda model?

A reasonable strategy is to assume a macroscopic spacetime, temporarily, in order to do the urgent local calculations in spacetime. Eventually, the existence and dimension of macroscopic spacetime must be settled by calculation in the lambda model. The lambda model has to explain the observation of macroscopic spacetime. The technical question is, do harmonic surfaces in the manifold of spacetimes succeed in concentrating the *a priori* measure at macroscopic spacetimes, of which some are four dimensional?

A theory of large distance physics must give a definite explanation of cosmology. It must explain the observed cosmological data, especially the essential features of the rich spectrum of characteristic distance scales that are found in the observed universe. But cosmology is still a diffuse, data rich subject compared to high energy particle physics. The standard model of particle physics is a sharp theoretical target. The standard model provides a definite, succinct *theoretical* structure to be explained, and a small set of precisely measured parameters to be calculated. A candidate theory of physics needs credibility before it can usefully take on the relatively nebulous theoretical problems of cosmology. The only reliable way I can see for a theory of physics to establish credibility is to explain the verified part of the standard model in detail. If that can be done in the lambda model, then the project of extracting cosmology from the lambda model will become promising.

An explanation of cosmology will require a construction of cosmological time. The general nonlinear model, to be well-defined as a two dimensional quantum field theory, needed its target manifold, spacetime, to be riemannian and compact. The general nonlinear model has to be well-defined in order that the manifold of general nonlinear models be usable as the target manifold of the lambda model. The problem is to construct real time.

A real time quantum field theory can be obtained from the *a priori* measure of the lambda model by making an ad hoc Wick rotation locally in spacetime, at distances where the spacetime curvature is insignificant, if a macroscopic spacetime is singled out by the lambda model. But cosmological time presumably needs a global construction that is everywhere consistent with local Wick rotation. I have no clear idea how this might be done. Perhaps there is a global analytic continuation of the manifold of spacetimes, which looks like Wick rotation locally in any macroscopic spacetime. Perhaps cosmological time can be related to the sliding large spacetime distance L, through the relation between L^2 and the logarithm of the characteristic two dimensional distance.

In the lambda model, where spacetime is taken to be riemannian for technical reasons, an *explanation* of real time is needed. There should be principle that explains *why* Wick rotation should be done locally in a macroscopic riemannian spacetime.

The basic question is the existence of the short distance limit in two dimensions, the limit at $\Lambda^{-1} = 0$. The lambda model is built as a two dimensional quantum field theory by integrating out the fluctuations at short two dimensional distances, starting from $\Lambda^{-1} = 0$. So the lambda model is well-defined as a two dimensional quantum field theory only if the limit exists. The short distance limit $\Lambda^{-1} = 0$ in two dimensions is the limit $L = \infty$ in spacetime, the limit in which only the spacetime zero modes fluctuate. The lambda model generates and controls the large distance spacetime physics, acting downward in spacetime distance from the limit at $L = \infty$. The theory is well-founded only if the $L = \infty$ limit exists.

My argument that the lambda model has a scale invariant limit at asymptotically short two dimensional distance is only formal. In the limit $\Lambda^{-1} = 0$, the fluctuating two dimensional fields of the lambda model are dimensionless. Their fluctuations can be expected to explore the entire manifold of spacetimes. Control over the global structure of the manifold of spacetimes will be needed before a rigorous argument can be made for the scale invariant short distance limit of the lambda model.

The question is probably not a practical one. For practical purposes, it is enough to construct the lambda model for sufficiently large finite values of L. The limit at $L = \infty$ is only needed to make the theory secure.

2. The structure of the theory

2.1 The lambda model

The lambda model is a two dimensional nonlinear model. Its target manifold is the manifold of renormalized general nonlinear models of the string worldsurface, which is the manifold of compact riemannian background spacetimes. The lambda model acts on the general nonlinear model at short two dimensional distances, determining the large distance physics of spacetime.

Consider a renormalized general nonlinear model of the worldsurface. The perturbations of this reference general nonlinear model are parametrized by the nearly dimensionless coupling constants λ^i. The λ^i are coupled to the approximately marginal spin 0 quantum fields $\phi_i(z,\bar{z})$, the fields that have scaling dimension near 2. The nearby general nonlinear models are made by inserting

$$\mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\, \lambda^i \phi_i(z,\bar{z})} \tag{2.1}$$

into the reference general nonlinear model. The two dimensional distance μ^{-1} is the distance at which the general nonlinear model is normalized. The coupling constants λ^i are local coordinates for the manifold of spacetimes. The λ^i are the large distance wave modes of the spacetime metric and other spacetime fields.

The lambda model is a nonlinear model whose field is a fluctuating map $\lambda(z,\bar{z})$ from the worldsurface to the manifold of spacetimes. In coordinates, the lambda field $\lambda(z,\bar{z})$ is expressed by component lambda fields $\lambda^i(z,\bar{z})$ which act as local sources coupled to the quantum fields of the reference general nonlinear model. The worldsurface is defined locally as a function of the lambda field by inserting

$$\mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\, \lambda^i(z,\bar{z}) \phi_i(z,\bar{z})} \tag{2.2}$$

into the reference general nonlinear model.

The map $\lambda(z,\bar{z})$ fluctuates at short two dimensional distances, from a two dimensional cutoff distance Λ_0^{-1} up to a sliding characteristic two dimensional distance Λ^{-1} which is still very much shorter than μ^{-1}. The fluctuations are described by a functional integral

$$\int \mathrm{D}\lambda\ \mathrm{e}^{-S(\lambda)}\ \mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\, \lambda^i(z,\bar{z}) \phi_i(z,\bar{z})} \tag{2.3}$$

inserted in the reference general nonlinear model. The action $S(\lambda)$ depends on the two dimensional distance. The action of fluctuations at two dimensional distance Λ^{-1} is

$$S(\Lambda,\lambda) = \int \mathrm{d}^2 z \frac{1}{2\pi} T^{-1} g_{ij}(\Lambda,\lambda)\, \partial\lambda^i\, \bar{\partial}\lambda^j\,. \tag{2.4}$$

The metric coupling $T^{-1}g_{ij}$ varies with the two dimensional distance Λ^{-1}.

At large spacetime distance, there are only a finite number of spacetime wave modes, because of spacetime being assumed compact and riemannian. So there are only a finite number of nearly marginal coupling constants λ^i in the general nonlinear model. So there are only a finite number of fields $\lambda^i(z,\bar{z})$ in the lambda model. The target manifold of the lambda model is finite dimensional.

2.2 Cancelling handles at short two dimensional distance

The lambda model is designed to cancel the effects of handles at short distance on the worldsurface. A handle attached to the worldsurface at two dimensional distance Λ^{-1} has the effect of a bi-local insertion

$$\frac{1}{2}\int \mathrm{d}^2 z_1\, \mu^2 \frac{1}{2\pi} \int \mathrm{d}^2 z_2\, \mu^2 \frac{1}{2\pi}\, \phi_i(z_1,\bar{z}_1)\, T g^{ij}(\Lambda,\lambda)\ \ln(\Lambda^2 |z_1 - z_2|^2)\ \phi_j(z_2,\bar{z}_2) \tag{2.5}$$

where z_1 and z_2 are the points where the two ends of the handle are attached to the worldsurface. The sum over indices i, j is the sum over states flowing through the handle. The fields $\phi_i(z_1, \bar{z}_1)$ and $\phi_j(z_2, \bar{z}_2)$ are produced in the worldsurface by the states flowing through the ends of the handle. The handle connects the fields at its two ends by the gluing matrix $Tg^{ij}(\Lambda, \lambda)$.

The bi-local insertion, equation (2.5), will calculated explicitly in section 3 below. For now, its form follows from general principles of two dimensional quantum field theory. The logarithmic dependence on the separation $|z_1 - z_2|$ between the two ends of the handle, for separations near Λ^{-1}, follows from the fact that the fields $\phi_i(z, \bar{z})$ are approximately marginal.

The metric coupling $T^{-1}g_{ij}(\Lambda, \lambda)$ of the lambda model is formulated as the inverse of the handle gluing matrix $T\, g^{ij}(\Lambda, \lambda)$. This formulation is designed so that the propagator of the lambda fields at two dimensional distance Λ^{-1} is

$$\left\langle\, \lambda^i(z_1, \bar{z}_1)\; \lambda^j(z_2, \bar{z}_2) \,\right\rangle = -Tg^{ij}(\Lambda, \lambda)\; \ln(\Lambda^2\, |z_1 - z_2|^2)\,. \tag{2.6}$$

The lambda model, equation (2.3), then produces the bi-local insertion

$$\frac{1}{2} \int \mathrm{d}^2 z_1\, \mu^2 \frac{1}{2\pi} \int \mathrm{d}^2 z_2\, \mu^2 \frac{1}{2\pi}\; \phi_i(z_1, \bar{z}_1)\; \left\langle\, \lambda^i(z_1, \bar{z}_1)\; \lambda^j(z_2, \bar{z}_2) \,\right\rangle\; \phi_j(z_2, \bar{z}_2) \tag{2.7}$$

which cancels the effects of the handle.

The properties that define the lambda model — its form as a two dimensional nonlinear model, its field as a map from the worldsurface to the manifold of spacetimes, its specific metric coupling — are all naturally determined by the short distance properties of the worldsurface, which determine the effects of handles at short distance in the worldsurface, which the lambda model is designed to cancel.

The number T^{-1} is the partition function of the worldsurface without handles, the 2-sphere. In a macroscopic spacetime of volume V,

$$T^{-1} = g_s^{-2}\, V \tag{2.8}$$

where g_s is the spacetime coupling constant. The metric coupling of the lambda model has a form that is local in spacetime,

$$T^{-1} g_{ij} = g_s^{-2}\, V g_{ij} \tag{2.9}$$

where $V g_{ij}$ is properly normalized so that it is expressible as the spacetime integral of the product of the corresponding spacetime wave modes. The coupling strength of the lambda model is therefore the spacetime coupling constant g_s.

2.3 Generalized scale invariance

The general nonlinear model is renormalizable, so it depends on the characteristic short distance Λ^{-1} only through the running coupling constants $\lambda_r^i(\Lambda/\mu, \lambda)$ which satisfy the renormalization group equation

$$\Lambda \frac{\partial}{\partial \Lambda}_{/\mu, \lambda} \lambda_r^i = \beta^i(\lambda_r)\,. \tag{2.10}$$

The running coupling constants couple to the two dimensional quantum fields $\phi_i^\Lambda(z,\bar{z})$ normalized at the short two dimensional distance Λ^{-1}. The general nonlinear model can be described at short distance by the insertion of the running couplings,

$$\mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\, \lambda^i \phi_i(z,\bar{z})} = \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi}\, \lambda_r^i \phi_i^\Lambda(z,\bar{z})} \tag{2.11}$$

obeying the renormalization group equation

$$\left(\Lambda \frac{\partial}{\partial \Lambda}_{/\lambda_r} + \beta^i(\lambda_r) \frac{\partial}{\partial \lambda_r^i} \right) \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi}\, \lambda_r^i \phi_i^\Lambda(z,\bar{z})} = 0\,. \tag{2.12}$$

The handle gluing matrix and its inverse matrix, the metric coupling $T^{-1}g_{ij}$, are natural structures of the worldsurface at two dimensional distance Λ^{-1}, so they depend on Λ^{-1} only through the running coupling constants λ_r. The action of the lambda model therefore depends only on the running sources $\lambda_r^i(z,\bar{z})$,

$$S(\Lambda,\lambda) = S(\lambda_r) \tag{2.13}$$

$$S(\lambda_r) = \int \mathrm{d}^2 z\, \frac{1}{2\pi}\, T^{-1} g_{ij}(\lambda_r)\, \partial\lambda_r^i\, \bar{\partial}\lambda_r^j \tag{2.14}$$

where the metric coupling $T^{-1}g_{ij}(\lambda_r)$ is independent of the two dimensional distance, as a function of the running coupling constants. The lambda model takes the same form at every short two dimensional distance Λ^{-1}

$$\int \mathrm{D}\lambda_r\, \mathrm{e}^{-S(\lambda_r)}\, \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi}\, \lambda_r^i(z,\bar{z}) \phi_i^\Lambda(z,\bar{z})}\,, \tag{2.15}$$

when it is expressed in terms of running fields $\lambda_r^i(z,\bar{z})$. The running fields transform, with an increase of the two dimensional distance $\Lambda^{-1} \to (1+\epsilon)\Lambda^{-1}$, by the renormalization group flow $\lambda_r^i \to \lambda_r^i - \epsilon\beta^i(\lambda_r)$.

The lambda model is therefore a scale invariant nonlinear model in the generalized sense [1, 2, 3]. The metric coupling $T^{-1}g_{ij}(\Lambda,\lambda)$, written in terms of the original renormalized coupling constants, is not literally invariant under a change of the characteristic two dimensional distance Λ^{-1}. Rather, the metric coupling is invariant under the combination of changing scale, $\Lambda^{-1} \to (1+\epsilon)\Lambda^{-1}$, and simultaneously flowing in the target manifold, $\lambda^i \to \lambda^i + \epsilon\beta^i(\lambda)$. The transformation of the target manifold is only a change of variables in the functional integral that defines the nonlinear model, so all observable quantities are scale invariant. The lambda model is novel in that its scale invariance is of the generalized kind even at the classical level.

2.4 The lambda model acts at short two dimensional distance

The lambda model is constructed, starting at a vanishingly short two dimensional cutoff distance Λ_0^{-1}, by integrating over more and more short distance fluctuations, up to the characteristic two dimensional distance Λ^{-1}. The characteristic two dimensional distance Λ^{-1} slides *outwards*, as in any two dimensional quantum field theory. The lambda model acts entirely at short two dimensional distance. The fluctuations at two dimensional distances up to Λ^{-1} act on the short distance structure of the general nonlinear model to produce an effective general nonlinear model at two dimensional distances longer than Λ^{-1}. The effective general nonlinear model of the worldsurface defines an effective string theory.

2.5 The general nonlinear model at short distance

Formally, the general nonlinear model is parametrized by infinitely many coupling constants λ^i, corresponding to infinitely many spacetime wave modes. But the general nonlinear model is renormalizable, so almost all of the coupling constants are irrelevant. The irrelevant coupling constants λ^i are coupled to irrelevant quantum fields $\phi_i(z,\bar{z})$. An irrelevant quantum field has negligible renormalized effect when inserted at short two dimensional distance. In particular, the irrelevant quantum fields that are inserted at short distance at the ends of handles have negligible renormalized effect. There is no need to cancel this negligible effect, so there is no need for the irrelevant coupling constants λ^i to fluctuate in the lambda model. It would not matter if the irrelevant λ^i did fluctuate. Their fluctuations would have negligible effect.

Consider a reference general nonlinear model that is scale invariant, satisfying $\beta = 0$. In coordinates around this reference point, the beta function takes the form

$$\beta^i(\lambda) = \gamma(i)\lambda^i + O(\lambda^2) \tag{2.16}$$

so

$$\lambda^i = (\mu\Lambda^{-1})^{\gamma(i)}\,\lambda^i_r \tag{2.17}$$

up to higher order corrections. Each coupling constant λ^i has definite scaling dimension $-\gamma(i)$. The corresponding two dimensional quantum field $\phi_i(z,\bar{z})$ is a scaling field with scaling dimension $2+\gamma(i)$. The number $\gamma(i)$ is the anomalous dimension of the field ϕ_i.

All coupling constants with $\gamma(i) > 0$ are irrelevant in the extreme short distance limit $\Lambda^{-1} = 0$. Their effects are driven to zero at the long two dimensional distance μ^{-1}. They have no effect in the renormalized two dimensional quantum field theory.

If a coupling constant λ^i had $\gamma(i) < 0$, it would be a relevant coupling constant. But a relevant coupling constant in the general nonlinear model of the string worldsurface would correspond to a tachyonic spacetime wave mode. There are no relevant coupling constants in a sensible string worldsurface. That is, all the anomalous dimensions satisfy

$$\gamma(i) \geq 0\,. \tag{2.18}$$

At $\Lambda^{-1} = 0$, the general nonlinear model is parametrized by the marginal coupling constants, the λ^i with $\gamma(i) = 0$.

Because all the anomalous dimensions $\gamma(i)$ are nonnegative, the renormalization group flow drives the general nonlinear model to the submanifold of fixed points, the submanifold where $\beta = 0$. The manifold of scale invariant general nonlinear models is the attracting manifold for the renormalization group flow. Renormalization *forces* the general nonlinear model to be scale invariant, so there is no possible choice of spacetime besides the manifold of scale invariant general nonlinear models.

The sharp distinction between the irrelevant and the marginal coupling constants does not persist when Λ^{-1} is greater than zero. Define the number L by

$$L^2 = \ln(\mu^{-1}\Lambda)\,. \tag{2.19}$$

The suppression of irrelevant coupling constants is

$$\lambda^i = \mathrm{e}^{-L^2\gamma(i)}\,\lambda^i_r \tag{2.20}$$

so λ^i is irrelevant at two dimensional distance Λ^{-1} if $L^2\gamma(i) \gg 1$. The property of irrelevance changes with the short two dimensional distance Λ^{-1}, depending on the value of the number L.

At a macroscopic spacetime, the coupling constants λ^i in the general nonlinear model are the wave modes of spacetime fields, including the spacetime metric. For wave modes that are localized in the macroscopic spacetime, the numbers $\gamma(i)$ are the eigenvalues of covariant second order differential operators acting on the spacetime wave modes. For wave modes localized at spacetime distances where spacetime curvature is insignificant, the anomalous dimensions $\gamma(i)$, being quadratic in the spacetime derivatives, take the form

$$\gamma(i) = p(i)^2 + m(i)^2 \tag{2.21}$$

where $p(i)$ is the spacetime wave number and $m(i)$ is the spacetime mass of the wave mode λ^i.

The number L is therefore a spacetime distance. The manifold of spacetimes is parametrized by the spacetime wave modes that have wave number $p(i)$ and mass $m(i)$ not many times larger than $1/L$. These are the coupling constants λ^i that fluctuate in the lambda model, so the number L is the characteristic ultraviolet spacetime distance in the lambda model.

Each coupling constant λ^i is associated to a spacetime distance $L(i)$ given by

$$L(i)^2 = \gamma(i)^{-1}\,. \tag{2.22}$$

The coupling constant λ^i is irrelevant if $L(i)/L \ll 1$. To be specific, call λ^i irrelevant if $L(i)/L < 1/20$. With this definition, the irrelevant coupling constants are suppressed by drastic scaling factors $\mathrm{e}^{-L^2\gamma(i)}$, factors of e^{-400} or less. The effects of the irrelevant coupling constants can be omitted without significant loss of accuracy. The ratio $1/20$ is more or less arbitrary. The details of the definition of irrelevance do not matter, as long as enough accuracy is maintained.

The coupling constants that are not irrelevant have $\delta(i) < 400/L^2$. If L^2 is a large number, the non-irrelevant coupling constants are very nearly marginal. Their scaling dimensions $-\gamma(i)$ differ only slightly from zero. They might be called the *quasi-marginal* or *L-marginal* coupling constants. The coupling constants that are irrelevant at two dimensional distance Λ^{-1} might be called the *L-irrelevant* coupling constants.

The renormalization of the general nonlinear model decouples the irrelevant coupling constants. The decoupling is accomplished by defining the renormalized quasi-marginal coupling constants so that all effects of the L-irrelevant coupling constants are absorbed into the effects of the quasi-marginal coupling constants. This is the basic principle of renormalization in quantum field theory.

The coupling constants that describe the structure of the general nonlinear model at short two dimensional distance Λ^{-1} are the spacetime wave modes that describe spacetime at spacetime distances larger than L. The coupling constants that are spacetime wave modes at distances much smaller than L are decoupled. Thus the short distance structure of the renormalized general nonlinear model encodes the large distance structure of spacetime. The lambda model acts on the short distance structure of the general nonlinear model, constructing an effective short distance structure in two dimensions, thereby constructing the effective large distance structure of spacetime.

The principles of renormalization can be applied accurately to the general nonlinear model at the short distance Λ^{-1} as long as $L^2 = \ln(\Lambda/\mu)$ is a very large number, large enough to be effectively a divergence in the two dimensional field theory. Precisely how large is necessary will have to be settled by detailed calculations. It seems reasonable to assume, tentatively, that numbers on the order of $L^2 = 10^{24}$ or $L^2 = 10^{20}$ are more than large enough. If that is so, and if the unit of spacetime distance is within a few orders of magnitude of the Planck length, then L can be taken smaller than any distance practical for experiment, while $L^2 = \ln(\Lambda/\mu)$ still remains large enough for the lambda model to work.

Write $M(L)$ for the manifold of renormalized general nonlinear models at two dimensional distance Λ^{-1}. The manifold of renormalized general nonlinear models depends on L because the property of coupling constant irrelevance depends on the ratio $\mu\Lambda^{-1} = \mathrm{e}^{-L^2}$ between the short two dimensional distance Λ^{-1} and the long two dimensional distance μ^{-1}. The L-irrelevant coupling constants are the coupling constants whose scaling dimensions $-\gamma(i)$ are far from zero on the scale set by L^{-2}, say $\gamma(i)L^2 > 400$. The L-irrelevant coupling constants are decoupled by the renormalization of the general nonlinear model and are ignored, without significant loss of accuracy. The manifold $M(L)$ is parametrized by the coupling constants that are not L-irrelevant, the quasi-marginal coupling constants λ^i, those whose scaling dimensions are not far from zero on the scale set by L^{-2}, say $\gamma(i)L^2 < 400$. The structure of the general nonlinear model at two dimensional distance Λ^{-1} depends only on the quasi-marginal coupling constants λ^i. So these λ^i parametrize the manifold $M(L)$.

Each coupling constant λ^i is associated to a spacetime distance $L(i)$ by equation (2.22). The L-irrelevant coupling constants are the spacetime wave modes at spacetime distances $L(i)$ which are small on the scale set by L, say $L(i) < L/20$. The quasi-marginal coupling constants λ^i are the spacetime wave modes at spacetime distances that are not much smaller than L, say $L(i) > L/20$. The manifold $M(L)$ is parametrized by the spacetime wave modes at spacetime distances on the order of L and larger. $M(L)$ is the manifold of spacetimes at spacetime distances on the order of L and larger.

$M(\infty)$ is the manifold of scale invariant general nonlinear models, the exact solutions of the fixed point equation $\beta = 0$. In the strict continuum limit, the limit $\Lambda^{-1} = 0$, renormalization forces the general nonlinear model to lie in $M(\infty)$, because $M(\infty)$ is the stable attracting manifold of the renormalization group flow. There are no relevant coupling constants in $M(\infty)$, only marginal and irrelevant coupling constants.

$M(\infty)$ is the foundation on which all the manifolds $M(L)$ are built. Scaling fields $\phi_i(z,\bar{z})$ are constructed in a scale invariant general nonlinear model, a model in $M(\infty)$. The anomalous dimensions $\gamma(i)$ are calculated there. These calculations identify the quasi-marginal coupling constants λ^i, out of which the manifolds $M(L)$ are built.

JHEP10(2003)063

As Λ^{-1} increases from zero, as L decreases from infinity, the set of quasi-marginal coupling constants grows. As L becomes smaller, the spacetime wave modes at spacetime distances somewhat smaller than L become available as quasi-marginal coupling constants. The dimension of the manifold $M(L)$ increases.

The manifolds $M(L)$ are built up incrementally as L decreases. If $L > L'$, the manifold $M(L')$ is made by extending $M(L)$.

A general nonlinear model λ in $M(L)$ satisfies $\beta^i(\lambda) = 0$ in the directions of the L-irrelevant coupling constants λ^i. As a spacetime, λ satisfies the classical field equation $\beta = 0$ at spacetime distances smaller than L. The spacetime λ might be pictured as composed of spacetime regions or cells, each of linear size L, satisfying $\beta = 0$ inside each cell. The spacetime wave modes localized inside the cells are the L-irrelevant coupling constants. They are coupled to the L-irrelevant scaling fields $\phi_i(z, \bar{z})$. These are scaling fields because they see only the spacetime at distances smaller than L, where $\beta = 0$. The properties of the L-irrelevant scaling fields $\phi_i(z, \bar{z})$, including the anomalous scaling dimensions $\gamma(i)$, are calculated locally in the spacetime λ. The properties of an L-irrelevant scaling field localized inside a spacetime cell depend on the values of the quasi-marginal coupling constants only through the values of the spacetime fields in the neighborhood of the spacetime cell where the L-irrelevant scaling field is localized.

At each λ in $M(L)$, the L-irrelevant coupling constants are determined by calculations in the spacetime λ that are local on the spacetime distance scale set by L. In particular, the coupling constants that are L'-marginal but L-irrelevant are determined locally in spacetime. These are the additional coupling constants that parametrize the extension of $M(L)$ to $M(L')$ at the spacetime λ in $M(L)$.

In this way, the structure of the manifold of spacetimes is built from large spacetime distance towards smaller, as the renormalized general nonlinear model is built from short two dimensional distance towards longer. As the characteristic spacetime distance L decreases, the spacetime wave modes at the distances below L appear as ripples on the larger wave modes of spacetime.

For L larger than L', the manifold $M(L)$ is a submanifold in $M(L')$. It is the submanifold defined by the vanishing of the L-irrelevant coupling constants. It is the submanifold of spacetimes in $M(L')$ that solve the classical field equations $\beta = 0$ at spacetime distances smaller than L.

$M(L)$ is also a quotient manifold of $M(L')$. At two dimensional distance Λ^{-1}, the L-irrelevant coupling constants that extend $M(L)$ to $M(L')$ are decoupled from the quasi-marginal coupling constants that parametrize $M(L)$. Nothing of the structure of the general nonlinear model at two dimensional distance Λ^{-1} depends on the L-irrelevant coupling constants.

A general nonlinear model at particular values of the quasi-marginal coupling constants λ^i describes a string worldsurface at two dimensional distances longer than Λ^{-1}. The properties of the general nonlinear model of the worldsurface depend on the values of the quasi-marginal coupling constants λ^i. The quasi-marginal coupling constants parametrize the classical background spacetime in which the strings scatter.

The quasi-marginal coupling constants change only very slowly with the two dimensional distance, so the worldsurface is approximately scale invariant at distances longer than Λ^{-1}. String scattering amplitudes are computed by integrating scaling fields $\phi_i(z,\bar{z})$ over the worldsurface. The two dimensional distance Λ^{-1} acts as short distance cutoff in these worldsurface integrals. As a result, the string propagator has the cut off form

$$T\, g^{ij} \left[\frac{1 - \mathrm{e}^{-2L^2\gamma(i)}}{\gamma(i)} \right]. \tag{2.23}$$

The propagation of string modes associated to spacetime distances $L(i)$ larger than L is suppressed. The spacetime distance L acts as infrared spacetime cutoff in the string scattering amplitudes. The string scattering is taking place inside a spacetime region, or cell, of linear size L.

The manifold $M(L)$ of general nonlinear models at two dimensional distance Λ^{-1} is the manifold of classical background spacetimes where strings scatter within spacetime regions of size L. The amplitudes for string scattering within a spacetime cell are calculated by integrating over the worldsurface the L-irrelevant scaling fields that are localized inside the spacetime cell. The worldsurface integrals are cutoff at the short distance two dimensional distance Λ^{-1}. The spacetime cell of size L might be considered a hypothetical experimental region, within which strings are hypothetically scattered at spacetime distances smaller than L. The string scattering amplitudes are determined by the background spacetime in the spacetime neighborhood where the experiment takes place, which is parametrized by the quasi-marginal coupling constants.

The quasi-marginal coupling constants are not precisely dimensionless. Some have $\beta(\lambda) \neq 0$, so can act as sources and detectors for string modes at spacetime distance L. It should be possible, in principle, to use the quasi-marginal coupling constants in this way to describe the experimental apparatus for scattering experiments at spacetime distances smaller than L. The theoretical representation of nature is divided, for every large spacetime distance L, into a background spacetime at spacetime distances larger than L, containing observers measuring hypothetical string scattering amplitudes in the background spacetime at spacetime distances smaller than L.

From the point of view of an observer situated in spacetime at spacetime distance L, the condition $\beta(\lambda) = 0$ on the L-marginal coupling constants is the consistency condition for extending tree-level string scattering amplitudes to spacetime distances larger than L. It expresses the condition that the background spacetime is a classical vacuum at spacetime distances L and larger. When the tree-level string scattering calculations are extended to spacetime distances larger than L, the general nonlinear model of the worldsurface is probed at two dimensional distances shorter than Λ^{-1}. If the background spacetime were not a classical vacuum, if $\beta(\lambda)$ were not zero on the L-marginal coupling constants, then the running coupling constants would blow up at two dimensional distances shorter than Λ^{-1}. The worldsurface would be pathological at short two dimensional distance. From this point of view, the equation $\beta = 0$ is a consistency condition.

But if the classical spacetime is the renormalized general nonlinear model, then the condition $\beta(\lambda) = 0$ is inevitable. It is forced by the renormalization. The renormalized general nonlinear model at nonzero two dimensional distance Λ^{-1} derives from the renormalized model at $\Lambda^{-1} = 0$, where necessarily $\beta(\lambda) = 0$ because there are no relevant coupling constants. Spacetime is necessarily a classical vacuum. From this point of view, the equation $\beta = 0$ holds necessarily. In this classical picture, there does not seem to be any room for an observer, in the absence of fluctuations.

2.6 What the cancelling does

The condition $\beta(\lambda) = 0$ on the quasi-marginal coupling constants allows the tree-level string scattering amplitudes to be extended to larger spacetime distances. But string loop corrections are divergent in the spacetime infrared. The divergence is logarithmic in the short two dimensional distance Λ^{-1}. Because of the divergence, the two dimensional cutoff cannot be removed. The spacetime infrared cutoff cannot be relaxed.

The divergence, equation (2.5), is a bi-local insertion at short two dimensional distance. It disturbs the short distance structure of the general nonlinear model, disturbing the selected background spacetime. The divergence signals that the possible background spacetimes are not properly determined, at any large spacetime distance L. A mechanism is missing that will determine the background spacetime so as to nullify the effects of the string loop divergence at short two dimensional distances.

The lambda model does this. The fluctuating lambda fields cancel the string loop effects at short two dimensional distance, eliminating the dependence on Λ^{-1}. The lambda model acts at two dimensional distances from Λ_0^{-1} up to Λ^{-1}, producing an effective general nonlinear model of the worldsurface at each two dimensional distance between Λ_0^{-1} and Λ^{-1}. These are effective background spacetimes at every spacetime distance from L_0 down to L, where L_0 is defined by

$$L_0^2 = \ln\left(\frac{\Lambda_0}{\mu}\right). \tag{2.24}$$

The cancelling implies that infrared string loop corrections do not ever have to be calculated, because their effects are already built into the effective background spacetimes produced by the lambda model. Infrared string loop corrections, starting at spacetime distance L in the effective background spacetime, would merely undo what the lambda model already did when it built the effective background spacetime from larger spacetime distance down to L. Because of the cancelling, the infrared string loop corrections would disturb the effective background spacetime at spacetime distance L exactly so as to produce the effective background spacetimes at spacetime distances larger than L. The lambda model produces the effective background spacetime depending on L exactly so it nullifies the infrared divergent string loop corrections.

The lambda model thus acts autonomously at short two dimensional distances, from a very short two dimensional cutoff distance Λ_0^{-1} up to Λ^{-1}. It acts autonomously at large spacetime distances from a very large infrared spacetime cutoff distance L_0 down to L. The structure of the effective general nonlinear model of the worldsurface is determined, at any short distance Λ^{-1}, entirely by two dimensional nonlinear model calculations in the lambda model. Handles are dispensed with completely, at two dimensional distances shorter than Λ^{-1}. Infrared string loop corrections are dispensed with, at spacetime distances larger than

L. String theory is used only at spacetime distances smaller than L. There is no practical use for string theory, if L can be pushed smaller than any observable spacetime distance.

Moreover, the lambda model makes sense nonperturbatively, as a two dimensional nonlinear model, while string theory is formulated only perturbatively. The lambda model constructs the effective background spacetime nonperturbatively. It *defines* nonperturbative string theory, at large distance in spacetime. If a nonperturbative version of string theory did exist, then its infrared quantum corrections, accumulated from small distance to large, would undo the work of the lambda model. Even if a nonperturbative version of string theory did exist, there would be no need to use it at large spacetime distances.

2.7 The effective general nonlinear model

The lambda model constructs the effective general nonlinear model of the worldsurface by a local process in two dimensions, acting entirely at short distance. So the effective general nonlinear model at two dimensional distance Λ^{-1} is independent of the two dimensional cutoff distance Λ_0^{-1}. It depends only on effective coupling constants $\lambda_e^i(\Lambda/\Lambda_0, \lambda_r)$. The lambda model builds the effective coupling constants starting at the two dimensional cutoff distance, starting from the running coupling constants at that distance,

$$\lambda_e^i\left(\frac{\Lambda_0}{\Lambda_0}, \lambda_r\right) = \lambda_0^i = \lambda_r^i\left(\frac{\mu}{\Lambda_0}, \lambda\right). \tag{2.25}$$

The effective coupling constants λ_e^i are coupled to effective two dimensional fields $\phi_i^{\Lambda,e}(z,\bar{z})$ at two dimensional distance Λ^{-1}. The effective general nonlinear model is described by the insertion

$$\mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi} \lambda_e^i \phi_i^{\Lambda,e}(z,\bar{z})}. \tag{2.26}$$

It satisfies an effective renormalization group equation

$$\left(\Lambda \frac{\partial}{\partial \Lambda}_{/\lambda_e} + \beta_e^i(\lambda_e)\frac{\partial}{\partial \lambda_e^i}\right) \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi} \lambda_e^i \phi_i^{\Lambda,e}(z,\bar{z})} = 0. \tag{2.27}$$

The effective beta function β_e consists of the beta function of the general nonlinear model, β, plus corrections $\delta\beta$ generated by the lambda model,

$$\beta_e = \beta + \delta\beta. \tag{2.28}$$

2.8 The effective lambda model

The lambda fluctuations produce an effective lambda model as well as an effective general nonlinear model. The lambda model is itself a two dimensional nonlinear model, so it is renormalizable. The effective lambda model at two dimensional distance Λ^{-1} is described by the effective metric coupling $T^{-1}g_{ij}^e(\Lambda/\Lambda_0, \lambda_r)$.

The crucial principle is that the effective lambda model and the effective general nonlinear model of the worldsurface evolve in tandem as the two dimensional distance Λ^{-1} increases. The propagator of the effective lambda model cancels handles at distance Λ^{-1} in the effective worldsurface. The effective metric coupling $T^{-1}g_{ij}^e(\Lambda/\Lambda_0, \lambda_r)$ is the inverse of the effective handle gluing matrix.

JHEP10(2003)063

This principle of *tandem renormalization* follows from the cancelling of handles by lambda fluctuations over a finite range of two dimensional distances. The cancelling of handles by lambda fluctuations between two dimensional distances Λ_0^{-1} and Λ_1^{-1}, with $\Lambda_0^{-1} < \Lambda^{-1} < \Lambda_1^{-1}$, can be broken up into the cancelling between Λ_0^{-1} and Λ^{-1}, and the cancelling between Λ^{-1} and Λ_1^{-1}. The second cancelling can be expressed in terms of handles in the effective worldsurface and fluctuations in the effective lambda model, both at two dimensional distance Λ^{-1}. The second cancelling therefore implies that the effective metric coupling $T^{-1}g^e_{ij}(\Lambda/\Lambda_0, \lambda_r)$ is the inverse of the effective handle gluing matrix.

The effective metric coupling of the lambda model is identified with a natural structure in the effective general nonlinear model of the worldsurface, so it depends only on the effective coupling constants λ^i_e. It depends on Λ^{-1} only through the λ^i_e. Therefore the effective lambda model is scale invariant in the generalized sense, taking the same form

$$\int \mathrm{D}\lambda_e \, \mathrm{e}^{-S_e(\lambda_e)} \, \mathrm{e}^{-\int \mathrm{d}^2 z \, \Lambda^2 \frac{1}{2\pi} \, \lambda^i_e(z,\bar z)\phi_i^{\Lambda,e}(z,\bar z)} \tag{2.29}$$

$$S_e(\lambda_e) = \int \mathrm{d}^2 z \, \frac{1}{2\pi} \, T^{-1} g^e_{ij}(\lambda_e) \, \partial\lambda^i_e \, \bar\partial\lambda^j_e \tag{2.30}$$

at every short two dimensional distance Λ^{-1}.

2.9 The *a priori* measure

In a two dimensional nonlinear model, the fluctuations at two dimensional distances shorter than the characteristic distance Λ^{-1} distribute themselves over the target manifold of the model to form a measure $\mathrm{d}\rho(\Lambda, \lambda)$ on the target manifold, called the *a priori* measure [1, 2, 3]. The *a priori* measure summarizes the fluctuations at two dimensional distances shorter than Λ^{-1}.

The *a priori* measure is calculated as the one point expectation value at two dimensional distance Λ^{-1},

$$\int \mathrm{d}\rho(\Lambda, \lambda) \, f(\lambda) = \langle \, f(\lambda(z, \bar z)) \, \rangle \tag{2.31}$$

where the expectation value is calculated by integrating over the lambda fluctuations at two dimensional distances up to Λ^{-1}. Equivalently,

$$\mathrm{d}\rho(\Lambda, \lambda') = \mathrm{dvol}(\lambda') \, \left\langle \, \delta(\lambda' - \lambda(z, \bar z)) \, \right\rangle \, . \tag{2.32}$$

The *a priori* measure of the lambda model is a measure on the manifold of spacetimes. Locally on the manifold of spacetimes, it takes the form of a measure on the spacetime wave modes λ^i, so it has the potential to be a quantum field theory in spacetime. The spacetime quantum field theory correlation functions of the wave modes λ^i are to be the integrals of functions $f(\lambda)$ with respect to the *a priori* measure on the manifold of spacetimes. So the spacetime correlation functions are to be calculated as the one point expectation values in the lambda model, as in equation (2.31).

In a two dimensional nonlinear model, the renormalization group acts on the target manifold of the model by a diffusion process [1, 2, 3]. As the characteristic two dimensional distance Λ^{-1} increases, additional fluctuations are included in the *a priori* measure, causing the it to diffuse outwards in the target manifold. The metric governing the diffusion is the effective metric coupling. When the nonlinear model is scale invariant, the diffusion process has stationary coefficients. When the scale invariance is of the generalized kind, as in the lambda model, the stationary diffusion process is driven by the flow in the target manifold that provides the scale invariance.

With each infinitesimal increase in the two dimensional distance, $\Lambda^{-1} \to (1+\epsilon)\Lambda^{-1}$, the *a priori* measure diffuses outwards because of the additional fluctuations. At the same time, the effective coupling constants flow, $\lambda_e^i \to \lambda_e^i - \epsilon\beta_e^i(\lambda_e)$, along the driving vector field $-\beta_e^i$.

Writing the *a priori* measure in the variables λ_e^i as $\mathrm{d}\rho_e(\Lambda, \lambda_e)$, the driven diffusion process is

$$-\Lambda \frac{\partial}{\partial\Lambda}_{/\lambda_e} \mathrm{d}\rho_e(\Lambda, \lambda_e) = \nabla_i^e \left(T\, g_e^{ij}(\lambda_e)\nabla_j^e + \beta_e^i(\lambda_e)\right) \mathrm{d}\rho_e(\Lambda, \lambda_e) \tag{2.33}$$

where ∇_i^e is the covariant derivative with respect to the effective metric $T^{-1}g_{ij}^e$. The coefficients, $T\, g_e^{ij}$ and β_e^i, of the diffusion process are stationary, independent of Λ^{-1}, because of the generalized scale invariance of the effective lambda model.

In the very long diffusion time $\ln(\Lambda_0/\Lambda)$, the *a priori* measure converges to the equilibrium measure $\mathrm{d}\rho_{eq}(\lambda_e)$ of the stationary diffusion process, no matter what its initial value at Λ_0^{-1}. The equilibrium *a priori* measure is peaked near the attracting submanifold where $\beta_e = 0$. If the equilibrium *a priori* measure concentrates near a macroscopic spacetime, then the *a priori* measure, as a measure on the wave modes of spacetime fields, is potentially a quantum field theory in spacetime, uniquely produced by the lambda model.

Assuming that the lambda model makes only small corrections to β, the *a priori* measure is driven by $-\beta_e$ first to the submanifold where $\beta = 0$. Then the corrections to β determine the subset of the $\beta = 0$ submanifold at which the *a priori* measure actually concentrates.

The *a priori* measure of the lambda model is generally covariant in spacetime, the target manifold of the general nonlinear model, because the renormalization of the general nonlinear model is carried out in a way that is invariant under reparametrization of its target manifold [1, 2, 3].

If a particular background spacetime is chosen by hand, arbitrarily, then it serves as the initial condition for the diffusion of the *a priori* measure. By the time a finite value of Λ^{-1} is reached, the *a priori* measure has diffused to the equilibrium measure, erasing all dependence on the initial choice of spacetime. The lambda model dynamically produces independence from the arbitrary choice of background spacetime.

2.10 The spacetime action principle

The beta function of the general nonlinear model, $\beta^i(\lambda)$, is a gradient vector field on the manifold of general nonlinear models [1, 2, 3, 9, 10, 11, 12]. In the proof of the gradient property [11, 12], the beta function is expressed as a gradient with respect to an intrinsic metric on the manifold of two dimensional quantum field theories, defined by the two point correlation function of the fields $\phi_i(z, \bar{z})$ on the plane. The metric coupling of the lambda model, $T^{-1}g_{ij}(\lambda)$, defined as the inverse of the handle gluing matrix, is that same intrinsic metric, multiplied by T^{-1}. So the beta function of the general nonlinear model is the gradient with respect to the metric coupling $T^{-1}g_{ij}(\lambda)$

$$T^{-1}g_{ij}(\lambda)\,\beta^j(\lambda) = \frac{\partial}{\partial\lambda^i}\,T^{-1}a(\lambda) \tag{2.34}$$

of the *potential function* $T^{-1}a(\lambda)$ on the manifold of spacetimes.

In a macroscopic spacetime of volume V,

$$T^{-1}a(\lambda) = g_s^{-2}\,V\,a(\lambda) \tag{2.35}$$

where $V\,a(\lambda)$ is properly normalized so as to be the spacetime integral of a local functional of the spacetime wave modes λ^i. The potential function $T^{-1}a(\lambda)$ is the spacetime action of the massless spacetime field theory whose scattering amplitudes are the same as the tree-level string scattering amplitudes at large distance [13, 10].

The gradient property will be derived as well for the effective beta function of the effective general nonlinear model,

$$T^{-1}g^e_{ij}(\lambda_e)\,\beta^j_e(\lambda_e) = \frac{\partial}{\partial\lambda^i_e}\,T^{-1}a_e(\lambda_e)\,. \tag{2.36}$$

The stationary diffusion process for the *a priori* measure is therefore driven by a gradient flow. The equilibrium measure then satisfies the first order differential equation

$$0 = \left(T\,g_e^{ij}(\lambda_e)\nabla^e_j + \beta^i_e(\lambda_e)\right)\mathrm{d}\rho_{eq}(\lambda_e) \tag{2.37}$$

whose solution is

$$\mathrm{d}\rho_{eq}(\lambda_e) = \mathrm{dvol}_e(\lambda_e)\;e^{-T^{-1}a_e(\lambda_e)} \tag{2.38}$$

where $\mathrm{dvol}_e(\lambda_e)$ is the metric volume element associated to the effective metric coupling.

The first order equation (2.37) for the *a priori* measure is the equation of motion $\beta^i_e(\lambda_e) = 0$ in the sense of spacetime quantum field theory. If the lambda model is successful, spacetime quantum field theory will be explained as the *a priori* measure of the lambda model. The quantum action principle of spacetime physics will then derive from the gradient property of the beta function of the general nonlinear model.

The *a priori* measure of the lambda model is nontrivial even at the classical level, because the lambda model is scale invariant in the generalized sense even at the classical level. The effective potential function is the classical potential function plus corrections generated by the lambda fluctuations

$$T^{-1}a_e = T^{-1}a + \delta(T^{-1}a)\,. \tag{2.39}$$

Before quantum corrections are taken into account, the *a priori* measure and the diffusion process are written in terms of the uncorrected running coupling constants λ^i_r, around a spacetime satisfying $\beta = 0$, to leading order in the λ^i_r,

$$-\,\Lambda\frac{\partial}{\partial\Lambda}_{/\lambda_r}\,\mathrm{d}\rho_r(\Lambda,\lambda_r) = \partial_i\left(Tg^{ij}\partial_j + \gamma(i)\lambda^i_r\right)\mathrm{d}\rho_r(\Lambda,\lambda_r)\,. \tag{2.40}$$

The uncorrected potential function is

$$T^{-1}a(\lambda_r) = \frac{1}{2}T^{-1}g_{ij}\gamma(i)\lambda_r^i\lambda_r^j + O(\lambda_r^3)\,. \tag{2.41}$$

The uncorrected equilibrium *a priori* measure is

$$\mathrm{dvol}(\lambda_r)\ e^{-\frac{1}{2}T^{-1}g_{ij}\gamma(i)\lambda_r^i\lambda_r^j} \tag{2.42}$$

in the gaussian approximation. The equilibration time for the wave mode λ^i is $1/\gamma(i)$.

2.11 Complementarity between spacetime quantum field theory and string theory

For each value of the characteristic spacetime distance L, the lambda model produces two complementary descriptions of spacetime physics. The *a priori* measure describes the spacetime physics at distances larger than L as spacetime quantum field theory. The effective general nonlinear model of the worldsurface describes spacetime physics at distances smaller than L by effective string scattering amplitudes. Both descriptions apply in local regions of spacetime, on the scale of spacetime distance set by L. L is the characteristic ultraviolet distance in the effective spacetime quantum field theory, and the characteristic infrared distance in the effective string scattering amplitudes.

The lambda model constructs the *a priori* measure and the effective general nonlinear model so that the two descriptions agree at spacetime distance L, where both apply. By the tandem renormalization principle, the data of the effective lambda model matches the data of the effective general nonlinear model. The effective potential function on the wave modes λ_e^i at spacetime distance L will be the generating functional for the effective string scattering amplitudes at distance L, by a version of the argument [13] that connected the beta function $\beta^i(\lambda)$ of the general nonlinear model to the tree-level string scattering amplitudes at large distance.

Given that the effective string scattering amplitudes match the scattering amplitudes of the effective spacetime quantum field theory, and that the effective spacetime quantum field theory has to be produced by the lambda model before the effective string scattering amplitudes become nonperturbatively reliable, there does not seem to be any practical use for the string scattering amplitudes at any spacetime distance L large enough to be reached by the lambda model. On the other hand, there could be circumstances where the effective string theory would be useful as an alternate technical algorithm for calculating scattering amplitudes.

2.12 The fermionic spacetime modes

The target manifold of the lambda model is a graded manifold. The manifold of spacetimes has both bosonic and fermionic dimensions. The bosonic coupling constants λ^i are the wave modes of bosonic spacetime fields, the fermionic λ^i are the wave modes of fermionic spacetime fields. The *a priori* measure is a measure on the graded manifold of spacetimes, a quantum field theory of bosonic and fermionic fields in spacetime.

The lambda model needs a technical construction of the general nonlinear model of the worldsurface in which the bosonic and fermionic coupling constants λ^i occur on an equal footing. The bosonic λ^i must couple to bosonic fields $\phi_i(z,\bar{z})$, the fermionic λ^i to fermionic fields $\phi_i(z,\bar{z})$. The metric coupling $T^{-1}g_{ij}$ must be symmetric in the bosonic directions and antisymmetric in the fermionic directions. And the construction must accomodate a general compact riemannian spacetime.

A construction of the fermionic coupling constants λ^i and their antisymmetric metric coupling $T^{-1}g_{ij}$ is given below, in section 9. The construction is based on the locally supersymmetric string worldsurface, in which worldsurface reparametrization invariance is implemented by means of supersymmetric worldsurface ghost fields [14]. The odd parameters of the supersymmetric worldsurface are eliminated, and the bosonic worldsurface ghost fields are used to construct an ordinary scale invariant worldsurface that is spacetime covariant [15]. Bosonic and fermionic scaling fields $\phi_i(z,\bar{z})$ couple to the spacetime wave modes λ^i, which are correspondingly bosonic and fermionic.

The technical drawback of the covariant worldsurface is the picture ambiguity. The worldsurface scaling fields fall into infinitely many formally equivalent subspaces called pictures, distinguished by a discrete picture charge. The bosonic scaling fields have integer picture charge. The fermionic scaling fields have picture charge integer plus half. The canonical pictures are distinguished by the condition that the dimensions of the scaling fields are bounded below. These are the natural pictures to use in analyzing the effects of handles at short distance. For the bosonic scaling fields, there is exactly one canonical picture, which has picture charge -1. The metric is symmetric on the canonical bosonic picture. For the fermionic scaling fields, there are two canonical pictures, the pictures of charge $-1/2$ and $-3/2$. The metric pairs the two canonical fermionic pictures. In the sum over states flowing through a handle, one of the canonical fermionic pictures is inserted at one end of the handle, the other canonical fermionic picture at the other end of the handle. There is no single space of fermionic coupling constants λ^i with an antisymmetric metric coupling $T^{-1}g_{ij}$.

A small technical innovation is devised to put the canonical fermionic scaling fields effectively in a single picture that effectively has picture charge -1, and on which there is an antisymmetric metric $T^{-1}g_{ij}$. The bosonic and fermionic coupling constants λ^i then combine to form a single graded space, with a graded metric coupling $T^{-1}g_{ij}$.

When the bosonic and fermionic coupling constants λ^i are put on the same footing as coordinates for the graded manifold of spacetimes, the spacetime dynamics of the fermionic wave modes takes a nonstandard form. The wave operators acting on the fermionic wave modes are quadratic in the spacetime derivatives, like the wave operators acting on the bosonic wave modes. They are not the standard first order Dirac wave operators. The unphysical degrees of freedom of the fermionic spacetime fields are eliminated by gauge symmetries, leaving the standard physical content of the Dirac operators.

The lambda model needs the metric coupling $T^{-1}g_{ij}$ to be positive definite in the bosonic directions. Otherwise, there would be instability under short distance fluctuations of the bosonic lambda fields $\lambda^i(z,\bar{z})$. This positivity condition is not satisfied if the worldsurface contains a Ramond-Ramond sector, because the metric on that sector is the tensor product of two antisymmetric matrices, which is not positive definite. Only the heterotic string worldsurface [8] is without a Ramond-Ramond sector. For this purely technical reason, it appears that the lambda model can only work in the heterotic string worldsurface. The metric coupling $T^{-1}g_{ij}$ on the manifold of general nonlinear models of the heterotic worldsurface is positive definite in the bosonic directions, because there is no Ramond-Ramond sector and because spacetime is assumed riemannian.

2.13 Physics is built from large distance towards small

The degrees of freedom of the lambda model are the coupling constants λ^i of the renormalized general nonlinear model, varying locally in two dimensions as fields $\lambda^i(z,\bar{z})$. The target manifold of the lambda model at two dimensional distance Λ^{-1} is the manifold $M(L)$. The renormalized general nonlinear model at two dimensional distance Λ^{-1} provides the data on the target manifold $M(L)$, the metric $T^{-1}g_{ij}(\lambda_r)$ and the vector field $\beta^i(\lambda_r)$, which specify the couplings of the lambda model. The renormalized general nonlinear model provides the lambda model with its degrees of freedom and its interactions. The renormalized general nonlinear model is the raw material on which the lambda model works.

The coupling constants λ^i are the wave modes of spacetime. The renormalization of the general nonlinear model arranges the degrees of freedom λ^i over the range of short two dimensional distances Λ^{-1}, in a hierarchy organized according to spacetime distance L, following the formula $L^2 = \ln(\Lambda/\mu)$. At the shortest two dimensional distances, the degrees of freedom λ^i are at the largest spacetime distances. The renormalization of the general nonlinear model decouples the small distance spacetime wave modes at short two dimensional distance Λ^{-1}. As the two dimensional distance Λ^{-1} increases, as the spacetime distance L decreases, the renormalization introduces, as additional degrees of freedom, the spacetime wave modes at smaller spacetime distances on the distance scale set by L.

The hierarchy of degrees of freedom is put in place by the renormalization of the general nonlinear model, before the lambda model is set to work. The lambda model acts autonomously on the renormalized general nonlinear model, from short two dimensional distance towards long. The lambda model operates on the degrees of freedom λ^i as it finds them, distributed by the renormalization of the general nonlinear model across the range of short two dimensional distances. The lambda model begins to operate at the extremely short two dimensional cutoff distance $\Lambda_0^{-1} \approx 0$, seeing only the extreme infrared wave modes in spacetime. As the lambda model operates at longer two dimensional distance Λ^{-1}, it sees spacetime wave modes at decreasing spacetime distance L. At each stage, at each characteristic spacetime distance L, the lambda model can ignore the decoupled spacetime wave modes at distances smaller than L, without significant loss of accuracy. The lambda model never notices the small distance structure of spacetime.

Locality in spacetime is expressed by the hierarchy of degrees of freedom λ^i, distributed in the two dimensional distance Λ^{-1} according to the spacetime distance L. This form of locality in spacetime is left undisturbed by the lambda model. The lambda model produces corrections to the interactions at each two dimensional distance Λ^{-1}, at each characteristic spacetime distance L. Nonperturbative effects in the lambda model might possibly change the form in which the degrees of freedom effectively appear. But the lambda model does not change the distribution of the basic degrees of freedom λ^i in the two dimensional distance Λ^{-1}, and in the spacetime distance L. Nor does it change the decoupling of small distance spacetime wave modes, for every large value of L.

The *a priori* measure of the lambda model is the distribution of the fluctuations of the fields $\lambda^i(z,\bar{z})$ at two dimensional distances shorter than Λ^{-1}. It is a measure on the target manifold of the lambda model, so it is a measure, for each L, on the graded manifold of spacetimes, $M(L)$. The *a priori* measure is thus a measure on the bosonic and fermionic spacetime wave modes λ^i at spacetime distances larger than L. If this measure has appropriate technical properties, then it is a quantum field theory in spacetime, cut off in the ultraviolet at spacetime distance L, describing physics at spacetime distances larger than L.

The lambda model is a local two dimensional quantum field theory, so it is built starting from its short distance limit at $\Lambda^{-1} = 0$, and proceeding outwards to longer two dimensional distances Λ^{-1} by integrating over the short distance modes of the fields $\lambda^i(z,\bar{z})$.

As the two dimensional distance Λ^{-1} *increases*, the spacetime distance L *decreases*. Additional spacetime modes λ^i begin to fluctuate, at smaller and smaller spacetime distances. Each wave mode λ^i is at a characteristic spacetime distance $L(i)$. The mode λ^i starts fluctuating in the lambda model when L has dropped below a value $L_0(i)$ not much larger than $L(i)$, say $L_0(i) = 20L(i)$. The fluctuations of the two dimensional field $\lambda^i(z,\bar{z})$ are cut off at a short two dimensional distance $\Lambda_0(i)^{-1}$ given by $L_0(i)^2 = \ln(\Lambda_0(i)/\mu)$. The field $\lambda^i(z,\bar{z})$ fluctuates only at two dimensional distances Λ^{-1} longer than its individualized cutoff distance $\Lambda_0(i)^{-1}$.

The *a priori* measure of the lambda model evolves with the increasing two dimensional distance Λ^{-1}, starting from the cutoff distance Λ_0^{-1}. As Λ^{-1} increases, as L decreases, the spacetime wave modes λ^i at distance L begin to fluctuate, and are *integrated in* to the *a priori* measure. The *a priori* measure is effectively a delta function in the variable λ^i concentrated at $\lambda^i = 0$, until L drops below $L_0(i)$, until Λ^{-1} becomes longer than $\Lambda_0(i)^{-1}$. When L drops below $L_0(i)$, the *a priori* measure begins to diffuse outward in the spacetime wave mode λ^i.

The characteristic equilibration time of the variable λ^i is $1/\gamma(i) = L(i)^2$, according to equation (2.40). The available diffusion time, $L_0(i)^2 - L(i)^2$, is more than enough to allow λ^i to reach equilibrium well before L decreases from $L_0(i)$ to $L(i)$. The *a priori* measure at two dimensional distance Λ^{-1} therefore gives an accurate description of the physics at all spacetime distances greater than L, the physics of the spacetime wave modes λ^i at all spacetime distances $L(i) > L$. The assumption here is that the equilibration times estimated from the uncorrected diffusion process are not significantly different from the actual diffusion times in the effective diffusion process.

The lambda model thus builds its *a priori* measure, which is to be spacetime quantum field theory, starting from the large distance limit at $L = \infty$, by *integrating in* the spacetime wave modes at smaller and smaller spacetime distances L. Despite this top down method of construction, from large spacetime distance to small, the resulting quantum field theory appears local in spacetime. Spacetime locality is expressed in the functional integral formulation of quantum field theory by the possibility of integrating out the small distance wave

JHEP10(2003)063

modes of the spacetime fields without losing information about the functional measure on the wave modes at larger distances. The *a priori* measure of the lambda model is local in spacetime, in this sense, because integrating out the spacetime wave modes at small spacetime distance merely reverses the process of *integrating in* that was performed by the lambda model. Integrating out a small distance wave mode λ^i replaces the equilibrium *a priori* measure in that variable with the value of the integral over λ^i, multiplied by the delta function concentrated at $\lambda^i = 0$. This simply reverses the diffusion accomplished by the lambda model, which starts from the delta function concentrated at $\beta^i(\lambda) = 0$ and tends to the equilibrium measure, since diffusion conserves the total weight of the measure. The *a priori* measure at a larger spacetime distance $L > L'$ is recovered from the *a priori* measure at the smaller spacetime distance L', by integrating out the wave modes λ^i at spacetime distances $L(i)$ between L and L'.

This appearance of locality in the spacetime quantum field theory depends crucially on the fact that, as Λ^{-1} increases, the newly fluctuating fields $\lambda^i(z,\bar{z})$ at spacetime distances smaller than L do not disturb the equilibrium *a priori* measure on the wave modes at spacetime distances larger than L. As L decreases, the *a priori* measure stabilizes on the spacetime wave modes at distances larger than L. Stability at large distance is ensured by the decoupling of irrelevant coupling constants in the renormalization of the general nonlinear model. The spacetime wave modes λ^i at small distance on the scale set by L are relatively irrelevant coupling constants in the renormalized general nonlinear model. The small distance wave modes are decoupled from the large distance wave modes, which are less irrelevant. The fluctuations of the small distance wave modes have no effect on the *a priori* measure at distances larger than L, as a measure on the large distance spacetime wave modes.

The building of the *a priori* measure from short two dimensional distance Λ^{-1} towards longer, is based on the hierarchy of submanifolds $M(L) \rightarrow M(L')$, $L > L'$. As Λ^{-1} increases, the *a priori* measure diffuses outwards from the submanifold $M(L)$ in $M(L')$, along the L-irrelevant directions λ^i that parametrize the extension of $M(L)$ to $M(L')$.

The stability of the *a priori* measure on the large distance wave modes, its independence from the additional degrees of freedom that enter at relatively small spacetime distance, is based on the hierarchy of quotient manifolds $M(L') \rightarrow M(L)$. The decoupling by renormalization, expressed in the quotient structure, ensures that integration over the fibers of the quotient reverses the diffusion of the *a priori* measure, which gives the property of spacetime locality to the *a priori* measure as spacetime quantum field theory.

Because of the stability of the *a priori* measure on the large distance wave modes, it should not be necessary to calculate quantum corrections in the lambda model at all spacetime distances from $L = \infty$ down to L, in order to find the effective spacetime physics at distance L. That much calculation would not be practical. It would accumulate enormous amounts of information about very large distance physics that would not be relevant to the physics at spacetime distance L. The only new calculations that are needed in the lambda model at two dimensional distance Λ^{-1} are to be done at spacetime distances on the order of L. The effective interactions in the lambda model of the degrees of freedom at larger spacetime distances have already been calculated at two dimensional distances shorter than Λ^{-1}. The effective lambda model has already stabilized in the large distance degrees of freedom, so the large distance calculations do not need to be redone. Only the properties of the coupling constants λ^i, on the border between L-marginal and L-irrelevant, are in flux at two dimensional distance Λ^{-1}. What new calculations are needed can be made in local spacetime regions, on the spacetime distance scale set by L, which is where the borderline coupling constants λ^i are found.

Spacetime quantum field theory, as constructed by the lambda model, is local in spacetime, but it is not constructed from a microscopic dynamics. There is no guarantee, outside the perturbation expansion, that the dynamical laws of spacetime physics at large distance can be derived from the dynamical laws at small distance. The effective potential function $g_s^{-2}\, V\, a_e(\lambda_e)$ at spacetime distance L cannot be derived from the effective potential function at a smaller spacetime distance L', except perturbatively. The actual form of the effective degrees of freedom λ_e^i at spacetime distance L are not determined by spacetime quantum field theory effects acting on the degrees of freedom at spacetime distances smaller than L, but rather by nonperturbative two dimensional effects in the lambda model acting at two dimensional distances shorter than Λ^{-1}, so at spacetime distances larger than L. Perturbatively, the effective degrees of freedom and the effective potential function of the lambda model will be consistent with perturbative spacetime quantum field theory. But nonperturbative effects (such as quark confinement) will have to be found in the lambda model.

For the lambda model to be well-defined, its target manifold should be finite dimensional. Calculations requiring an infinite number of fields $\lambda^i(z,\bar{z})$ would be difficult to control. The target spacetime of the general nonlinear model is assumed to be compact and riemannian. It follows that the general nonlinear model has a discrete spectrum of scaling dimensions $\gamma(i)$. The renormalization of the general nonlinear model then suppresses the ultraviolet wave modes in spacetime by factors $e^{-L^2\gamma(i)}$, leaving only a finite number of quasi-marginal coupling constants λ^i to parametrize the target manifold $M(L)$. The target manifold $M(L)$ is therefore finite dimensional, at each point λ in $M(L)$. The lambda model is well-defined, at least for small fluctuations around any point λ in its target manifold, $M(L)$, the manifold of compact riemannian spacetimes.

For any finite L, no matter how large, it is possible that fluctuations in $M(L)$ will lead to spacetimes of linear size much larger than L, spacetimes which are macroscopic on the distance scale set by L. Fluctuations in the lambda model can lead to the places in $M(L)$ where spacetime grows arbitrarily large. The dimension of $M(L)$ would then increase without bound. In particular, there are harmonic surfaces $\lambda_H(z,\bar{z})$ in $M(L)$ that pass through such places, as described in section 11 below. Weak coupling nonperturbative effects in the lambda model will be dominated by harmonic surfaces in the target manifold, including such *decompactifying* harmonic surfaces. So it will be necessary to control the effects of the unbounded growth in the dimension of $M(L)$ on a decompactifying harmonic surface.

The potential difficulty should only arise when the limit $L \to \infty$ is studied. At any finite value of L, calculations in the lambda model see only spacetime regions that are bounded on the distance scale set by L. Only a finite number of quasi-marginal spacetime wave modes λ^i can fit into any such local spacetime region. The difficulty posed by the unbounded dimensionality of the target manifold is a problem of the extreme infrared in spacetime, relevant to the problem of sending the two dimensional cutoff distance Λ_0^{-1} to zero.

To finding the spacetime physics at finite distance L, the lambda model must be built from the two dimensional cutoff distance all the way out to Λ^{-1}. This would seem to require starting with control over the extreme infrared in spacetime, the limit $L \to \infty$ at zero two dimensional distance. But, in the most favorable circumstance, it might be possible to postpone the problem of controlling the limit $L \to \infty$. The limit $L \to \infty$ might become purely a question of principle. At any finite value of L, it might be possible to obtain the spacetime physics of a local region of spacetime, on the distance scale set by L, in terms of the spacetime fields over a bounded neighborhood of that region. The spacetime wave modes λ^i in the extreme infrared, however many there are, will only make their influence known through the values of the local spacetime fields in the spacetime neighborhood under study. The two dimensional cutoff distance can be removed, so far as the local spacetime physics is concerned, without needing to worry about the possibly infinite number of spacetime wave modes in the extreme infrared.

Formally, the scale invariance of the effective lambda model allows the two dimensional cutoff distance Λ_0^{-1} to be taken to zero. According to the principle of tandem renormalization, the effective lambda model and the effective general nonlinear model evolve together in the two dimensional distance Λ^{-1}. It follows that the effective lambda model is automatically scale invariant, in the generalized sense. This *tautological scale invariance* means that any divergent dependence on the two dimensional cutoff distance Λ_0^{-1} can be absorbed into a change of integration variable in the functional integral defining the lambda model. The lambda model is thus finite in the limit $\Lambda_0^{-1} \to 0$. This formal argument should be effective in calculations of spacetime physics at finite spacetime distance L, allowing the two dimensional cutoff to be removed in such calculations. But it will be necessary to control the extreme infrared spacetime wave modes in arbitrarily large spacetimes, before it will be possible to establish the tautological generalized scale invariance of the lambda model at asymptotically short two dimensional distance. The existence of the limit $L \to \infty$ will need to be established in order to make the foundation of the theory secure.

3. The infrared divergence in string theory

The infrared failure of string theory, which is due to the existence of a manifold of possible background spacetimes, expresses itself as a technical disease of the string worldsurface, a divergence at short two dimensional distance. Each degenerating handle in the string worldsurface depends logarithmically on the two dimensional cutoff distance. The logarithmic divergence is due to marginal scaling fields flowing as string states through the degenerating handle. The divergence is in the infrared in spacetime, because the marginal scaling fields of the general nonlinear model correspond to the zero modes of the spacetime fields that characterize spacetime. The marginal scaling fields of the general nonlinear model generate the infinitesimal variations in the manifold of spacetimes. The infrared divergence occurs precisely when there is infinitesimal continuous degeneracy of the possible background spacetimes. The divergence occurs because of the existence of a continuous manifold of spacetimes. A background spacetime cannot be chosen from the manifold of possible background spacetimes, because perturbative string theory is divergent in any one of the manifold of possible background spacetimes.

3.1 A degenerating handle

Suppose that a background spacetime is chosen, satisfying $\beta = 0$ at all spacetime distances. The string worldsurface is described by an exactly scale invariant general nonlinear model. Consider a degenerating handle in the scale invariant worldsurface.

A handle is a tube that connects the worldsurface to itself. Making a transverse cut through the tube displays the handle to be formed by gluing together the boundaries of two holes in the worldsurface. The two holes can be anywhere on the worldsurface. The handle degenerates when the two holes shrink, each to a single point. The limit is a node, consisting of two distinct points on the worldsurface identified together as a single point.

A degenerating handle is parametrized by its two endpoints on the worldsurface, z_1 and z_2, and by a complex number q. The absolute value of q measures the thickness of the handle. Each hole has radius $|q|^{1/2}$. The phase of q measures the twist imparted when the two holes in the worldsurface are glued to form the handle. The endpoints z_1 and z_2 are the centers of the two holes. Let w_1 and w_2 be complex coordinates for the two regions of the worldsurface where the holes are located. The first hole is formed by removing the disk $|w_1 - z_1| < |q|^{1/2}$ from the first region, the second hole by removing the disk $|w_2 - z_2| < |q|^{1/2}$ from the second region. The boundaries of the two resulting holes are identified by the equation $(w_1 - z_1)(w_2 - z_2) = q$. The result is an almost degenerate handle whose complex structure is parametrized by the two points, z_1 and z_2, and by the complex number q which lies near 0. At $q = 0$, the handle degenerates to a node. The two points z_1 and z_2 become identified together as a single point.

A degenerating handle can also be pictured as a long tube of length $-\ln|q|$ connecting the two regions of the worldsurface. The long tube is parametrized by the complex coordinate $u = \ln(w_1 - z_1) - \frac{1}{2}\ln q = -\ln(w_1 - z_1) + \frac{1}{2}\ln q$. In this view, the degenerating handle represents string states propagating between two regions of the worldsurface in the very long world time $-\ln|q|$, during which the string can explore the largest distances in spacetime.

A string scattering amplitude is calculated in perturbative string theory by integrating the partition function of the worldsurface with respect to all the parameters of its complex structure. The worldsurface partition function is non-singular in the integration parameters, except where a handle in the worldsurface degenerates to a node. Only there can the integral diverge.

3.2 The contribution of a degenerating handle

The contribution of a degenerating handle to the worldsurface partition function is made explicit by summing over a complete set of string states flowing through the handle. Each end of the handle is the boundary circle of a hole in the worldsurface. A string state flowing through an end of the handle shows itself on the worldsurface as a boundary condition on the boundary of the hole. A hole in the worldsurface with a boundary condition on the boundary of the hole is a local field in the worldsurface. The local fields in a scale invariant worldsurface are linear combinations of the scaling fields. Each sum over string states flowing through an end of the handle is a sum over scaling fields inserted at the point in the worldsurface where the end of the handle is attached. The integral over the phase of q eliminates the scaling fields of nonzero spin, leaving only spin 0 scaling fields at the ends of the handle.

Summing over string states flowing through a degenerating handle replaces the handle with a double insertion of scaling fields in the worldsurface,

$$\begin{aligned}&\frac{1}{2}\int \mathrm{d}^2 z_1\,\mu^2\frac{1}{2\pi}\int \mathrm{d}^2 z_2\,\mu^2\frac{1}{2\pi}\int \mathrm{d}^2 q\,\mu^4\frac{1}{2\pi}(\mu\,|q|^{1/2})^{-8}\times\\&\times\phi_i(z_1,\bar{z}_1)(\mu\,|q|^{1/2})^{2+\gamma(i)}Tg^{ij}(\mu\,|q|^{1/2})^{2+\gamma(j)}\,\phi_j(z_2,\bar{z}_2)\,.\end{aligned}\tag{3.1}$$

The $\phi_i(z,\bar{z})$ form a complete set of linearly independent spin 0 scaling fields, normalized at two dimensional distance μ^{-1}. The scaling dimension of the field ϕ_i is $2+\gamma(i)$, the anomalous dimension is $\gamma(i)$. The sums over indices i,j are sums over the string states flowing through the two ends of the handle. The summation convention is used for indices i,j in place of explicit sums. The scaling field $\phi_i(z_1,\bar{z}_1)$ represents the string state flowing through the handle at endpoint z_1. The scaling field $\phi_j(z_2,\bar{z}_2)$ represents the string state flowing through the handle at endpoint z_2. The endpoints z_1 and z_2 range over the entire worldsurface.

The factor $(\mu\,|q|^{1/2})^{2+\gamma(i)}$ scales the field ϕ_i from the circle of radius μ^{-1} to the circle of radius $|q^{1/2}|$. Similarly, ϕ_j is scaled by $(\mu\,|q|^{1/2})^{2+\gamma(j)}$.

The factor $(\mu\,|q|^{1/2})^{-8}$ is for two dimensional scale invariance. The gluing equation $(w_1-z_1)(w_2-z_2)=q$ is left invariant when w_1, w_2, z_1, z_2, and $q^{1/2}$ are simultaneously scaled by the same scaling factor, so the integral over the handle parameters must also be invariant.

There is an overall factor $1/2$ because the two ends of the handle are indistinguishable. The factor $1/2\pi$ in each two dimensional integral is conventional.

The matrix $T\,g^{ij}$ implements the gluing of the two boundary circles to form the handle, tying together the string states passing through the two ends of the handle. A more specific identification of the gluing matrix $T\,g^{ij}$ is given below, in section 4.5. One property is needed now, the fact that $T\,g^{ij}=0$ if $\gamma(i)\neq\gamma(j)$. This follows from scale invariance of the gluing process when the two holes are scaled inversely.

The handle degenerates at $q=0$. The integral over the parameter $|q|$ diverges near $q=0$ if and only if there are spin 0 scaling fields $\phi_i(z,\bar{z})$ with anomalous scaling dimension $\gamma(i)=0$. These are the *marginal* scaling fields, the scaling fields with scaling dimension exactly equal to 2. The spin 0 scaling fields are the possible infinitesimal variations

$$\epsilon\int \mathrm{d}^2 z\frac{1}{2\pi}\,\phi_i(z,\bar{z})\tag{3.2}$$

of the action of the general nonlinear model. The marginal scaling fields are the infinitesimal variations that preserve two dimensional scale invariance. The marginal scaling fields are the infinitesimal variations of the background spacetime. So a degenerating handle produces a divergence if and only if there is an infinitesimal continuous degeneracy in the set of possible background spacetimes.

The divergence is in the infrared in spacetime, because the marginal scaling fields correspond to the zero modes of the spacetime field equations $\beta = 0$. Also, a handle with thickness parameter q near 0 describes a string propagating for very long world time, exploring the largest distances in spacetime.

In order to regulate the perturbative string theory, integrals over worldsurface parameters are cut off at a short two dimensional distance Λ_0^{-1}. In particular, the radius of the hole at each end of a degenerating handle is bounded below by Λ_0^{-1}. The integral over q in equation (3.1) is regulated by the cutoff $|q|^{1/2} > \Lambda_0^{-1}$. The cutoff dependence is extracted by integrating up to some limit $|q|^{1/2} = \Lambda_1^{-1}$, where Λ_1^{-1} is a fixed worldsurface distance, independent of the cutoff. The cutoff dependent part of the handle insertion, integral (3.1), becomes

$$\int \mathrm{d}^2 z_1\, \mu^2 \frac{1}{2\pi} \int \mathrm{d}^2 z_2\, \mu^2 \frac{1}{2\pi} \phi_i(z_1, \bar{z}_1) T\, g^{ij} \left[\frac{(\mu^2 \Lambda_1^{-2})^{\gamma(i)} - (\mu^2 \Lambda_0^{-2})^{\gamma(i)}}{\gamma(i)} \right] \phi_j(z_2, \bar{z}_2)\,. \tag{3.3}$$

The expression

$$T\, g^{ij} \left[\frac{(\mu^2 \Lambda_1^{-2})^{\gamma(i)} - (\mu^2 \Lambda_0^{-2})^{\gamma(i)}}{\gamma(i)} \right] \tag{3.4}$$

is the string propagator of the large distance string modes, with the two dimensional short distance cutoff acting as infrared regularizer in spacetime. The infrared spacetime cutoff distance L_0 is given by

$$L_0^2 = \ln\left(\frac{\Lambda_0}{\mu}\right). \tag{3.5}$$

For small $\gamma(i)$, the string propagator behaves as

$$T\, g^{ij}\, \frac{1}{\gamma(i)} \tag{3.6}$$

until $\gamma(i)$ becomes smaller than L_0^{-2}. Then the pole is regularized, becoming

$$T\, g^{ij}\, \ln(\Lambda_0^2 \Lambda_1^{-2}) \tag{3.7}$$

a logarithm of the two dimensional cutoff distance Λ_0^{-1}. The two dimensional cutoff distance, by acting as the short distance cutoff in worldsurface integrals, cuts off the propagator of the string modes at infrared spacetime distance L_0.

If the spacetime is held fixed, the two dimensional cutoff distance Λ_0^{-1} can be taken so close to zero that there is cutoff dependence only when $\gamma(i) = 0$. The divergent part of the handle insertion, integral (3.1), then becomes

$$\frac{1}{2} \int \mathrm{d}^2 z_1 \mu^2\, \frac{1}{2\pi} \int \mathrm{d}^2 z_2 \mu^2\, \frac{1}{2\pi} \phi_j(z_2, \bar{z}_2) T\, g^{ij}\, \ln(\Lambda_0^2 \Lambda_1^{-2}) \phi_i(z_1, \bar{z}_1) \tag{3.8}$$

where now the indices i, j range only over the marginal scaling fields. They will continue to do so until further notice.

The restriction to marginal scaling fields $\phi_i(z, \bar{z})$ will have to be relaxed, because it will become untenable to assume that Λ_0^{-1} can be taken to zero with the spacetime held fixed.

3.3 The effects of the divergence at short distance

The divergence is a symptom of deficiency in the string worldsurface. The divergence signals that string theory is incomplete, that the string worldsurface is not adequately formulated. A mechanism is missing to cancel the divergence. The divergence is in the spacetime infrared, so the missing mechanism should operate at large distance in spacetime.

The renormalization of the general nonlinear model exhibits the large distance spacetime physics to be encoded in the short distance structure of the general nonlinear model of the worldsurface. So the missing mechanism should operate at short two dimensional distance. But a degenerating handle is not necessarily local on the worldsurface. A degenerating handle may connect two regions on the worldsurface which, in the absence of the handle, are distant from each other or even disconnected from each other. First, it is necessary to isolate the divergent effects of degenerating handles on the short distance structure of the general nonlinear model of the worldsurface. These are the effects on the large distance physics of spacetime. Then a mechanism can be designed to cancel the divergent effects at short two dimensional distance.

The short distance structure of the worldsurface is visible in an arbitrary local two dimensional neighborhood. So the short distance effects of degenerating handles are produced by those degenerating handles whose two endpoints lie in the same two dimensional neighborhood. These are the *local* handles. The missing mechanism can then be designed to cancel the divergence produced by the local degenerating handles.

The endpoints z_1 and z_2 of a local handle lie in the same two dimensional neighborhood. The two dimensional distance $|z_1 - z_2|$ is independent of the two dimensional cutoff distance, so the divergent effects of a local handle can be extracted naturally, by putting $\Lambda_1^{-1} = |z_1 - z_2|$ as upper bound on the handle thickness parameter $|q|^{1/2}$. The availability of the two dimensional distance $|z_1 - z_2|$ between the endpoints of a local handle allows the short distance effects of the local handle to be isolated naturally.

Substituting $|z_1 - z_2|$ for Λ_1^{-1} in equation (3.8), the divergent effects of a local handle are described by a bi-local insertion in the local two dimensional neighborhood

$$\frac{1}{2}\int \mathrm{d}^2 z_1 \mu^2 \,\frac{1}{2\pi}\int \mathrm{d}^2 z_2 \mu^2 \,\frac{1}{2\pi}\phi_j(z_2,\bar{z}_2) T\, g^{ij} \ln(\Lambda_0^2\, |z_1 - z_2|^2)\phi_i(z_1,\bar{z}_1)\,. \tag{3.9}$$

As mentioned, the indices i, j are now ranging only over the marginal scaling fields, but this restriction will have to be lifted, because it derives from the assumption that Λ_0^{-1} can be taken arbitrarily close to zero with the spacetime held fixed.

4. A local mechanism to cancel the divergence

4.1 The restricted lambda model

The lambda model is formulated to cancel the effects of local handles at short two dimensional distance. The construction of the lambda model will be formal, at first, because of the artificial restriction to marginal scaling fields $\phi_i(z,\bar{z})$ in the description of the divergent effects of handles.

To cancel the divergent bi-local insertion made by a local handle, the marginal coupling constants λ^i are made into local sources $\lambda^i(z,\bar{z})$ which are coupled to the marginal scaling fields $\phi_i(z,\bar{z})$ by inserting

$$\mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\, \lambda^i(z,\bar{z})\, \phi_i(z,\bar{z})} \tag{4.1}$$

into the general nonlinear model of the worldsurface. Then the sources $\lambda^i(z,\bar{z})$ are set fluctuating with gaussian propagator

$$\left\langle\, \lambda^i(z_1,\bar{z}_1)\; \lambda^j(z_2,\bar{z}_2)\,\right\rangle = -T\, g^{ij}\; \ln(\Lambda_0^2\, |z_1 - z_2|^2)\,. \tag{4.2}$$

The fluctuating sources $\lambda^i(z,\bar{z})$, coupled to the marginal scaling fields $\phi_i(z,\bar{z})$, produce at leading order the insertion

$$\frac{1}{2}\int \mathrm{d}^2 z_1 \mu^2\, \frac{1}{2\pi}\int \mathrm{d}^2 z_2 \mu^2\, \frac{1}{2\pi}\phi_j(z_2,\bar{z}_2)\,\left\langle\, \lambda^i(z_1,\bar{z}_1)\lambda^j(z_2,\bar{z}_2)\,\right\rangle\, \phi_i(z_1,\bar{z}_1) \tag{4.3}$$

which cancels the effects of a single local handle at two dimensional distances near Λ_0^{-1}. Exponentiated, the insertions of the lambda propagator cancel the effects of arbitrarily many independent local handles on the worldsurface. But multiple handles are independent only when widely separated on the worldsurface. The sub-leading effects of colliding local handles remain to be cancelled.

The gaussian fluctuations are generated by inserting into the general nonlinear model of the worldsurface a functional integral over the sources $\lambda^i(z,\bar{z})$

$$\int \mathrm{D}\lambda \mathrm{e}^{-\int \mathrm{d}^2 z\, \frac{1}{2\pi}\, T^{-1} g_{ij}\, \partial\lambda^i\, \bar{\partial}\lambda^j}\, \mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2\, \frac{1}{2\pi}\, \lambda^i(z,\bar{z})\phi_i(z,\bar{z})} \tag{4.4}$$

The sources $\lambda^i(z,\bar{z})$ have become dimensionless, massless, scalar quantum fields. The propagator of a massless scalar field in two dimensions is logarithmic, so must be normalized at a characteristic two dimensional distance, which is Λ_0^{-1}.

The sub-leading effects of colliding local handles are cancelled by making non-gaussian corrections to the fluctuations. The corrections are generated by adding interaction terms to the gaussian action

$$S(\lambda) = \int \mathrm{d}^2 z\, \frac{1}{2\pi}\left(T^{-1} g_{ij}\, \partial\lambda^i\, \bar{\partial}\lambda^j + T^{-1} g_{ij,k}\lambda^k\, \partial\lambda^i\, \bar{\partial}\lambda^j + T^{-1} g_{ij,kl}\lambda^k\lambda^l\, \partial\lambda^i\, \bar{\partial}\lambda^j + \cdots\right). \tag{4.5}$$

Only local interactions are needed to cancel the short distance effects of the local handles. A rough argument is that collisions between handles produce the insertions of scaling fields that are to be cancelled by the interactions, only local interactions are needed. A better argument is given later. The interaction terms must be scale invariant because the effects of the handles are given by scale invariant integrals over worldsurface parameters. The interaction terms must therefore all contain two derivatives of the dimensionless scalar fields $\lambda^i(z,\bar{z})$, multiplied by any number of scalar fields. Infinitely many such interaction terms are possible. The infinite number of coefficients are calculated, in principle, from the worldsurface integrals for multiple handles. Fortunately there is a much simpler way.

JHEP10(2003)063

The marginal coupling constants λ^i are parameters for the scale invariant perturbations of the reference general nonlinear model that was chosen initially. The λ^i are local coordinates on the manifold $M(\infty)$ of scale invariant general nonlinear models, which is the manifold of spacetimes. The sources $\lambda^i(z,\bar{z})$ are therefore the components of a map $\lambda(z,\bar{z})$ from the worldsurface to the manifold $M(\infty)$, written in coordinates.

The reference general nonlinear model is a point λ_1 in $M(\infty)$, the origin of the coordinate system, the point with coordinates $\lambda_1^i = 0$. If the sources were nonzero constants $\lambda^i(z,\bar{z}) = \lambda_2^i$, their effect would be to change the general nonlinear model to a nearby scale invariant general nonlinear model λ_2 in $M(\infty)$. The fluctuating sources $\lambda^i(z,\bar{z})$ describe spacetime fluctuating locally on the worldsurface.

Once spacetime is set fluctuating in two dimensions, the mechanism that cancels the divergence must operate within any local fluctuation. Within a local fluctuation, the worldsurface might be in a nearby spacetime λ_2 in $M(\infty)$. The fluctuating spacetime $\lambda(z,\bar{z})$ can be considered to be nearly constant, locally in two dimensions, because the fields $\lambda^i(z,\bar{z})$ are dimensionless, and their variations in two dimensions are suppressed in the functional integral.

Within a local fluctuation to a spacetime λ_2, the short distance effects of local handles are given by the handle gluing matrix $T\, g^{ij}(\lambda_2)$ of the general nonlinear model λ_2. Local fluctuations around λ_2 will be needed to cancel the effects of these local handles. The gaussian fluctuations around λ_2 will be governed by the metric $T^{-1} g_{ij}(\lambda_2)$ which is the inverse of the handle gluing matrix in the general nonlinear model λ_2. The non-gaussian corrections at λ_2 are given by an infinite series of interaction terms, as in equation (4.5), with coefficients $T^{-1} g_{ij,k}(\lambda_2)$, $T^{-1} g_{ij,kl}(\lambda_2)$, and so on.

To cancel the effects of local handles, once spacetime is set fluctuating locally in two dimensions, there must be a cancelling set of local fluctuations around each point λ in $M(\infty)$. For each point λ in $M(\infty)$, the cancelling mechanism is a functional integral over maps $\lambda(z,\bar{z})$ from the worldsurface to a coordinate neighborhood of λ in $M(\infty)$.

But the fluctuations around λ_2 are completely determined by the fluctuations around λ_1, and *vice versa,* since one set of fluctuations is obtained from the other simply by a translation of coordinates in $M(\infty)$. The cancelling mechanisms for two nearby points, λ_1 and λ_2, must be equivalent, under the dictionary that translates sources $\lambda^i(z,\bar{z})$ in λ_1 to equivalent sources in λ_2. The cancelling mechanisms must operate simultaneously, and coherently, in all the spacetimes λ in the manifold $M(\infty)$.

A coherent collection of such functional integrals over dimensionless scalar fields is a two dimensional nonlinear model [1, 2, 3]. The target manifold of the nonlinear model is the manifold $M(\infty)$. The field of the nonlinear model is a map $\lambda(z,\bar{z})$ from the worldsurface to $M(\infty)$. The metric coupling of the nonlinear model is completely determined by the gaussian fluctuations at each point λ in $M(\infty)$. The handle gluing matrix $T\, g^{ij}(\lambda)$ in each spacetime λ gives all the information needed to determine the higher order interactions of the fluctuations. The action of the nonlinear model is globally defined as a function of the map $\lambda(z,\bar{z})$ from the worldsurface to $M(\infty)$,

$$S(\lambda) = \int \mathrm{d}^2 z \, \frac{1}{2\pi} \, T^{-1} g_{ij}(\lambda) \, \partial\lambda^i \, \bar{\partial}\lambda^j \, . \tag{4.6}$$

The coherence condition on the local fluctuations avoids a laborious calculation of the effects of collisions of multiple handles in a fixed spacetime.

The local mechanism that is inserted to cancel the divergence is the functional integral

$$\int \mathrm{D}\lambda \mathrm{e}^{-S(\lambda)} \mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \, \frac{1}{2\pi}\, \lambda^i(z,\bar{z}) \phi_i(z,\bar{z})} \,. \tag{4.7}$$

This is a two dimensional nonlinear model whose target manifold is the manifold of spacetimes $M(\infty)$. The metric coupling is the natural metric $T^{-1} g_{ij}(\lambda)$ on the manifold of spacetimes, the inverse of the handle gluing matrix. The field is the *lambda field,* $\lambda(z,\bar{z})$. The small fluctuations around a given reference spacetime are described in coordinates by the *lambda fields,* $\lambda^i(z,\bar{z})$.

This nonlinear model completely accomplishes the cancelling of the short distance effects of the local handles. The argument for complete cancelation is based on the coherence condition over the manifold of spacetimes $M(\infty)$. Suppose that some part of the divergence is left uncancelled. For each spacetime λ in $M(\infty)$, the uncancelled divergence would be cancelled by additional interactions among the lambda fields $\lambda^i(z,\bar{z})$. The gaussian part of the divergence is already cancelled, by design, so the additional interactions must be at least tri-linear in the lambda fields $\lambda^i(z,\bar{z})$. There must be coherence of these remaining interactions as λ varies in $M(\infty)$, so the additional interactions must involve only derivatives of the lambda fields. Otherwise, varying λ would produce quadratic interaction terms. The additional interactions must be scale invariant in two dimensions. Finally, the additional interactions cannot increase at long two dimensional separations because the effects being cancelled are made by handles in collision. There are no interactions compatible with all these conditions.

The lambda fields $\lambda^i(z,\bar{z})$ are dimensionless scalar fields in two dimensions. Once they are set fluctuating, large fluctuations are inevitable. Locality in two dimensions requires that *all* configurations $\lambda(z,\bar{z})$ participate in the functional integral, not merely the configurations that can be represented as perturbations $\lambda^i(z,\bar{z})$ around a constant spacetime λ. The functional integral must contain field configurations $\lambda(z,\bar{z})$ that make large excursions in the manifold of spacetimes. The most interesting configurations will be those that wrap around nontrivial topological features in the manifold of spacetimes, producing semi-classical nonperturbative effects.

On the other hand, because the lambda fields are dimensionless scalars, every fluctuation can be regarded locally as almost constant in two dimensions. Every fluctuation, however large, can be regarded as pieced together out of locally almost constant fluctuations.

The action $S(\lambda)$ is defined by the coherent family of actions for small fluctuations around constant configurations of $\lambda(z,\bar{z})$. But $S(\lambda)$, as given by equation (4.6), is well-defined globally, for all maps $\lambda(z,\bar{z})$ to the manifold of spacetimes. It does not depend on any choice of reference point λ_1 in the target manifold $M(\infty)$, nor on any choice of coordinates for the target manifold. The global definition of the nonlinear model is equivalent to the local definition.

Also needed is a global construction of the worldsurface, as a functional of the map $\lambda(z,\bar{z})$. Each local two dimensional region can be constructed by inserting sources $\lambda^i(z,\bar{z})$ into some general nonlinear model, as in equation (4.1). The local regions can then be patched together, in principle, to construct the global worldsurface, as a function of the map $\lambda(z,\bar{z})$. But I do not give any effective method for doing such patching. Instead, in section 5, I point out a way to avoid worldsurface calculations entirely.

The cancelling mechanism described by equation (4.7) is the *restricted* lambda model. It is only a formal mechanism. It works only perturbatively, and only in a generic submanifold of the target manifold $M(\infty)$, and only at sufficiently short two dimensional distances Λ_0^{-1}. The formal nature of this mechanism is due to the assumption that the spacetime λ can be held fixed while the two dimensional cutoff distance Λ_0^{-1} is taken to zero. Once spacetime has been set fluctuating in two dimensions, this assumption becomes untenable. The two dimensional cutoff distance Λ_0^{-1} must be held fixed while the spacetime fluctuates. It is possible that fluctuations of $\lambda(z,\bar{z})$ will reach places in $M(\infty)$ where some scaling fields $\phi_i(z,\bar{z})$ become only slightly irrelevant. Some anomalous dimensions $\gamma(i)$ will becomve very small. The calculation of the divergence produced by a local handle must then be revised to include the insertions of slightly irrelevant scaling fields. To cancel the divergence, it will be necessary to extend the target manifold of the lambda model beyond $M(\infty)$.

4.2 Macroscopic spacetimes

As the spacetime λ fluctuates, it may come upon places within $M(\infty)$ where spacetime becomes large in some or all of its dimensions. In such a *macroscopic* spacetime, there are many large distance wave modes which are not zero modes, but which have $\gamma(i)$ very small, small enough that the corresponding scaling fields $\phi_i(z,\bar{z})$ are not suppressed by the factors $(\mu\Lambda_0^{-1})^{\gamma(i)}$ in the effects of a local handle. These scaling fields $\phi_i(z,\bar{z})$ are almost marginal. It is no longer possible to separate distinctly the marginal scaling fields from the irrelevant scaling fields in the analysis of the degenerating handle. It is no longer possible to ignore as cutoff independent the contribution of the irrelevant scaling fields, and describe the divergence entirely in terms of the marginal scaling fields. The slightly irrelevant scaling fields must be included in the analysis of the divergence.

Write the spacetime metric on the macroscopic dimensions of spacetime as $\frac{1}{\alpha'}h_{\mu\nu}(x)$, explicitly proportional to $1/\alpha'$. The linear size of the macroscopic spacetime spacetime goes as $(\alpha')^{-1/2}$. The macroscopic spacetime becomes infinitely large as $\alpha' \to 0$. The spacetime wave operators acting on the massless spacetime fields are proportional to α', up to corrections that are higher order in α'. The eigenvalues $\gamma(i)$ go to zero as α'. The characteristic spacetime distances $L(i)$ of the massless wave modes go as $(\alpha')^{-1/2}$. More and more of the massless wave modes λ^i become almost marginal coupling constants in the general nonlinear model.

As spacetime fluctuates in the lambda model, the parameter α' might approach zero. In the handle insertion, the coefficient of the almost marginal scaling fields would approach a logarithm of the two dimensional cutoff distance Λ_0^{-1}. Once the spacetime is set fluctuating at fixed two dimensional cutoff distance, it becomes impossible to separate the marginal coupling constants from the irrelevant coupling constants.

The manifold of spacetimes $M(\infty)$ can be completed, and made locally compact, by adding a set of points $M(\infty)_d$ which corresponds to the limits $\alpha' \to 0$, subject to some identifications in the remaining parameters of the spacetime metric, which lose their significance in the limit. The completed manifold of spacetimes is the union

$$\overline{M(\infty)} = M(\infty) \cup M(\infty)_d \,. \tag{4.8}$$

The submanifold $M(\infty)_d$ might be called the *locus of decompactification.* The macroscopic spacetimes are the points λ in $M(\infty)$ which lie near the locus of decompactification.

The prototype for this completion of $M(\infty)$ is the manifold of toroidal two dimensional spacetimes. A two dimensional spacetime torus has Kahler metric proportional to a complex number σ. The spacetime volume is $\mathrm{Im}(\sigma)$. When $\mathrm{Im}(\sigma)$ is large, the manifold of spacetimes is parametrized by the complex parameter $q = \mathrm{e}^{2\pi i \sigma}$. The locus of decompactification is the single point $q = 0$. The real part of σ loses significance in the general nonlinear model in the limit $q \to 0$.

Besides the macroscopic spacetimes, there are also exceptional submanifolds within $M(\infty)$ where some scaling fields that are generically irrelevant become marginal. Such a scaling fields has anomalous dimension $\gamma(i) = 0$ on the exceptional submanifold, but is *not* a tangent vector to the manifold $M(\infty)$. The beta function $\beta(\lambda)$ vanishes to first order in the coupling constants, but becomes nonzero at some higher order. Near the exceptional submanifold, the anomalous dimension $\gamma(i)$ is slightly larger than 0.

There are also combinations of these phenomena, places where spacetime becomes macroscopic in some dimensions and goes to an exceptional point in other dimensions. These are the circumstances under which large distance spacetime wave modes get small nonzero masses $m(i)$.

Call the *singular locus* the entire submanifold $\overline{M(\infty)}$ where some generically nonzero anomalous dimensions $\gamma(i)$ go to zero, where some generically irrelevant scaling fields $\phi_i(z, \bar{z})$ become marginal. The singular locus includes the locus of decompactification. Almost marginal coupling constants λ^i occur in the spacetimes which lie near the singular locus.

When the almost marginal coupling constants λ^i are set fluctuating, the manifold of spacetimes $M(\infty)$ is *thickened* near the singular locus. Near the singular locus, the manifold $M(\infty)$ is extended in the directions parametrized by the slightly irrelevant coupling constants λ^i.

4.3 Gaussian fluctuations of quasi-marginal sources

The short distance effects of a degenerating local handle must be re-calculated, to include the scaling fields that are slightly irrelevant. The reference spacetime is still assumed to be in $M(\infty)$. The general nonlinear model is still assumed to be scale invariant. Again, the integral over the handle thickness parameter $|q|$ in equation (3.1) is bounded below by the two dimensional cutoff $|q|^{1/2} > \Lambda_0^{-1}$, and above by the separation between the endpoints of the handle, $|q|^{1/2} < |z_1 - z_2|$. The complete short distance contribution of the local handle, after integrating over the parameter q, is the bi-local insertion

$$\frac{1}{2}\int \mathrm{d}^2 z_1\, \mu^2 \frac{1}{2\pi}\int \mathrm{d}^2 z_2\, \mu^2 \frac{1}{2\pi}\phi_i(z_1,\bar{z}_1) T\, g^{ij}\left[\frac{(\mu^2\,|z_1-z_2|^2)^{\gamma(i)}-(\mu^2\Lambda_0^{-2})^{\gamma(i)}}{\gamma(i)}\right]\phi_j(z_2,\bar{z}_2)\,. \tag{4.9}$$

The indices i, j now range over the entire collection of marginal and slightly irrelevant spin 0 scaling fields. The logarithmic divergence appears in the limit $\gamma(i)\to 0$.

To cancel the effects of the degenerating local handle, again insert local sources $\lambda^i(z,\bar{z})$ in the worldsurface,

$$\mathrm{e}^{-\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\lambda^i(z,\bar{z})\,\phi_i(z,\bar{z})} \tag{4.10}$$

and again set the sources fluctuating with a gaussian propagator, which now is

$$\langle\, \lambda^i(z_1,\bar{z}_1)\ \lambda^j(z_2,\bar{z}_2)\,\rangle = T\, g^{ij}\ \left[\frac{(\mu^2\Lambda_0^{-2})^{\gamma(i)}-(\mu^2\,|z_1-z_2|^2)^{\gamma(i)}}{\gamma(i)}\right]. \tag{4.11}$$

This gaussian propagator is scale invariant, given that λ^i has scaling dimension $-\gamma(i)$. Only the additive normalization constant depends on Λ_0^{-1}.

Even though the lambda fields are not all dimensionless, their fluctuations are still described by a two dimensional nonlinear model. But now the metric coupling of the nonlinear model depends on the two dimensional distance. At distances $|z_1-z_2|$ close to Λ_0^{-1}, the lambda propagator, equation (4.11), is approximately

$$\langle\, \lambda^i(z_1,\bar{z}_1)\lambda^j(z_2,\bar{z}_2)\,\rangle \approx -T\, g^{ij}\,(\mu\Lambda_0^{-1})^{2\gamma(i)}\ \ln(\Lambda_0^2\,|z_1-z_2|^2)\,. \tag{4.12}$$

This is the gaussian propagator of a nonlinear model with a metric coupling

$$T^{-1}g_{ij}(\Lambda_0) = (\Lambda_0\mu^{-1})^{2\gamma(i)}\,T^{-1}g_{ij} \tag{4.13}$$

that varies with the two dimensional distance Λ_0^{-1}.

Now consider the lambda propagator at a two dimensional distance Λ^{-1} longer than Λ_0^{-1} but still much shorter than μ^{-1}. For $|z_1-z_2|$ near Λ^{-1}, the lambda propagator is

$$\begin{aligned}\langle \lambda^i(z_1,\bar{z}_1)\lambda^j(z_2,\bar{z}_2)\,\rangle \approx{}& -T\, g^{ij}\,(\mu\Lambda^{-1})^{2\gamma(i)}\ln(\Lambda^2\,|z_1-z_2|^2)+\\ &+T\, g^{ij}\left[\frac{(\mu^2\Lambda_0^{-2})^{\gamma(i)}-(\mu^2\Lambda^{-2})^{\gamma(i)}}{\gamma(i)}\right].\end{aligned} \tag{4.14}$$

After the constant term is subtracted, this is the gaussian propagator of a nonlinear model with metric coupling

$$T^{-1}g_{ij}(\Lambda) = (\Lambda\mu^{-1})^{2\gamma(i)}\,T^{-1}g_{ij}\,. \tag{4.15}$$

The constant term

$$T\, g^{ij}\left[\frac{(\mu^2\Lambda_0^{-2})^{\gamma(i)}-(\mu^2\Lambda^{-2})^{\gamma(i)}}{\gamma(i)}\right] = T\, g^{ij}\left[\frac{\mathrm{e}^{-2L_0^2\gamma(i)}-\mathrm{e}^{-2L^2\gamma(i)}}{\gamma(i)}\right] \tag{4.16}$$

makes a contribution to the renormalization of the effective lambda model, and to the renormalization of the effective general nonlinear model. For spacetime wave modes λ^i at spacetime distances lying between L_0 and L, $L^2\gamma(i)$ is small and $L_0^2\gamma(i)$ is large, so the constant term is

$$-T\, g^{ij}\, \frac{1}{\gamma(i)}\,. \tag{4.17}$$

Except for the minus sign, this is the tree-level spacetime propagator for the wave modes at distances between L_0 and L. The minus sign is there because, as the two dimensional distance increases from Λ_0^{-1} to Λ^{-1}, the wave modes from spacetime distance L_0 down to L are being *integrated in*. String theory works in the opposite direction, from L to L_0. In string theory, spacetime wave modes are integrated *out* by making contractions using the spacetime propagator, with positive sign. The difference in sign expresses the cancelling between lambda fluctuations and worldsurface handles. Integrating out the spacetime wave modes from L up to L_0 *undoes* the integrating in that is done in the lambda model, goint from L_0 down to L.

When the characteristic two dimensional distance increases from Λ_0^{-1} to Λ^{-1}, the lambda propagator becomes

$$T\, g^{ij}\, (\mu\Lambda^{-1})^{2\gamma(i)} \left[\frac{1-(\Lambda^2\,|z_1-z_2|^2)^{\gamma(i)}}{\gamma(i)}\right] \tag{4.18}$$

while, in the handle insertion, the coefficient of the scaling fields

$$T\, g^{ij}\, (\mu\Lambda^{-1})^{2\gamma(i)} \left[\frac{(\Lambda^2\,|z_1-z_2|^2)^{\gamma(i)}-1}{\gamma(i)}\right]. \tag{4.19}$$

Comparing the two expressions shows that the string worldsurface is at two dimensional distances longer than Λ^{-1}, while the lambda model operates at two dimensional distances shorter than Λ^{-1}. The handle insertion makes sense in the regime $|z_1-z_2| > \Lambda^{-1}$, where handles contribute positively. In the short distance regime, $|z_1-z_2| < \Lambda^{-1}$, the handle insertion is defined only by analytic continuation. On the other hand, the lambda propagator makes sense for $|z_1-z_2| < \Lambda^{-1}$. There are fluctuations at all two dimensional distances y from Λ_0^{-1} up to Λ^{-1}. The fluctuations contribute positively to the lambda correlations when $y > |z_1-z_2|$

$$\left\langle\, \lambda^i(z_1,\bar z_1)\; \lambda^j(z_2,\bar z_2)\,\right\rangle = \int_{\Lambda_0^{-1}}^{\Lambda^{-1}} \mathrm{d}y \frac{2}{y}\theta(y-|z_1-z_2|) T\, g^{ij}\, (\mu y)^{2\gamma(i)}\,. \tag{4.20}$$

This formula for the lambda propagator parallels the integral over handle thickness. It exhibits the lambda fluctuations as a random process indexed by the two dimensional distance,

$$\lambda^i(z,\bar z) = \int_{\Lambda_0^{-1}}^{\Lambda^{-1}} \mathrm{d}y\; \lambda^i(y,z,\bar z) \tag{4.21}$$

$$\left\langle\, \lambda^i(y_1,z_1,\bar z_1)\; \lambda^j(y_2,z_2,\bar z_2)\,\right\rangle = \theta(y_1-|z_1-z_2|)\; \delta(y_1-y_2) T\, g^{ij}\, (\mu y_1)^{\gamma(i)}\, (\mu y_2)^{\gamma(j)}\,. \tag{4.22}$$

The lambda propagator is defined for $|z_1-z_2| > \Lambda^{-1}$ only by continuation.

The gaussian fluctuations are generated by inserting into the general nonlinear model of the worldsurface a functional integral over the sources $\lambda^i(z,\bar{z})$,

$$\int \mathrm{D}\lambda \; \mathrm{e}^{-\int \mathrm{d}^2 z \, \frac{1}{2\pi} T^{-1} g_{ij}(\Lambda)\, \partial\lambda^i \, \bar{\partial}\lambda^j} \, \mathrm{e}^{-\int \mathrm{d}^2 z \, \mu^2 \, \frac{1}{2\pi} \lambda^i(z,\bar{z})\phi_i(z,\bar{z})} \, . \tag{4.23}$$

This is the gaussian approximation to a nonlinear model whose metric coupling is explicitly scale dependent.

The gaussian fluctuations of the lambda field $\lambda^i(z,\bar{z})$ at two dimensional distance Λ^{-1} are suppressed by the factor

$$(\mu\Lambda^{-1})^{\gamma(i)} = \mathrm{e}^{-L^2\gamma(i)} \tag{4.24}$$

which is just the suppression of the irrelevant coupling constants in the renormalized general nonlinear model of the worldsurface. Only the quasi-marginal coupling constants fluctuate significantly, the coupling constants λ^i with $L^2\gamma(i)$ not larger than, say, 400. These are the coupling constants that parametrize the extension of $M(\infty)$ into $M(L)$. The spacetime wave modes are cut off in the ultraviolet. The spacetime wave modes λ^i fluctuate only at characteristic spacetime distances $L(i)$, given by $L(i)^2 = 1/\gamma(i)$, that are larger than the ultraviolet spacetime distance $L/20$.

The fluctuations of the quasi-marginal coupling constants λ^i extend the target manifold of the lambda model from $M(\infty)$ into $M(L)$. The target manifold becomes the manifold of spacetimes $M(L)$, which is a *thickening* of the manifold of spacetimes $M(\infty)$ near the locus of decompactification, and near the rest of the singular locus. The thickening is controlled in the spacetime ultraviolet at spacetime distance L. The control is in place before λ is set fluctuating, having been provided by the renormalization of the general nonlinear model. The thickening is suppressed away from the singular locus by the renormalization of the general nonlinear model. Away from the singular locus, the target manifold of the lambda model is simply $M(\infty)$. The meaning of *away from* is set by the spacetime distance L, which derives from the ratio $\mu\Lambda^{-1}$ of the two dimensional distances in the renormalization of the general nonlinear model. The meaning of *near the locus of decompactification* is set by the spacetime distance L. A macroscopic spacetime is a spacetime of linear size much larger than L.

Once the fluctuations of the quasi-marginal coupling constants λ^i extend away from $M(\infty)$ into $M(L)$, the cancelling mechanism must act in any spacetime λ in $M(L)$. The cancelling mechanism must act on worldsurfaces described by general nonlinear models that are not exactly scale invariant. In retrospect, it was an oversimplification to start the analysis of the string theory divergence in a spacetime λ_1 that was an exact solution of $\beta = 0$. At the nonzero two dimensional cutoff distance Λ_0^{-1}, the possible general nonlinear models form the manifold $M(L_0)$. The initial spacetime λ_1 should have been chosen from the manifold of spacetimes $M(L_0)$.

4.4 The full lambda model

Reconsider the gaussian fluctuations in a spacetime in $M(\infty)$, described by equation (4.23). Define running coupling constants

$$\lambda_r^i = (\mu^{-1}\Lambda)^{\gamma(i)}\lambda^i \tag{4.25}$$

which couple to scaling fields

$$\phi_i^\Lambda(z,\bar z) = (\mu\Lambda^{-1})^{2+\gamma(i)}\,\phi_i(z,\bar z) \tag{4.26}$$

which are normalized at two dimensional distance Λ^{-1}. The gaussian fluctuations at distance Λ^{-1} take the form

$$\int \mathrm{D}\lambda_r \mathrm{e}^{-\int \mathrm{d}^2 z\,\frac{1}{2\pi}T^{-1} g_{ij}\partial\lambda_r^i\bar\partial\lambda_r^j}\mathrm{e}^{-\int \mathrm{d}^2 z\,\Lambda^2\frac{1}{2\pi}\,\lambda_r^i(z,\bar z)\phi_i^\Lambda(z,\bar z)} \tag{4.27}$$

which is the same at every two dimensional distance Λ^{-1}. This is the generalized scale invariance of the lambda model, in the gaussian approximation.

The effects of the local handle also take the same form at every two dimensional distance Λ^{-1}, when the states flowing through the ends of a handle are represented by the scaling fields ϕ_i^Λ. The handle gluing matrix takes the same form $T\,g^{ij}$ at every two dimensional distance Λ^{-1}. The bi-local handle insertion for $|z_1 - z_2|$ near Λ^{-1} is

$$\frac{1}{2}\int \mathrm{d}^2 z_1\,\Lambda^2\frac{1}{2\pi}\int \mathrm{d}^2 z_2\,\Lambda^2\frac{1}{2\pi}\phi_i^\Lambda(z_1,\bar z_1)T\,g^{ij}\ln(\Lambda^2\,|z_1-z_2|^2)\phi_j^\Lambda(z_2,\bar z_2) \tag{4.28}$$

at every two dimensional distance Λ^{-1}.

Now consider a spacetime λ_1 in $M(L)$. The general nonlinear models in $M(L)$ near λ_1 are parametrized by the quasi-marginal coupling constants λ^i. The general nonlinear model near λ_1 is given by the insertion

$$\mathrm{e}^{-\int \mathrm{d}^2 z\,\mu^2\frac{1}{2\pi}\,\lambda^i\phi_i(z,\bar z)} \tag{4.29}$$

which is interpreted as a perturbation of the general nonlinear model λ_1 with coefficients $\lambda^i - \lambda_1^i$,

$$\mathrm{e}^{-\int \mathrm{d}^2 z\,\mu^2\frac{1}{2\pi}\,\lambda_1^i\phi_i(z,\bar z)}\,\mathrm{e}^{-\int \mathrm{d}^2 z\,\mu^2\frac{1}{2\pi}\,(\lambda^i-\lambda_1^i)\phi_i(z,\bar z)}\,. \tag{4.30}$$

The renormalized general nonlinear model at short two dimensional distance Λ^{-1} depends on Λ^{-1} only through the running coupling constants $\lambda_r^i(\Lambda/\mu,\lambda)$, which satisfy the full renormalization group equation

$$\Lambda\frac{\partial}{\partial\Lambda}_{/\mu,\lambda}\lambda_r^i = \beta^i(\lambda_r)\,. \tag{4.31}$$

The running coupling constants couple to the two dimensional quantum fields $\phi_i^\Lambda(z,\bar z)$ that are normalized at the short two dimensional distance Λ^{-1}. The general nonlinear models are equally well described by insertion of the running coupling constants

$$\mathrm{e}^{-\int \mathrm{d}^2 z\,\mu^2\frac{1}{2\pi}\,\lambda^i\phi_i(z,\bar z)} = \mathrm{e}^{-\int \mathrm{d}^2 z\,\Lambda^2\frac{1}{2\pi}\,\lambda_r^i\,\phi_i^\Lambda(z,\bar z)}\,. \tag{4.32}$$

The scale dependence of the renormalized general nonlinear model is expressed by the full renormalization group equation

$$\left(\Lambda\frac{\partial}{\partial\Lambda}_{/\lambda_r} + \beta^i(\lambda_r)\frac{\partial}{\partial\lambda_r^i}\right)\,\mathrm{e}^{-\int \mathrm{d}^2 z\,\Lambda^2\frac{1}{2\pi}\,\lambda_r^i\,\phi_i^\Lambda(z,\bar z)} = 0\,. \tag{4.33}$$

JHEP10(2003)063

Once the quasi-marginal coupling constants λ^i are set fluctuating locally in two dimensions, there will be local fluctuations into spacetimes where the general nonlinear model of the worldsurface is not scale invariant. The effect of a local handle in that region of the worldsurface is a bi-local insertion of local fields in the scale non-invariant general nonlinear model.

Consider a general nonlinear model λ_1 in $M(L)$. At two dimensional distances close to Λ^{-1}, the departure from scale invariance is slight, because the quasi-marginal coupling constants are nearly marginal. The departure from scale invariance becomes significant only over a range of two dimensional distances. The analysis of the effects of a local handle at $|z_1 - z_2|$ near Λ^{-1} is just as in an exactly scale invariant worldsurface. The general nonlinear model λ_1 is described at two dimensional distance Λ^{-1} by the running coupling constants

$$\lambda^i_{1,r} = \lambda^i_r(\mu\Lambda^{-1}, \lambda_1)\,. \tag{4.34}$$

The effects of the local handle are given by the bi-local insertion

$$\frac{1}{2}\int \mathrm{d}^2 z_1\, \Lambda^2 \frac{1}{2\pi} \int \mathrm{d}^2 z_2\, \Lambda^2 \frac{1}{2\pi} \phi^\Lambda_i(z_1, \bar{z}_1) T\, g^{ij}(\lambda_{1,r}) \ln(\Lambda^2\, |z_1 - z_2|^2) \phi^\Lambda_j(z_2, \bar{z}_2) \tag{4.35}$$

where $T\, g^{ij}(\lambda_{1,r})$ is the handle gluing matrix at two dimensional distance Λ^{-1} in the spacetime λ_1.

The gaussian mechanism that cancels the effects of the handle consists of sources $\lambda^i_r(z, \bar{z})$ fluctuating in two dimensions around the constant values $\lambda^i_{1,r}$,

$$\int \mathrm{D}\lambda_r\; \mathrm{e}^{-\int \mathrm{d}^2 z\, \frac{1}{2\pi} T^{-1} g_{ij}(\lambda_{1,r})\, \partial\lambda^i_r\, \bar{\partial}\lambda^j_r}\; \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi}\, \lambda^i_r(z,\bar{z})\, \phi^\Lambda_i(z,\bar{z})} \tag{4.36}$$

where the metric coupling $T^{-1} g_{ij}(\lambda_{1,r})$ is the inverse of the handle gluing metric at two dimensional distance Λ^{-1} in the spacetime λ_1.

It remains to patch together the collection of gaussian functional integrals, consistently, over the manifold of spacetimes $M(L)$. Again, the interactions are completely determined by the collection of gaussian functional integrals around the points λ_1 in $M(L)$, and by the condition that the interactions be coherent under shifting of the origin of coordinates in $M(L)$. Again, the model is a nonlinear model. The target manifold is $M(L)$. The fields of the nonlinear model are the maps $\lambda_r(z, \bar{z})$ from the worldsurface to the manifold of spacetimes $M(L)$.

The nonlinear model, the lambda model, is the functional integral over maps $\lambda_r(z, \bar{z})$

$$\int \mathrm{D}\lambda_r \mathrm{e}^{-S(\lambda_r)}\; \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi}\, \lambda^i_r(z,\bar{z})\, \phi^\Lambda_i(z,\bar{z})} \tag{4.37}$$

with action

$$S(\lambda_r) = \int \mathrm{d}^2 z\, \frac{1}{2\pi} T^{-1} g_{ij}(\lambda_r)\, \partial\lambda^i_r\, \bar{\partial}\lambda^j_r\,. \tag{4.38}$$

Again, the coherence condition on the local fluctuations avoids a laborious calculation of the effects of collisions of multiple handles in a fixed spacetime.

The lambda model is manifestly scale invariant in the generalized sense, as written in terms of the field $\lambda_r(z,\bar z)$. Re-written in terms of the field $\lambda(z,\bar z)$, the lambda model is

$$\int \mathrm{D}\lambda \mathrm{e}^{-S(\Lambda,\lambda)}\ \mathrm{e}^{-\int \mathrm{d}^2z\,\mu^2 \frac{1}{2\pi}\,\lambda^i(z,\bar z)\,\phi_i(z,\bar z)} \tag{4.39}$$

$$S(\Lambda,\lambda) = \int \mathrm{d}^2z\,\frac{1}{2\pi}\,T^{-1}g_{ij}(\Lambda,\lambda)\,\partial\lambda^i\,\bar\partial\lambda^j\,. \tag{4.40}$$

The metric coupling depends on the two dimensional distance Λ^{-1}, but only through a transformation of the target manifold

$$T^{-1}g_{ij}(\Lambda,\lambda) = \frac{\partial\lambda_r^k}{\partial\lambda^i}\,T^{-1}g_{kl}(\lambda_r)\frac{\partial\lambda_r^l}{\partial\lambda^j}\,. \tag{4.41}$$

The metric coupling satisfies the natural renormalization group equation

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+\beta_*\right)T^{-1}g_{ij}(\Lambda,\lambda)=0 \tag{4.42}$$

where

$$\beta_*(T^{-1}g_{ij}) = \beta^k\partial_k(T^{-1}g_{ij}) + (\partial_i\beta^k)\,T^{-1}g_{kj} + T^{-1}g_{ik}\,(\partial_j\beta^k) \tag{4.43}$$

is the infinitesimal change of the metric coupling under the flow generated by the vector field $\beta^i(\lambda)$.

The generalized scale invariance of the full lambda model is of course consistent with the generalized scale invariance of the gaussian fluctuations around a scale invariant general nonlinear model. There, the linearized beta function is

$$\beta^i(\lambda) = \gamma(i)\lambda^i + \cdots \tag{4.44}$$

and the infinitesimal equation for generalized scale invariance is

$$\left(\Lambda\frac{\partial}{\partial\Lambda}+\gamma(i)+\gamma(j)\right)T^{-1}g_{ij}(\Lambda)=0 \tag{4.45}$$

in the gaussian approximation.

The lambda model differs from the nonlinear models with generalized scale invariance as originally contemplated [1, 2, 3], in that the *classical* metric coupling of the lambda model depends nontrivially on the two dimensional distance. The flow on the target manifold, generated by the vector field $\beta^i(\lambda)$, is present already in the classical lambda model, instead of arising from the quantum corrections.

Given $\beta^i(\lambda)$, the metric coupling $T^{-1}g_{ij}(\Lambda,\lambda)$ can be determined entirely from its value at one specific short two dimensional distance, for example its value $T^{-1}g_{ij}(\Lambda_0,\lambda)$ at two dimensional distance Λ_0^{-1}, by integrating the renormalization group equation (4.42). The data that gives the couplings of the lambda model can be determined entirely at small two dimensional distance in the renormalized general nonlinear model. The long two dimensional distance μ^{-1} enters only in the determination of the target manifold $M(L)$, by the decoupling of the irrelevant coupling constants in the renormalization of the general nonlinear model.

The lambda model is formulated at each two dimensional distance Λ^{-1} as a nonlinear model whose metric coupling depends on Λ^{-1}. The metric coupling at two dimensional distance Λ^{-1} governs the fluctuations of the lambda field $\lambda(z,\bar{z})$ at that distance. The lambda model is built up incrementally in the two dimensional distance, from the cutoff Λ_0^{-1} to longer two dimensional distances Λ^{-1}, using the nonlinear model at each distance to make the next incremental step. The building of the lambda model expresses the fundamental principle of renormalization, that information propagates locally in the distance scale.

4.5 Identification of the metric coupling

The metric coupling $T^{-1}g_{ij}$ of the lambda model is defined as the inverse of the handle gluing matrix, $T\,g^{ij}$, because the lambda model is designed to cancel the effects of the local handles. But the handle gluing matrix is not a directly accessible object in the general nonlinear model of the worldsurface. For calculation, it is useful to express the handle gluing matrix in terms of more usual field theory quantities.

First consider a scale invariant general nonlinear model. Make a worldsurfurce by connecting a pair of 2-spheres to each other by a handle. This worldsurface is equivalent to a single 2-sphere. Place a scaling field $\phi_i(z_1,\bar{z}_1)$ in one of the 2-spheres, and a second scaling field $\phi_j(z_2,\bar{z}_2)$ in the other 2-sphere. Calculate the partition function function of the worldsurface, summing over scaling fields at the ends of the handle, at points z_3 and z_4. Schematically, the result is

$$Z\,\langle\,\phi_i(1)\;\phi_k(3)\,\rangle\,T\,g^{kl} Z\,\langle\,\phi_l(4)\;\phi_j(2)\rangle\;. \tag{4.46}$$

This can also be calculated as the partition function of the single 2-sphere containing the two scaling fields,

$$Z\,\langle\,\phi_i(1)\;\phi_j(2)\rangle\;. \tag{4.47}$$

The equivalence of the two calculations implies that the metric coupling is identical to the un-normalized two point expectation value of the scaling fields at separation $|z_1-z_2| = \mu^{-1}$,

$$T^{-1}g_{ij} = Z\,\langle\,\phi_i(z_1,\bar{z}_1)\;\phi_j(z_2,\bar{z}_2)\,\rangle\;. \tag{4.48}$$

The two point expectation value at arbitrary separation is

$$Z\,\langle\,\phi_i(z_1,\bar{z}_1)\phi_j(z_2,\bar{z}_2)\,\rangle = T^{-1}g_{ij}(\mu\,|z_1-z_2|)^{-2-\gamma(i)-2-\gamma(j)}\;. \tag{4.49}$$

The scale-dependent metric $T^{-1}g_{ij}(\Lambda)$ is given by the two point expectation value at separation $|z_1 - z_2| = \Lambda^{-1}$

$$\begin{aligned} T^{-1}g_{ij}(\Lambda) &= Z\,\langle\,\phi_i(z_1,\bar{z}_1)\phi_j(z_2,\bar{z}_2)\,\rangle\,(\mu^{-1}\Lambda)^4 && (4.50)\\ T^{-1}g_{ij} &= Z\left\langle\,\phi^{\Lambda}_i(z_1,\bar{z}_1)\;\phi^{\Lambda}_j(z_2,\bar{z}_2)\right\rangle\;. && (4.51)\end{aligned}$$

These un-normalized expectation values are normalized to geive ordinary expectation values. The normalizing factor is the partition function without insertions

$$Z\,\langle 1\rangle = T^{-1} \tag{4.52}$$

which is the factor T^{-1} in the metric coupling $T^{-1}g_{ij}$. The normalized metric g_{ij} is identical to the normalized two point expectation value

$$g_{ij} = \left\langle\, \phi_i^\Lambda(z_1,\bar z_1)\phi_j^\Lambda(z_2,\bar z_2)\right\rangle \tag{4.53}$$

at $|z_1 - z_2| = \Lambda^{-1}$. Equivalently, the normalized metric is the coefficient of the identity operator in the operator product expansion

$$\phi_i^\Lambda(z_1,\bar z_1)\phi_j^\Lambda(z_2,\bar z_2) = (\Lambda\,|z_1 - z_2|)^{-2-\gamma(i)-2-\gamma(j)} g_{ij}\, 1 + \cdots \,. \tag{4.54}$$

When the spacetime is macroscopic, it makes sense to calculate the volume V of spacetime, which is a large number. The spacetime coupling constant g_s in the macroscopic spacetime is given by

$$T^{-1} = g_s^{-2}\, V\,. \tag{4.55}$$

The number V is the factor in the partition function $Z\,\langle 1\rangle$ that comes from the integral over the zero mode of the worldsurface position $x^\mu(z,\bar z)$ in the macroscopic spacetime. The metric coupling at the macroscopic spacetime is written

$$T^{-1}g_{ij} = g_s^{-2}\, V\, g_{ij}\,. \tag{4.56}$$

The metric $V\, g_{ij}$ is properly normalized to be a local inner product on the wave modes of the spacetime fields, an integral over the macroscopic spacetime of a product of the two spacetime wave modes.

From the two dimensional point of view, the metric coupling should be written $T^{-1}g_{ij}$, because this form makes sense in the general spacetime, macroscopic or not. The form $g_s^{-2}\, V\, g_{ij}$ only makes sense when there is a macroscopic spacetime, in which case it is the appropriate form for expressing effects that are local in the macroscopic spacetime.

Now consider a general spacetime λ in $M(L)$. The general nonlinear model λ is not scale invariant. Again, the departure from scale invariance is not significant at two dimensional distances $|z_1 - z_2|$ close to Λ^{-1}. The argument identifying the gluing matrix can be repeated, since it depends only on the properties of the worldsurface at two dimensional distance Λ^{-1}, where the worldsurface appears scale invariant to a first approximation. The metric coupling at two dimensional distance Λ^{-1}, the inverse of the handle gluing matrix, is again identified with T^{-1} times the coefficient of the identity in the operator product at $|z_1 - z_2| = \Lambda^{-1}$,

$$\phi_i^\Lambda(z_1,\bar z_1)\phi_j^\Lambda(z_2,\bar z_2) = g_{ij}(\lambda_r)\, 1 + \cdots \,. \tag{4.57}$$

The metric coupling is also again given by the un-normalized two point expectation value at $|z_1 - z_2| = \Lambda^{-1}$

$$T^{-1}g_{ij}(\lambda_r) = Z\left\langle \phi_i^\Lambda(z_1,\bar z_1)\phi_j^\Lambda(z_2,\bar z_2)\right\rangle \tag{4.58}$$

but this formula obscures the crucial point that the metric coupling $T^{-1}g_{ij}(\lambda_r)$ is a purely short distance property of the worldsurface, because its inverse, the handle gluing matrix, is a purely short distance property of the worldsurface.

Except for the factor T^{-1}, the metric coupling $T^{-1}g_{ij}(\lambda_r)$, defined as the inverse of the handle gluing matrix, is identified with the intrinsic metric on the space of two dimensional quantum field theories used to prove the gradient property of $\beta^i(\lambda)$ [11, 12].

T is taken to be a fixed number. It will have to be fixed at an extremely small numerical value, if the volume V of macroscopic spacetime is to turn out proportional to T^{-1}. I leave untouched the question of whether the value of T is fixed by a dynamical mechanism within the lambda model.

5. $d = 2 + \epsilon$ dimensions

The first step in calculating the quantum corrections to the scaling behavior of the lambda model is to calculate the scale variation of the renormalized general nonlinear model in the presence of sources $\lambda^i(z,\bar{z})$ at a short two dimensional distance Λ^{-1}. The method is to calculate to second order in the sources, in an arbitrary spacetime λ_1 in $M(L)$, then patch together the results to get the result.

First assume a spacetime in $M(\infty)$. The general nonlinear model is scale invariant. Insert sources $\lambda^i(z,\bar{z})$. If the sources were constant, there would be no dependence on Λ^{-1}. Renormalization eliminates all dependence on the short two dimensional distance. The variation with respect to Λ^{-1} depends only on the derivatives of the sources $\lambda^i(z,\bar{z})$.

The calculation is done to second order in the sources, keeping only terms containing derivatives of the sources, giving

$$\Lambda\frac{\partial}{\partial\Lambda}_{\lambda}\, \mathrm{e}^{-\int \mathrm{d}^2z\,\mu^2\,\frac{1}{2\pi}\,\lambda^i\phi_i} = \int \mathrm{d}^2z\,\frac{1}{2\pi}\left(-\frac{1}{2}T\right)T^{-1}g_{ij}(\Lambda)\,\partial\lambda^i\,\bar{\partial}\lambda^j\,. \tag{5.1}$$

The computation is

$$\begin{aligned}\Lambda\frac{\partial}{\partial\Lambda}_{/\lambda}\,\frac{1}{2}\int \mathrm{d}^2z_1\,\mu^2\,\frac{1}{2\pi}\int \mathrm{d}^2z_2\,\mu^2\,\frac{1}{2\pi}\,\theta(|z_1-z_2|-\Lambda^{-1})\;\lambda^i(1)\lambda^j(2)\phi_i(1)\phi_j(2) =\\ = \frac{1}{2}\int \mathrm{d}^2z_1\,\mu^2\,\frac{1}{2\pi}\int \mathrm{d}^2z_2\,\mu^2\,\frac{1}{2\pi}\,\Lambda^{-1}\delta(|z_1-z_2|-\Lambda^{-1})\,\times\\ \times\,\lambda^i(1)\lambda^j(2)g_{ij}(\mu\,|z_1-z_2|)^{-2-\gamma(i)-2-\gamma(j)}\\ = \int \mathrm{d}^2z\,\frac{1}{2\pi}\left(-\frac{1}{2}\right)(\Lambda\mu^{-1})^{\gamma(i)+\gamma(j)}g_{ij}\,\partial\lambda^i\,\bar{\partial}\lambda^j\\ = \int \mathrm{d}^2z\,\frac{1}{2\pi}\left(-\frac{1}{2}T\right)T^{-1}g_{ij}(\Lambda)\,\partial\lambda^i\,\bar{\partial}\lambda^j\,.\end{aligned} \tag{5.2}$$

The two scaling fields $\phi_i\phi_j$ are replaced by their expectation value because all other operators that contribute to the product are down by powers of Λ^{-1}. The sources are assumed to vary only locally in two dimensions, which allows integration by parts.

Rewritten in terms of the running coupling constants λ_r^i, and the sources $\lambda_r^i(z,\bar{z})$, the scale variation is

$$D = \Lambda\frac{\partial}{\partial\Lambda}_{/\lambda} = \Lambda\frac{\partial}{\partial\Lambda}_{/\lambda_r} + \beta^i(\lambda_r)\frac{\partial}{\partial\lambda_r^i}\,. \tag{5.3}$$

$$D\,\mathrm{e}^{-\int \mathrm{d}^2z\,\Lambda^2\,\frac{1}{2\pi}\,\lambda_r^i\phi_i^\Lambda} = \int \mathrm{d}^2z\,\frac{1}{2\pi}\left(-\frac{1}{2}T\right)T^{-1}g_{ij}\,\partial\lambda_r^i\,\bar{\partial}\lambda_r^j\,. \tag{5.4}$$

Next, the calculation is repeated for a general nonlinear model λ_1 in $M(L)$, using the approximate scale invariance at two dimensional distances $|z_1 - z_2|$ near Λ^{-1}. Then the quadratic calculations are patched together coherently over $M(L)$ to get the full scale variation formula

$$
\begin{aligned}
D\, \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2\, \frac{1}{2\pi}\, \lambda_r^i(z,\bar{z}) \phi_i^\Lambda(z,\bar{z})} = \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2\, \frac{1}{2\pi} \lambda_r^i(z,\bar{z}) \phi_i^\Lambda(z,\bar{z})} \times \\
\times \int \mathrm{d}^2 z\, \frac{1}{2\pi} \left(-\frac{1}{2} T\right) T^{-1} g_{ij}(\lambda_r)\, \partial \lambda_r^i\, \bar{\partial} \lambda_r^j\,.
\end{aligned}
\tag{5.5}
$$

The result is a correction to the metric coupling $T^{-1} g_{ij}(\lambda_r)$ of the lambda model, which is proportional to the metric coupling itself, with a coefficient $T/2$. That is, the entire effect of the general nonlinear model on the scaling behavior of the lambda model is to give the metric coupling a scaling dimension of $T/2$.

If the lambda model were continued from two dimensions to dimension $d = 2 + T/2$, the same effect would be obtained. The metric coupling of a nonlinear model in $d = 2 + \epsilon$ dimensions has scaling dimension ϵ.

The general nonlinear model can now be dispensed with. The lambda model interacting with the general nonlinear model is equivalent to the lambda model by itself in dimension $2 + T/2$. This technical device gives a way to avoid the technical difficulty of calculating properties of the general nonlinear model in the presence of sources when the lambda field $\lambda(z,\bar{z})$ makes large excursions in the manifold of spacetimes.

6. A formula for $S(\lambda)$

6.1 The global scale variation formula

The scale variation of the general nonlinear model in the presence of sources, equation (5.5), gives a formula for the action functional $S(\lambda_r)$ of the lambda model,

$$
-\, 2 T^{-1}\, D \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2\, \frac{1}{2\pi}\, \lambda_r^i \phi_i^\Lambda} = \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2\, \frac{1}{2\pi}\, \lambda_r^i \phi_i^\Lambda}\, S(\lambda_r)\,.
\tag{6.1}
$$

Calculating $S(\lambda_r)$ from the scale variation of the general nonlinear model is equivalent to calculating the short distance effects of the local handle, then designing the lambda propagator to cancel those effects, then finding the action $S(\lambda_r)$ that produces the needed lambda propagator. The equivalence between the two procedures for finding $S(\lambda_r)$ rests on the identification of the handle gluing matrix $T\, g^{ij}$ with the inverse of the un-normalized 2-point expectation value of the fields ϕ_i.

This is only a formula for $S(\lambda_r)$. It does not explain the lambda model as a mechanism. The scale variation formula only happens to give the same result as the handle calculation. But such a simple expression as equation (6.1) for the action of the lambda model invites speculation that the formula has a deeper explanation. Perhaps there is a way of writing the integral over string worldsurface parameters which makes it obvious that the effects of local handles are cancelled by the lambda model as defined by the scale variation formula. Perhaps there is a more complete model of the string worldsurface in which the lambda model does not have to be inserted by hand, but arises automatically.

Such speculations are not immediately useful. For now, it is enough to design the lambda model in order to cancel the local handles, and use the scale variation formula as an effective way to calculate the action functional $S(\lambda_r)$ from the short distance properties of the general nonlinear model.

6.2 The local scale variation and the gradient property

Let the characteristic two dimensional distance Λ^{-1} vary in two dimensions, defining a riemannnian metric

$$\mathrm{d}s^2 = \Lambda(z,\bar{z})^2 \; |dz|^2 \tag{6.2}$$

with scalar curvature density

$$\Lambda^2 \, R_2(\Lambda)(z,\bar{z}) = -4\partial\bar{\partial}\, \ln(\Lambda^2)\,. \tag{6.3}$$

The general nonlinear model is renormalized locally in each two dimensional neighborhood. The renormalization depends covariantly on the two dimensional riemannian metric. The characteristic short two dimensional distance is $\Lambda(z,\bar{z})^{-1}$ at the point z.

The local scale derivative is

$$D(z,\bar{z}) = \Lambda(z,\bar{z}) \frac{\partial}{\partial\Lambda(z,\bar{z})}_{/\lambda_r} + \beta^i(\lambda_r(z,\bar{z}))\frac{\partial}{\partial\lambda^i_r(z,\bar{z})} \tag{6.4}$$

$$D = \int \mathrm{d}^2 z \, D(z,\bar{z})\,. \tag{6.5}$$

The local scale variation of the general nonlinear model must be dimensionless, it must be covariant in the two dimensional metric, it must be a local functional of $\Lambda(z,\bar{z})$ and $\lambda_r(z,\bar{z})$, and it must vanish if both $\Lambda(z,\bar{z})$ and $\lambda_r(z,\bar{z})$ are locally constant. It must take the form

$$\begin{aligned} D(z,\bar{z})\; \mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi}\lambda^i_r \phi^\Lambda_i} &= \mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi}\lambda^i_r \phi^\Lambda_i} \times \\ &\times \frac{1}{2\pi}\left(-\frac{1}{2}T\right)\left[T^{-1} g_{ij}(\lambda_r)\partial\lambda^i_r\bar{\partial}\lambda^j_r - \frac{1}{4}\Lambda^2 R_2(\Lambda) T^{-1} a(\lambda_r)\right] \end{aligned} \tag{6.6}$$

where $T^{-1}a(\lambda_r)$ is some function on $M(L)$.

The local scale derivatives commute,

$$[\, D[\epsilon_1],\, D[\epsilon_2]\,] = 0 \tag{6.7}$$

where

$$D[\epsilon] = \int \mathrm{d}^2 z \, \epsilon(z,\bar{z})\, D(z,\bar{z})\,. \tag{6.8}$$

The commutator of two local scale derivatives acting on the general nonlinear model is

$$\begin{aligned} &[D[\epsilon_1],\, D[\epsilon_2]\,]\mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi}\lambda^i_r \phi^\Lambda_i} = \\ &\quad = \mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi}\lambda^i_r \phi^\Lambda_i} \frac{1}{2} T \int \mathrm{d}^2 z \frac{1}{2\pi}\Big\{ (\epsilon_1\partial\epsilon_2 - \epsilon_2\partial\epsilon_1)[\beta^i T^{-1} g_{ij}(\lambda_r)\bar{\partial}\lambda^j_r - \bar{\partial} T^{-1} a(\lambda_r)] + \\ &\qquad + \left[\partial\lambda^i_r\, T^{-1} g_{ij}\, \beta^j(\lambda_r) - \partial\, T^{-1} a(\lambda_r)\right] (\epsilon_1\bar{\partial}\epsilon_2 - \epsilon_2\bar{\partial}\epsilon_1)\Big\} \end{aligned} \tag{6.9}$$

This must vanish, because the local scale derivatives commute, so

$$0 = \beta^i\, T^{-1} g_{ij}(\lambda_r)\, \bar\partial\lambda_r^j - \bar\partial\, T^{-1} a(\lambda_r) \tag{6.10}$$

$$0 = \partial\lambda_r^i\, T^{-1} g_{ij}\, \beta^j(\lambda_r) - \partial\, T^{-1} a(\lambda_r)\,. \tag{6.11}$$

The function $T^{-1}a(\lambda_r)$ depends only on λ_r, so

$$0 = \beta^i\, T^{-1} g_{ij} - \partial_j(T^{-1}a) \tag{6.12}$$

$$0 = T^{-1} g_{ij}\, \beta^j - \partial_i(T^{-1}a) \tag{6.13}$$

proving that the vector field β^i on $M(L)$ is the gradient of the function $T^{-1}a$ with respect to the metric $T^{-1}g_{ij}$. The function $T^{-1}a(\lambda_r)$ is the *potential function.*

This proof of the gradient property, using the commutativity of the local scale derivatives, is equivalent to the original proof [11, 12]. The first attempts to show the gradient property of the beta function of the general nonlinear model failed because the dilaton couplings were missing from the general nonlinear model [1, 2, 3]. Given the dilaton couplings to the two dimensional scalar curvature [9], the gradient property of $\beta^i(\lambda)$ was shown by direct calculation, in the limit of large target spacetime [10]. The beta function $\beta^i(\lambda)$ was shown to be the gradient of the classical spacetime field theory action that reproduced the large distance tree-level string scattering amplitudes. That same spacetime field theory action had been identified with the coefficient of the two dimensional scalar curvature in the local scale variation of the general nonlinear model [9]. The proof of the gradient property [11, 12] can be interpreted as an explanation of the coincidence between the calculations [9, 10] which found the same spacetime field theory action both as the coefficient of R_2 in the scale variation and as the potential function whose gradient was $\beta^i(\lambda)$. The proof [11, 12] was based on the axiomatic properties of two dimensional quantum field theory. It introduced an intrinsic metric to the space of two dimensional quantum field theories, $g_{ij}(\lambda_r)$, rather than the ad hoc metric defined on the spacetime wave modes that was used previously [1, 2, 3, 10]. The present version of the proof, using the commutativity of the local scale derivatives acting on the general nonlinear model, displays the historical genesis of the proof. The present version of the proof is particularly suited to the general nonlinear model at nonzero two dimensional distance Λ^{-1}, where the general nonlinear model is parametrized by slightly irrelevant coupling constants. The axiomatics of two dimensional quantum field theory do not quite apply.

6.3 The local scale variation formula

The local scale variation gives a formula for the covariant action density of the lambda model

$$-\,2\,T^{-1}\,D(z,\bar z)\mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi}\lambda_r^i\phi_i^\Lambda} = \mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi}\lambda_r^i\phi_i^\Lambda}\,\mathcal{L}(\lambda_r)(z,\bar z) \tag{6.14}$$

$$\mathcal{L}(\lambda_r)(z,\bar z) = \frac{1}{2\pi}\left[T^{-1} g_{ij}(\lambda_r)\,\partial\lambda_r^i\,\bar\partial\lambda_r^j - \frac{1}{4}\Lambda^2 R_2(\Lambda)\,T^{-1}a(\lambda_r)\right] \tag{6.15}$$

$$S(\lambda_r) = \int \mathrm{d}^2 z\ \mathcal{L}(\lambda_r)(z,\bar z)\,. \tag{6.16}$$

The renormalization of the general nonlinear model guarantees that the metric coupling $T^{-1}g_{ij}(\lambda_r)$ and the potential function $T^{-1}a(\lambda_r)$ depend on the two dimensional distance Λ^{-1} only through the running couplings λ_r^i.

This is the locally scale invariant form of a nonlinear model with generalized scale invariance [10]. The local scale variation of the action

$$D(z,\bar{z})\ S(\lambda_r) = \frac{1}{2\pi}\left[\beta_* T^{-1} g_{ij}(\lambda_r)\partial\lambda_r^i\bar{\partial}\lambda_r^j - \frac{1}{4}\Lambda^2 R_2(\Lambda)\,\beta^k\partial_k T^{-1}a(\lambda_r)\right]. \tag{6.17}$$

is equivalent to the local change in the couplings $T^{-1}g_{ij}$ and $T^{-1}a$ that is produced by the flow along the vector field $\beta^i(\lambda)$ in the target manifold. The potential function $T^{-1}a(\lambda_r)$ plays the role that the dilaton potential plays in the general nonlinear model [9, 10].

The local scale variation formula, equation (6.14), gives an effective method to calculate the metric coupling $T^{-1}g_{ij}(\lambda_r)$ and the potential function $T^{-1}a(\lambda_r)$, the couplings of the locally covariant lambda model. This data determines the couplings of the lambda model at every short two dimensional distance Λ^{-1}.

7. The *a priori* measure

A nonlinear model such as the lambda model is specified by two pieces of data, the metric coupling, which is a riemannian metric on the target manifold, and the *a priori* measure which is a measure on the target manifold [1, 2, 3]. In the functional integral, equation (4.39), defining the lambda model

$$\int \mathrm{D}\lambda\ \mathrm{e}^{-S(\Lambda,\lambda)}\ \mathrm{e}^{-\int \mathrm{d}^2 z\,\mu^2\,\frac{1}{2\pi}\,\lambda^i\phi_i} \tag{7.1}$$

the functional measure $\mathrm{D}\lambda$ on the lambda fields is formally a product over the points $(z,\bar{z})$ of the worldsurface, at characteristic two dimensional distance Λ^{-1},

$$\mathrm{D}\lambda = \prod_{(z,\bar{z})} \mathrm{d}\rho(\Lambda,\lambda(z,\bar{z})) \tag{7.2}$$

where the measure at each point

$$\mathrm{d}\rho(\Lambda,\lambda(z,\bar{z})) = \mathrm{dvol}(\Lambda,\lambda(z,\bar{z}))\,\rho(\Lambda,\lambda(z,\bar{z})) \tag{7.3}$$

is the *a priori* measure, written as the metric volume element $\mathrm{dvol}(\Lambda,\lambda)$ associated to the metric coupling $T^{-1}g_{ij}(\Lambda,\lambda)$ multiplied by a function $\rho(\Lambda,\lambda)$.

The *a priori* measure of the lambda model is a normalized measure on the manifold of spacetimes $M(L)$. It describes the distribution of fluctuations at two dimensional distances shorter than Λ^{-1}. It is dynamically determined, fixed by the two dimensional generalized scale invariance of the lambda model.

JHEP10(2003)063

More concretely, the *a priori* measure governs the values of the lambda field at short distance on the worldsurface. The *a priori* measure at two dimensional distance scale Λ^{-1} is the distribution of the values of the field $\lambda(z,\bar{z})$. For any function $f(\lambda)$ on the manifold of spacetimes $M(L)$,

$$\int \mathrm{d}\rho(\Lambda,\lambda)\, f(\lambda) = \langle\, f(\lambda(z,\bar{z}))\,\rangle \tag{7.4}$$

where the expectation value is in the functional integral over all lambda fluctuations at two dimensional distances up to Λ^{-1}.

The calculations of the *a priori* measure in this section will be tree-level calculations, done at leading order in T, ignoring quantum corrections. The same calculations will be used, in the same form, in section 8.4 below, to find the effective *a priori* measure of the effective lambda model, which includes the quantum corrections.

7.1 Diffusion in λ

The *a priori* measure of a scale invariant nonlinear model is completely determined by the renormalization group of the model. As the two dimensional distance Λ^{-1} increases, the *a priori* measure diffuses outward in the target manifold, because of the fluctuations. The *a priori* measure is built outward from short two dimensional distance towards longer distance, as the fluctuations at longer distances are added in. If the nonlinear model is scale invariant, then the *a priori* measure diffuses to the equilibrium measure of the diffusion process. It does not matter what arbitrary measure is used for the *a priori* measure at the cutoff two dimensional distance Λ_0^{-1}. The *a priori* measure will diffuse to the equilibrium measure at distances Λ^{-1} much longer than the cutoff. When the cutoff distance Λ_0^{-1} is taken to zero, only the equilibrium *a priori* measure is visible. In the lambda model, whose scale invariance is of the generalized kind, the action takes the same form at every two dimensional distance Λ^{-1} when expressed in the running variables λ_r^i, so the equilibrium *a priori* measure takes the same form also when expressed in terms of the running variables λ_r^i.

The diffusion process is calculated in the usual way, expanding to second order about a reference point λ_1 in the target manifold,

$$f(\lambda) = f(\lambda_1) + \lambda^i\partial_i f(\lambda_1) + \frac{1}{2}\lambda^i\lambda^j\nabla_i\partial_j f(\lambda_1) + \cdots \tag{7.5}$$

$$\begin{aligned} -\Lambda\frac{\partial}{\partial\Lambda}\,\langle\, f\,\rangle &= -\Lambda\frac{\partial}{\partial\Lambda}\,\frac{1}{2}\left\langle\, \lambda^i(z,\bar{z})\lambda^j(z,\bar{z})\,\right\rangle \nabla_i\partial_j f(\lambda_1) \\ &= T\, g^{ij}(\Lambda,\lambda_1)\nabla_i\partial_j f(\lambda_1)\,. \end{aligned} \tag{7.6}$$

Only the propagator contributes to the scale variation because only two derivatives of the lambda fields $\lambda^i(z,\bar{z})$ occur in the interaction terms of the nonlinear model. The same scale variation formula can be obtained by canonically quantizing the lambda model, in the radial quantization, where the target manifold laplacian operator occurs as the zero mode piece of the dilation generator.

Patching coherently over the target manifold gives the diffusion equation on expectation values

$$-\Lambda\frac{\partial}{\partial\Lambda}\,\langle\, f\,\rangle = \left\langle\, T\, g^{ij}(\Lambda,\lambda)\,\nabla_i\partial_j f\,\right\rangle \tag{7.7}$$

or, equivalently, the diffusion equation directly on the measure

$$- \Lambda \frac{\partial}{\partial \Lambda_{/\lambda}} \rho = T\, g^{ij}(\Lambda,\lambda)\, \nabla_i \partial_j \rho\,. \tag{7.8}$$

Changing variables to the running coupling constants λ^i_r, the *a priori* measure is rewritten

$$\mathrm{dvol}(\lambda_r)\, \rho_r(\Lambda,\lambda_r) = \mathrm{dvol}(\Lambda,\lambda)\, \rho(\Lambda,\lambda)\,. \tag{7.9}$$

The diffusion equation, rewritten in the variables λ^i_r, is

$$-\Lambda \frac{\partial}{\partial \Lambda} \rho_r = \nabla_i \left[T\, g^{ij}(\lambda_r) \partial_j + \beta^i(\lambda_r) \right] \, \rho_r\,. \tag{7.10}$$

As Λ^{-1} increases, the distribution of fluctuations diffuses outwards on the target manifold, while the running coupling constants are driven by $-\beta^i(\lambda_r)$ towards the fixed point submanifold where $\beta = 0$.

Using the gradient property, the diffusion equation is written

$$-\Lambda \frac{\partial}{\partial \Lambda} \rho_r = \nabla_i\, T\, g^{ij} \, \left(\partial_j + \partial_j (T^{-1} a)\right) \, \rho_r\,. \tag{7.11}$$

The equilibrium *a priori* measure is simply

$$\mathrm{dvol}(\lambda_r)\, \mathrm{e}^{-T^{-1} a(\lambda_r)}\,. \tag{7.12}$$

It satisfies the first order equation

$$0 = (\partial_i + T^{-1} g_{ij}\, \beta^j)\, \mathrm{e}^{-T^{-1} a(\lambda_r)} \tag{7.13}$$

which is the equation of motion $\beta = 0$.

The lambda propagator, normalized at two dimensional distance Λ^{-1}, at a scale invariant general nonlinear model, is

$$\langle\, \lambda^i(1)\ \lambda^j(2)\, \rangle = T\, g^{ij} \, \left[\frac{(\mu^2 \Lambda^{-2})^{\gamma(i)} - (\mu^2\, |z_1 - z_2|^2)^{\gamma(i)})}{\gamma(i)} \right]. \tag{7.14}$$

Evaluating at $|z_1 - z_2| = \Lambda_0^{-1}$ gives

$$\langle\, \lambda^i\, \lambda^j\, \rangle = T\, g^{ij} \, \left[\frac{(\mu^2 \Lambda^{-2})^{\gamma(i)} - (\mu^2 \Lambda_0^{-2})^{\gamma(i)})}{\gamma(i)} \right] \tag{7.15}$$

which is the solution of the diffusion equation (7.8) starting at the cutoff distance Λ_0^{-1} with initial condition the delta function measure at $\lambda^i = 0$. The integral equation for the lambda propagator, equation (4.20), is the solution of the diffusion equation.

7.2 Gaussian approximation

The potential function at a scale invariant general nonlinear model has the gaussian approximation

$$T^{-1}a(\lambda_r) = \frac{1}{2}\lambda_r^i\, T^{-1} g_{ij}\, \gamma(i)\, \lambda_r^j + \cdots \tag{7.16}$$

which gives the tree-level two point correlation function

$$\langle\, \lambda_r^i \lambda_r^j \,\rangle = T g^{ij} \frac{1}{\gamma(i)} \tag{7.17}$$

in the *a priori* measure. This matches the tree-level one point expectation value of $f(\lambda_r) = \lambda_r^i \lambda_r^j$ in the lambda model, which is the lambda propagator

$$\langle\, \lambda_r^i(1)\ \lambda_r^j(2) \,\rangle = T\, g^{ij} \left[\frac{1 - (\Lambda^2\, |z_1 - z_2|^2)^{\gamma(i)})}{\gamma(i)} \right] \tag{7.18}$$

evaluated at $|z_1 - z_2| = 0$.

The potential function written in terms of the original renormalized coupling constants λ^i is

$$T^{-1}a(\Lambda, \lambda) = \frac{1}{2}(\mu^{-1}\Lambda)^{\gamma(i)}\, \lambda^i\, T^{-1} g_{ij}\, \gamma(i)\, (\mu^{-1}\Lambda)^{\gamma(j)} \lambda^j + \cdots \tag{7.19}$$

or

$$T^{-1}a(\Lambda, \lambda) = \frac{1}{2} \mathrm{e}^{L^2\gamma(i)} \lambda^i\, T^{-1} g_{ij}\, \gamma(i)\, \mathrm{e}^{L^2\gamma(j)} \lambda^j + \cdots \tag{7.20}$$

which exhibits the spacetime ultraviolet cutoff at $\gamma(i) = L^{-2}$, at least in a naive way. The actual implementation of the ultraviolet cutoff by the renormalization of the general nonlinear model depends on the decoupling of the L-irrelevant coupling constants in the interactions. The *a priori* measure does manifestly accomplish the basic task of keeping the irrelevant running coupling constants λ_r^i from becoming large, which keeps the irrelevant renormalized coupling constants λ^i close to zero, which is a necessary condition for the renormalization to be effective.

7.3 Spacetime quantum field theory

Near a macroscopic spacetime, the manifold of spacetimes $M(L)$ is parametrized by the spacetime wave modes λ_r^i at spacetime distances larger than L. The *a priori* measure takes the form of a functional integral over the spacetime wave modes. The potential function can be written

$$T^{-1}a(\lambda_r) = g_s^{-2}\, V\, a(\lambda_r) \tag{7.21}$$

where g_s is the spacetime coupling constant, and $V\, a(\lambda_r)$ is an integral over spacetime of a local functional of the spacetime fields. The *a priori* measure

$$\mathrm{dvol}(\lambda_r)\, \mathrm{e}^{-g_s^{-2} V a(\lambda_r)} \tag{7.22}$$

is the functional integral of a spacetime quantum field theory whose classical action is the potential function $g_s^{-2}\, V\, a(\lambda_r)$, and whose classical equation of motion is $\beta^i(\lambda_r) = 0$. It is the gradient property which that the classical action principle and implies that the classical spacetime equation of motion is $\beta = 0$.

It was known by explicit calculation [9, 10] that the potential function $T^{-1}a(\lambda_r)$ in a macroscopic spacetime has the form of a classical action functional of a spacetime field theory. It was shown [13] that the classical equation of motion $\beta = 0$ in a macroscopic spacetime generates the tree-level string scattering amplitudes at large spacetime distance. There should be a direct argument from the local scale variation, equation (6.6), showing that the potential function $T^{-1}a(\lambda)$ is the generating functional for the tree-level string scattering amplitudes at large spacetime distance. It should be possible to show directly that the coefficient of the two dimensional curvature density $\Lambda^2 R_2$ in the local scale variation gives precisely the generating functional for the one particle irreducible tree-level string scattering amplitudes.

The *a priori* measure governs the string worldsurface at short distance, which is to say, roughly, that the *a priori* measure governs strings when they look like points. In particular, when a handle degenerates to a node, the two dimensional curvature density accumulates in a delta-function at the node

$$\frac{1}{4\pi}\Lambda^2 R_2(\Lambda) = -2\delta^2(z,\bar{z}) \tag{7.23}$$

containing the contribution -2 to the Euler number of the worldsurface. The covariant action of the lambda model, equation (6.14), then contains a discrete contribution at the node which is exactly $T^{-1}a(\lambda_r)$. The lambda model therefore inserts the *a priori* measure at the node

$$\int \mathrm{dvol}(\lambda_r)\, \mathrm{e}^{-T^{-1}a(\lambda_r)}\,. \tag{7.24}$$

The *a priori* measure thus controls the propagation of strings at large spacetime distance. There should be a similarly direct, general argument that the *a priori* measure governs the large distance string scattering.

8. The effective lambda model

8.1 $S_e(\lambda_e)$

Integrating out the fluctuations of the lambda fields at short two dimensional distances from the cutoff distance Λ_0^{-1} up to Λ^{-1} produces an effective general nonlinear model of the worldsurface at two dimensional distances longer than Λ^{-1}. The effective model of the worldsurface is constructed out of two dimensional surface elements at distance Λ^{-1}, so it depends only on Λ^{-1} and on effective running coupling constants λ_e^i. The effective general nonlinear model satisfies an effective renormalization group equation

$$D_e\, \mathrm{e}^{-\int \mathrm{d}^2z\, \Lambda^2\, \frac{1}{2\pi}\, \lambda_e^i \phi_i^{\Lambda,e}(z,\bar{z})} = 0 \tag{8.1}$$

$$D_e = \Lambda \frac{\partial}{\partial \Lambda}_{/\lambda_e} + \beta_e^i(\lambda_e)\frac{\partial}{\partial \lambda_e^i}\,. \tag{8.2}$$

The effective model of the worldsurface is still approximately scale invariant at two dimensional distances near Λ^{-1}, so the calculations and arguments can be carried over from the general nonlinear model,

$$D_e\, \mathrm{e}^{-\int \mathrm{d}^2 z \Lambda^2 \frac{1}{2\pi} \lambda_e^i(z,\bar z)\phi_i^{\Lambda,e}(z,\bar z)} = \mathrm{e}^{-\int \mathrm{d}^2 z \Lambda^2 \frac{1}{2\pi} \lambda_e^i(z,\bar z)\phi_i^{\Lambda,e}(z,\bar z)} \times \\ \times \int \mathrm{d}^2 z\, \frac{1}{2\pi} \left(-\frac{1}{2} T_e\right) T_e^{-1} g_{ij}^e(\lambda_e)\, \partial\lambda_e^i\, \bar\partial\lambda_e^j\,. \tag{8.3}$$

The effective metric $T_e^{-1} g_{ij}^e(\lambda_e)$ is the inverse of the handle gluing matrix for the effective model of the worldsurface.

The effective lambda model and the effective model of the worldsurface evolve in tandem under the two dimensional renormalization group, in the sense that the effective fluctuations of the lambda fields at each two dimensional distance Λ^{-1} must automatically cancel the effects of local handles in the effective model of the worldsurface at two dimensional distance Λ^{-1}. The tandem renormalization principle states that $T_e^{-1} g_{ij}^e(\lambda_e)$, obtained from the scale variation of the effective model of the worldsurface, is the effective metric coupling of the effective lambda model at two dimensional distance Λ^{-1},

$$S_e(\lambda_e) = \int \mathrm{d}^2 z\, \frac{1}{2\pi}\, T_e^{-1} g_{ij}^e(\lambda_e)\, \partial\lambda_e^i\, \bar\partial\lambda_e^j\,. \tag{8.4}$$

Calculating the scale variation of the effective model of the worldsurface is equivalent to calculating the effective action $S_e(\lambda_e)$ of the lambda model,

$$D_e\, \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi} \lambda_e^i(z,\bar z)\phi_i^{\Lambda,e}(z,\bar z)} = \mathrm{e}^{-\int \mathrm{d}^2 z\, \Lambda^2 \frac{1}{2\pi} \lambda_e^i(z,\bar z)\phi_i^{\Lambda,e}(z,\bar z)} (-\frac{1}{2} T_e)\, S_e(\lambda_e)\ . \tag{8.5}$$

Calculating the local scale variation of the effective model of the worldsurface is equivalent to calculating the covariant action density of the effective lambda model,

$$D_e(z,\bar z) = \Lambda(z,\bar z) \frac{\partial}{\partial\Lambda(z,\bar z)}_{/\lambda_e} + \beta_e^i(\lambda_e(z,\bar z)) \frac{\partial}{\partial\lambda_e^i(z,\bar z)} \tag{8.6}$$

$$D_e(z,\bar z)\, \mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi} \lambda_e^i \phi_i^{\Lambda,e}} = \mathrm{e}^{-\int \Lambda^2 \frac{1}{2\pi} \lambda_e^i \phi_i^{\Lambda,e}} \left(-\frac{1}{2}\, T_e\right) \mathcal{L}_e(\lambda_e)(z,\bar z) \tag{8.7}$$

$$\mathcal{L}_e(\lambda_e)(z,\bar z) = \frac{1}{2\pi} \left[T_e^{-1} g_{ij}^e(\lambda_e)\, \partial\lambda_e^i\, \bar\partial\lambda_e^j - \frac{1}{4} \Lambda^2 R_2(\Lambda)\, T_e^{-1} a(\lambda_e) \right] \tag{8.8}$$

$$S_e(\lambda_e) = \int \mathrm{d}^2 z\ \mathcal{L}_e(\lambda_e)(z,\bar z)\,. \tag{8.9}$$

The commutativity of local scale derivatives again implies the gradient property

$$\begin{aligned} 0 &= \beta_e^i\, T_e^{-1} g_{ij}^e - \partial_j (T_e^{-1} a_e) \\ 0 &= T_e^{-1} g_{ij}^e\, \beta_e^j - \partial_i (T_e^{-1} a_e)\,. \end{aligned} \tag{8.10}$$

The results of integrating out the lambda fluctuations at two dimensional distances up to Λ^{-1} are described by the effective classical action $S_e(\lambda_e)$ of the effective lambda model. As before, at an effective spacetime solving $\beta_e(\lambda_e) = 0$, there is a coordinate system of effective coupling constants λ_e^i in which the matrix of effective anomalous dimensions is diagonalized,

$$T_e^{-1} a_e(\lambda_e) = T_e^{-1} a_e(0) + \frac{1}{2} \lambda_e^i\, T_e^{-1} g_{ij}^e\, \gamma_e(i)\, \lambda_e^j + \cdots \tag{8.11}$$

$$\beta_e^i(\lambda_e) = \gamma_e(i)\, \lambda_e^i + \cdots\,. \tag{8.12}$$

The effective lambda propagator is, as before,

$$\langle\, \lambda_e^i(z_1,\bar z_1)\; \lambda_e^j(z_2,\bar z_2)\,\rangle = T_e\, g_e^{ij}\; \left[\frac{1-(\Lambda^2\,|z_1-z_2|^2)^{\gamma_e(i)}}{\gamma_e(i)}\right]. \tag{8.13}$$

8.2 Self-sufficiency of the lambda model

For the classical lambda model, the scale variation of the general nonlinear model, equation (6.14), gives the couplings, the metric coupling $T^{-1}g_{ij}(\lambda_r)$ and the potential function $T^{-1}a(\lambda_r)$. The scale variation formula gives the same couplings that are derived from the formulation of the lambda model as the local mechanism that cancels the short distance effects of local handles.

The effective metric coupling $T_e^{-1}g^e_{ij}(\lambda_e)$ and the effective potential function $T_e^{-1}a_e(\lambda_e)$ can be calculated by constructing the effective model of the worldsurface, then taking the scale variation, equation (8.7). But the calculation can be done in the opposite direction. The effective lambda model can be constructed, as an autonomous nonlinear model in $2+\epsilon$ dimensions. Then the effective scale variation formula, equation (8.7), can be used to find the short distance properties of the effective model of the worldsurface. In principle, calculations in the effective model of the worldsurface can be entirely avoided, unless there is a reason to be interested in string scattering at small spacetime distance.

The manifold of renormalized general nonlinear models define the lambda model. Once defined, the lambda model becomes self-sufficient. All calculations of large distance physics can be done entirely within the lambda model. No worldsurface calculations are needed. On the other hand, it might well be useful to have a second means of calculating the effective couplings of the lambda model, from the effective model of the worldsurface.

8.3 Tautological scale invariance

The generalized scale invariance of the effective lambda model derives from the existence of the effective renormalizable model of worldsurface and from the tandem renormalization property. Enormously strong constraints are put on the renormalization of the lambda model. The lambda model is not a *general* nonlinear model. Its target manifold and its couplings are extremely special, mathematically natural objects. They have remarkable properties, which are realized in the scale invariance of the effective lambda model.

The scale invariance of the effective lambda model derives from the parametrization of the effective model of the worldsurface by effective running coupling constants λ_e^i, flowing under an effective renormalization group generated by an effective beta function $\beta_e^i(\lambda_e)$. The existence of such a parametrization of the effective general nonlinear model is a consequence of locality in the two dimensional distance scale, by the usual argument of effective field theory. The processes which accomplish change of two dimensional distance in the effective model of the worldsurface do not themselves depend on the distance, but depend only on the effective couplings constants at each distance Λ^{-1}. These processes now include integrating out the fluctuations of the lambda fields taking values in $M(L)$, using the effective lambda action given by the scale variation of the effective model of the worldsurface. All dependence on the two dimensional distance is absorbed into a flowing of the effective coupling constants λ_e^i. All local properties of the effective general nonlinear model at distance Λ^{-1}, such as the effective metric coupling $T_e^{-1} g^e_{ij}(\lambda_e)$ and the potential function $T^{-1} a_e(\lambda_e)$ depend only on the effective coupling constants, and are independent of the two dimensional distance Λ^{-1}.

The effective lambda model is *tautologically* scale invariant. Its scale invariance follows automatically from its tandem relation to the effective model of the worldsurface.

8.4 The effective *a priori* measure

The overall distribution of fluctuations in the effective lambda model at distances shorter than Λ^{-1} is described by an effective *a priori* measure on the manifold of spacetimes $M(L)$,

$$\mathrm{d}\rho_e(\Lambda, \lambda_e) = \mathrm{dvol}_e(\lambda_e)\, \rho_e(\Lambda, \lambda_e) \tag{8.14}$$

$$\int \mathrm{dvol}_e(\lambda_e)\, \rho_e(\Lambda, \lambda_e) f(\lambda_e) = \langle\, f(\lambda_e(z, \bar{z}))\, \rangle_e \tag{8.15}$$

where the expectation value is evaluated in the effective lambda model at two dimensional distance Λ^{-1}.

All the considerations that applied to the classical lambda model carry over to the effective lambda model. The effective *a priori* measure satisfies an effective diffusion equation, which takes the same form as the tree-level diffusion equation. The generalized scale invariance of the effective lambda model implies that the diffusion equation has stationary coefficients,

$$\begin{aligned} -\Lambda \frac{\partial}{\partial \Lambda}_{/\lambda_e}\, \rho_e(\Lambda, \lambda_e) &= \nabla^e_i \left(T_e g_e^{ij}\, \partial_j + \beta_e^i \right)\, \rho_e \\ &= \nabla^e_i\, T_e g_e^{ij}\, \left(\partial_j + \partial_j (T_e^{-1} a_e) \right)\, \rho_e\,. \end{aligned} \tag{8.16}$$

The effective *a priori* measure is the equilibrium measure

$$\mathrm{dvol}_e(\lambda_e)\; \mathrm{e}^{-T_e^{-1} a_e(\lambda_e)} \tag{8.17}$$

which satisfies the equation of motion $\beta_e = 0$,

$$0 = \left(\partial_i + T_e^{-1} g^e_{ij}\, \beta_e^j \right)\; \mathrm{e}^{-T_e^{-1} a_e(\lambda_e)}\,. \tag{8.18}$$

The effective *a priori* measure is a measure on the manifold of spacetimes $M(L)$. If it concentrates near a macroscopic spacetime, then it will take the form

$$\mathrm{dvol}_e(\lambda_e)\; \mathrm{e}^{-g_s^{-2}\, V\, a_e(\lambda_e)} \tag{8.19}$$

which will be an effective quantum field theory of the spacetime physics at distances larger than L. The gradient property of the effective beta function $\beta_e^i(\lambda_e)$ implies the quantum action principle in the spacetime quantum field theory, and the quantum equation of motion $\beta_e = 0$.

The effective potential function is the effective classical action of the spacetime quantum field theory. Correlation functions in the *a priori* measure are classical calculations in the effective *a priori* measure. For example, at an effective spacetime solving $\beta_e = 0$, the two-point correlation function in the effective *a priori* measure is the effective lambda propagator, equation (8.13), at $z_1 = z_2$,

$$\langle\, \lambda_e^i \, \lambda_e^j \,\rangle_e = T \, g_e^{ij} \, \frac{1}{\gamma_e(i)} \, . \tag{8.20}$$

8.5 Complementarity with effective string theory

The effective *a priori* measure concentrates at the effective spacetimes λ_e in $M(L)$ where $\beta_e(\lambda_e) = 0$. The effective model of the worldsurface is scale invariant. If the spacetime is macroscopic, then the effective model of the worldsurface can be used to calculate effective string scattering amplitudes at distances larger than L. The same relation will exist between the effective string scattering amplitudes and the effective spacetime action as in the tree-level theory, by the same arguments.

The effective *a priori* measure of the lambda model is a spacetime quantum field theory. It describes the spacetime physics at distances larger than L by the effective field equation $\beta_e = 0$, just as the uncorrected *a priori* measure is a classical spacetime field theory describing the spacetime physics at distances larger than L by the classical field equation $\beta = 0$. The tandem renormalization principle guarantees that the effective spacetime action is the effective potential function derived from the effective model of the worldsurface, which is also the generating functional for the effective string scattering amplitudes at spacetime distances on the order of L. The effective string scattering amplitudes therefore agree with the scattering amplitudes calculated from the effective spacetime quantum field theory.

Again, in principle there is no need to calculate the effective string scattering amplitudes except as a description of hypothetical physics at small distance in spacetime. The effective *a priori* measure can be calculated entirely within the lambda model, and gives all the large distance physics. In particular, it gives all the large distance string scattering amplitudes, via the effective spacetime quantum field theory.

Because the lambda model is designed to cancel the effects of local handles, string theory calculations proceeding from L to larger spacetime distances would reverse the evolution of the *a priori* measure down from larger spacetime distances to L, if nonperturbative string theory calculations could be done. In the absence of a nonperturbative formulation of string theory, all that can be said is that the perturbative evolution of the *a priori* measure is consistent with the string loop expansion, calculated using the effective model of the worldsurface. This will ensure that the effective action of the spacetime quantum field theories produced by the lambda model depends on the characteristic spacetime distance L in a way that is consistent with perturbative spacetime quantum field theory, whenever the perturbative theory is accurate.

At issue is the validity of the strange method by which the lambda model is to build spacetime quantum field theories. A spacetime quantum field theory, as a measure on the wave modes of the spacetime fields, is to be built up starting from the wave modes at the largest spacetime distances. As Λ^{-1} increases, as L becomes smaller, the spacetime wave modes at smaller and smaller spacetime distances are gradually included. This is opposite to the method used by renormalizable spacetime quantum field theory, which builds from small spacetime distance to large. The lambda model must be capable of reproducing the spectacular numerical successes that have been achieved in the real world by perturbative renormalizable spacetime quantum field theory. The lambda model cannot possibly be useful unless the method by which it builds spacetime quantum field theory is consistent with that of perturbative renormalizable spacetime quantum field theory.

9. The fermionic spacetime wave modes

The lambda model needs certain basic capabilities in the general nonlinear model of the worldsurface. Compact riemannian background spacetimes must be accomodated. There must be fermionic coupling constants λ^i on an equal footing with the bosonic λ^i, so that the manifold of spacetimes $M(L)$ will be a graded manifold. The *a priori* measure will then be a measure on fermionic as well as bosonic spacetime wave modes, which can be the functional integral of a spacetime quantum field theory containing both bosonic and fermionic fields.

As far as I know, the only model of the worldsurface that has these capabilities is the model that is constructed starting with a superconformal worldsurface in which two dimensional super-reparametrization invariance is implemented using worldsurface superconformal ghost fields [14]. The GSO projection then throws away the two dimensional spinor fields and the two dimensional supersymmetry, producing a two dimensional conformally invariant worldsurface. The resulting ordinary, scale invariant model of the worldsurface is covariant in spacetime. After the GSO projection, the spinor components of the worldsurface superconformal ghost fields are incorporated into the fermionic vertex operators, which are the two dimensional scaling fields $\phi_i(z, \bar{z})$ that represent the on-shell fermionic spacetime wave modes [15]. The GSO projection removes the tachyonic string modes, eliminating all scaling fields of scaling dimension less than 2 from flowing through degenerating handles. The drawback of the covariant worldsurface is the ambiguous characterization it gives of the two dimensional scaling fields. The scaling fields occur in a multiplicity of equivalent linear spaces, called *pictures*, requiring picture independence to be verified in global worldsurface calculations.

A technical obstacle stands in the way of using the covariant worldsurface to construct a graded manifold of general nonlinear models that can serve as the target manifold of the lambda model. The fermionic scaling fields occur in different pictures from the bosonic scaling fields. A unified description is needed of the bosonic and fermionic coupling constants λ^i, to serve as graded coordinates on the graded manifold of background spacetimes.

Only the on-shell fermionic vertex operators were needed for the covariant string perturbation theory. The on-shell fermionic vertex operators are the scaling fields that represent the on-shell fermionic string states. The lambda model needs all the fermionic scaling fields, on-shell and off-shell. All the marginal and nearly marginal fermionic scaling fields have to be constructed, since all the fermionic string states flow through degenerating handles. The fermionic scaling fields have to be constructed so that they appear in the worldsurface on an equal footing with the bosonic scaling fields, effectively in the same picture. There must be a single metric $T^{-1}g_{ij}$ on all the scaling fields, symmetric in the bosonic directions and antisymmetric in the fermionic directions. In the operator product

$$\phi_i(1)\,\phi_{\mathrm{j}}(2) = \mathrm{T}^{-1}\mathrm{g}_{\mathrm{ij}}\,|\mathrm{z}_1 - \mathrm{z}_2|^{-2-\gamma(\mathrm{i})-2-\gamma(\mathrm{j})}\,1 + \cdots \tag{9.1}$$

the metric $T^{-1}g_{ij}$ is antisymmetric if and only if ϕ_i and ϕ_j are fermionic.

The fermionic scaling fields $\phi_i(z,\bar{z})$ have to be constructed in the same picture as the bosonic scaling fields, with an antisymmetric metric arising from the operator product. Then the bosonic and fermionic scaling fields can be coupled to bosonic and fermionic lambda fields $\lambda^i(z,\bar{z})$, which can be interpreted as the even and odd components of a map $\lambda(z,\bar{z})$ from the worldsurface to the graded manifold of spacetimes.

Two notable technical consequences will follow from the construction of the antisymmetric metric on the Ramond sector scaling fields.

First, it seems that only on the heterotic worldsurfaces [8] can the lambda model make sense. A non-heterotic worldsurface would contain a Ramond-Ramond sector of bosonic scaling fields. The metric $T^{-1}g_{ij}$ on the Ramond-Ramond sector would be the tensor product of two antisymmetric metrics, which cannot be positive definite. The metric coupling of the lambda model would not then be positive definite on the bosonic part of the manifold of spacetimes. For this purely technical reason, it seems that the lambda model can only work on the heterotic worldsurface.

Second, the spacetime equation of motion for the fermionic wave modes is a second order wave equation, just as it is for the bosonic modes. The spacetime equation of motion takes the same form

$$0 = \beta^i(\lambda) = \gamma(i)\lambda^i + O(\lambda^2) \tag{9.2}$$

for all the wave modes λ^i, fermionic and bosonic. The anomalous dimension is quadratic in the spacetime wave number, $\gamma(i) = p(i)^2 + m(i)^2$, for the fermionic wave modes, as well as the bosonic ones. The unphysical states in the solutions of the second order wave equation are eliminated by a gauge symmetry, leaving the usual physical solutions of the first order Dirac equation.

The construction of the fermionic scaling fields is guided by the requirement that the linear space of scaling fields should match the space of states flowing through the handle, the need for an antisymmetric metric on the ferminonic scaling fields, and also the need to realize perturbative spacetime supersymmetry as a direct cancellation between the bosonic and fermionic lambda fields, whose simplest expression is the vanishing of the graded trace

$$\delta^i_i = T\,g^{ij}\,T^{-1}g_{ji} = 0 \tag{9.3}$$

which is the dimension of the graded manifold of spacetimes. In the space of string states, the vanishing of the graded trace $\delta^i_i = 0$ follows from perturbative spacetime supersymmetry applied to the one-loop vacuum string diagram. The same equation must hold in the corresponding space of scaling fields $\phi_i(z,\bar{z})$ or in the corresponding space of coupling constants λ^i.

The rest of this section is purely technical. The notation of [15] is used.

9.1 The antisymmetric metric

The covariant string worldsurface [15] is an ordinary bosonic worldsurface. The spinor components $\beta(z), \gamma(z)$ of the superconformal worldsurface ghost fields [14] are combined with the spacetime degrees of freedom to form the scaling fields $\phi_i(z, \bar{z})$. For simplicity, I only treat here the worldsurface in flat ten dimensional spacetime, and discuss only the z dependent parts of the worldsurface scaling fields. All the essential technical issues are resolved in this simplified context. The novel part of the construction of the fermionic scaling fields involves only the structure of the β, γ ghost fields. It is easily taken over to the general nonlinear model with a general target spacetime.

The space of z dependent scaling fields splits into two subspaces, the Ramond sector and the Neveu-Schwarz sector. The string states and the corresponding two dimensional scaling fields are described redundantly in an infinite set of pictures, labelled by the picture charge. The Neveu-Schwarz sector is represented by the pictures of integer picture charge, the Ramond sector by the pictures of charge integer plus half.

In analyzing the effects of degenerating handles in the worldsurface, there is an obvious benefit to choosing those special pictures in which the scaling dimensions are bounded below. In those special pictures, the scaling dimensions of the z dependent fields are bounded below by 1. The z dependent fields of dimension 1 are combined with $\bar{z}$ dependent fields of dimension 1 to form the marginal scaling fields, of scaling dimension 2.

For the Neveu-Schwarz sector, there is only one picture with bounded scaling dimensions, the picture of charge -1. The z dependent Neveu-Schwarz sector fields with picture charge -1 and scaling dimension 1 are, after GSO projection, the ten bosonic fields

$$\psi_\mu e^{-\phi} \quad \mu = 1, \cdots, 10 \tag{9.4}$$

plus two fermionic fields made entirely from the worldsurface ghost fields

$$\beta_{-1/2} e^{-\phi}, \qquad \gamma_{-1/2} e^{-\phi}. \tag{9.5}$$

The field $\phi(z)$ is the bosonization of the $\beta\gamma$ current, $\beta\gamma = -\partial\phi$. The exponentials of $\phi(z)$ correspond to the highest weight states of the β, γ algebra,

$$\beta_n e^{q\phi} = 0 \qquad n \geq -q - 1/2 \tag{9.6}$$

$$\gamma_n e^{q\phi} = 0 \qquad n \geq q + 3/2\,. \tag{9.7}$$

The operators β_n and γ_n lower the scaling dimension by n. The only pictures with scaling dimension bounded below are $q = -1/2$, $q = -1$, and $q = -3/2$

The graded trace δ^i_i in flat spacetime is the product of two factors,

$$\delta^i_i = (\delta^i_i)_z (\delta^i_i)_{\bar{z}}\,. \tag{9.8}$$

One factor is the graded trace over the z dependent fields, the other factor comes from the $\bar{z}$ dependent fields. The object will be to have

$$(\delta^i_i)_z = 0\,. \tag{9.9}$$

The Neveu-Schwarz sector fields make a net contribution of $10-2 = 8$. The Ramond sector fields must make a contribution of -8.

The metric on the bosonic z dependent fields of the N-S sector is symmetric and positive

$$\left\langle \psi_\mu e^{\phi}(z_1)\, \psi_\nu e^{\phi}(z_2) \right\rangle = \delta_{\mu\nu}\,(z_1 - z_2)^{-2}\,. \tag{9.10}$$

while the metric on the pair of fermionic N-S fields is antisymmetric

$$\left\langle \beta_{-1/2} e^{\phi}(z_1)\, \gamma_{-1/2} e^{\phi}(z_2) \right\rangle = (z_1 - z_2)^{-2} \tag{9.11}$$

$$\left\langle \gamma_{-1/2} e^{\phi}(z_1)\, \beta_{-1/2} e^{\phi}(z_2) \right\rangle = -(z_1 - z_2)^{-2}\,. \tag{9.12}$$

The Ramond sector has *two* pictures in which the scaling dimensions are bounded below, the pictures of charges $-1/2$ and $-3/2$. These two pictures are conjugate to each other in the metric on scaling fields, because a product of scaling fields can have nonzero expectation value on the 2-sphere, where the metric is calculated, only if the sum of the picture charges is -2.

Before GSO projection, the dimension 1 fields of picture charge $-1/2$ are

$$F_1(\gamma_0) e^{-\phi/2} S_\alpha(z) \tag{9.13}$$

and those of picture charge $-3/2$ are

$$F_2(\beta_0) e^{-3\phi/2} S_\beta(z) \tag{9.14}$$

where $S_\alpha(z)$ is the spin field of the spacetime degrees of freedom, a 32 component spacetime spinor in 10 dimensional spacetime; β_0 and γ_0 are the zero mode operators of the spinor ghost fields, satisfying the canonical commutation relations $[\gamma_0, \beta_0] = 1$; and $F_{1,2}$ are arbitrary functions.

The spacetime spinor fields $S_\alpha(z)$ have 32 components. The GSO projection cuts that number in half, to 16. Somehow, the infinite multiplicity of the β_0, γ_0 zero mode representation must give another factor of $1/2$, to obtain the contribution of -8 to $(\delta^i_i)_z$, in order to cancel the contribution of $+8$ from the N-S sector.

Several questions need to be answered. Which of the two pictures should go at each end of a degenerating handle? How can the Ramond sector fields, appearing in two different pictures, $-1/2$ and $-3/2$, play the same role as the N-S sector fields, appearing in the single charge -1 picture? How can a single *antisymmetric* metric on a space of graded dimension -8 be made from the symmetric metric $h_{\alpha\beta}$ on the spacetime spinors?

All of these questions are answered by finding a formalism in which the two conjugate $q = -1/2$ and $q = -3/2$ pictures appear effectively in a single picture of charge -1. The key is to represent the states of the quantized β, γ ghost fields in terms of distributions [16]. One crucial technical subtlety in the nature of these distributions has to be remarked.

Define the Ramond sector field

$$S_\alpha(t, z) = \delta(t - \gamma_0) e^{-\phi/2} S_\alpha(z) \tag{9.15}$$

which depends on a spacetime spinor index α and a complex number t. $S_\alpha(t, z)$ is a distribution in the complex number t, satisfying

$$\int \mathrm{d}t\, t^m\, S_\alpha(t,z) \;=\; \gamma_0^m e^{-\phi/2} S_\alpha(z) \tag{9.16}$$

$$\int \mathrm{d}t\, \delta^{(n)}(t)\, S_\alpha(t,z) \;=\; \beta_0^n e^{-3\phi/2} S_\alpha(z) \tag{9.17}$$

where the latter follows from the identity [16]

$$e^{-3\phi/2} = \delta(\gamma_0) e^{-\phi/2}\,. \tag{9.18}$$

The metric is

$$\langle S_{\alpha_1}(t_1,z_1)\, S_{\alpha_2}(t_2,z_2)\rangle = K(t_1,t_2)\, h_{\alpha_1,\alpha_2}\, (z_1-z_2)^{-2} \tag{9.19}$$

where h_{α_1,α_2} is the symmetric metric on the spacetime spinors, and

$$K(t_1,t_2) = \delta(t_1-t_2)\,. \tag{9.20}$$

The crucial technical subtlety is that this delta function distribution is an *odd* function of its argument,

$$\delta(t_1-t_2) = -\delta(t_2-t_1)\,. \tag{9.21}$$

This is not the real delta function distribution which is a measure on the real line, to be integrated against functions of a real variable. Rather, it is the formal delta function of a complex variable. It is to be integrated against analytic functions of the complex variable according to the rule

$$\int \mathrm{d}t\, \delta(t) f(t) = f(0)\,. \tag{9.22}$$

This formal delta function can be written as an equivalence class of ordinary distributional 1-forms on the complex plane,

$$\delta(t) = \mathrm{d}\bar{t}\, \delta^2(t,\bar{t}) \tag{9.23}$$

modulo $\partial/\partial\bar{t}$ of an arbitrary distribution with compact support on the complex plane. The formal delta function $\delta(t)$ is a 1-form, therefore an odd object. It satisfies, for any nonzero complex number a,

$$\delta(at) = \mathrm{d}(\bar{a}\bar{t})\, \delta^2(at,\bar{a}\bar{t}) = a^{-1}\delta(t)\,. \tag{9.24}$$

In particular

$$\delta(-t) = -\delta(t)\,. \tag{9.25}$$

To see that the formal delta function is needed for the distributional quantization of the $\beta(z)$, $\gamma(z)$ ghost fields, consider the identity

$$e^{-\phi}(z) = \delta(\gamma(z)) \tag{9.26}$$

which is justified by comparing the analytic operator product expansions

$$\gamma(z)\, e^{-\phi}(0) = z\, \gamma_{-1/2} e^{-\phi}(0) + \cdots \tag{9.27}$$

$$\gamma(z)\, \delta(\gamma(0)) = z\, \partial\gamma(0)\delta(\gamma(0)) + \cdots\,. \tag{9.28}$$

Then consider the operator product

$$e^{-\phi}(z)\, e^{-\phi}(0) = z^{-1}\, e^{-2\phi}(0) + \cdots \tag{9.29}$$

which is translated, for all complex numbers z,

$$\begin{aligned} \delta(\gamma(z))\, \delta(\gamma(0)) &= \delta(z\partial\gamma(0))\, \delta(\gamma(0)) + \cdots \\ &= z^{-1}\, \delta(\partial\gamma(0))\, \delta(\gamma(0)) + \cdots \end{aligned} \tag{9.30}$$

only if the formal delta function is used. This calculation illustrates how the usual quantization of the β, γ ghost fields [15] is systematically translated into the language of formal delta functions [16]. More details of the translation are given in section 9.5 below.

The metric K is antisymmetric,

$$K(t_1, t_2) = -K(t_2, t_1) \tag{9.31}$$

and it is an odd object, because it is a formal delta function.

The metric $h_{\alpha,\beta}$ on the spacetime spinors can also be interpreted as a distribution. The spacetime spinors s^α are the functions of 5 anticommuting variables $\hat{t}$. There are $2^5 = 32$ linearly independent functions $s^\alpha(\hat{t})$. The symmetric metric $h_{\alpha\beta}$ is represented by the distribution

$$\begin{aligned} \hat{K}(\hat{t}_1, \hat{t}_2) &= s^{\alpha_1}(\hat{t}_1)\, h_{\alpha_1\alpha_2} s^{\alpha_2}(\hat{t}_2) \\ &= \delta^5(\hat{t}_1 + \hat{t}_2)\,. \end{aligned} \tag{9.32}$$

The metric on the spacetime spinors $\hat{K}(\hat{t}_1, \hat{t}_2)$ is also an odd object, but symmetric.

The spinor fields $S_\alpha(z)$ are rewritten as functions of the 5 anticommuting variables $\hat{t}$,

$$S(\hat{t}, z) = S_\alpha(z)\, s^\alpha(\hat{t})\,. \tag{9.33}$$

The Ramond sector scaling field is a function of t and $\hat{t}$,

$$S(t, \hat{t}, z) = \delta(t - \gamma_0)\, e^{-\phi/2}\, S(\hat{t}, z)\,. \tag{9.34}$$

The Ramond sector metric, equation (9.19), is the product

$$K(t_1, t_2)\, \hat{K}(\hat{t}_1, \hat{t}_2) = \delta(t_1 - t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2)\,. \tag{9.35}$$

It is even, as the product of two odd objects. It is antisymmetric as the product of an antisymmetric metric and a symmetric metric. The Ramond sector field $S(t, \hat{t}, z)$ is therefore fermionic.

The GSO transformation sends $\hat{t} \to -\hat{t}$ and $t \to -t$, so the GSO projected fields are

$$S_+(t, \hat{t}, z) = \frac{1}{2} S(t, \hat{t}, z) + \frac{1}{2} S(-t, -\hat{t}, z)\,. \tag{9.36}$$

The metric on the GSO projected fields is

$$\frac{1}{2}\delta(t_1 - t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2) + \frac{1}{2}\delta(t_1 + t_2)\, \delta^5(\hat{t}_1 - \hat{t}_2)\,. \tag{9.37}$$

When the scaling field $S_+(t,\hat{t},z)$ is smeared with a polynomial function of t, it is in picture $-1/2$. When it is smeared with $\delta(t)$, or derivatives of $\delta(t)$, it is in picture $-3/2$. Effectively, in any worldsurface calculation, $S_+(t,\hat{t},z)$ is midway between the two pictures, which puts it in picture -1 along with the fields of the N-S sector.

The fermionic marginal scaling fields $\phi_i(z,\bar{z})$ are formed by combining the Ramond sector fields $S_+(t,\hat{t},z)$ with bosonic scaling fields depending on $\bar{z}$, for example

$$S_+(t,\hat{t},z)\,\bar{\partial}x^\mu(\bar{z})\,. \tag{9.38}$$

Fermionic coupling constants λ^i couple to these fermionic scaling fields.

The merging of the two pictures $-1/2$ and $-3/2$ to form a *virtual* picture -1 removes the ambiguity in the assignment of a picture at each of the two ends of a local handle. The handle can be represented as a sum of pairs of bosonic fields plus a sum of pairs of fermionic fields, each contracted with the handle gluing metric. Whatever picture changing operators are needed near the local handle will serve simultaneously to define the insertions of both the fermionic and the bosonic scaling fields.

The antisymmetric metric is presented as a kernel in equation (9.37). It is a well-defined generalized function of the variables $t_1,\hat{t}_1,t_2,\hat{t}_2$. But no concrete vector space is defined, on which the antisymmetric metric acts as a bilinear inner product. Formally, the Ramond sector fields $S(t,\hat{t},z)$ lie midway between picture $-1/2$ and picture $-3/2$. This formal description serves all practical purposes, since calculations in the lambda model require only contractions of products of the metric and its inverse. But the technical foundations of the theory would be more secure if the Ramond sector fields could be indexed by a concrete vector space. This should be a vector space of functions of t, lying midway between the analytic functions and the formal delta functions, perhaps some space of half-forms.

9.2 Lack of positivity in a Ramond-Ramond sector

If the string worldsurface has a Ramond-Ramond sector, as in any of the non-heterotic string theories, there is a serious technical difficulty for the lambda model, because the metric on a Ramond-Ramond scaling fields is not positive definite. For example, the Ramond-Ramond scaling fields

$$S_+(t_1,\hat{t}_1,z)\,\bar{S_+}(t_2,\hat{t}_2,\bar{z}) \tag{9.39}$$

have a metric that is the tensor product of two antisymmetric metrics. Such a tensor product always has directions with negative metric. Bosonic lambda fields would couple as sources to these Ramond-Ramond scaling fields. The metric coupling on those bosonic lambda fields would be negative. The action of the lambda model on these bosonic lambda fields would be unbounded below. I do not see how the lambda model could be made to work then. I do not see how there could be control over the short distance worldsurface fluctuations of the negative metric bosonic lambda fields, even if they are unphysical gauge artifacts.

This pathology seems to disqualify the non-heterotic string theories from being associated with a sensible large distance physics, at least as provided by the lambda model. But the pathology is purely technical. It should have a physical interpretation. There should be an explanation in physical terms of what goes wrong with the large distance physics in non-heterotic string theories.

9.3 $\delta^i_i = 0$

The goal now is to show that the Ramond sector fields contribute -8 to the graded trace. The inverse of the kernel

$$\delta(t_1 - t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2) \tag{9.40}$$

is the kernel

$$\mathrm{d}^5\hat{t}_1 \mathrm{d}t_1\, \delta(t_1 - t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2)\, \mathrm{d}^5\hat{t}_2 \mathrm{d}t_2 \tag{9.41}$$

because

$$\begin{aligned} \int_{t_2,\hat{t}_2} \delta(t_1 - t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2)\, \mathrm{d}^5\hat{t}_2 \mathrm{d}t_2\, \delta(t_2 - t_3)\, \delta^5(\hat{t}_2 + \hat{t}_3)\, \mathrm{d}^5\hat{t}_3 \mathrm{d}t_3 = \\ = \delta(t_1 - t_3)\, \delta^5(\hat{t}_1 - \hat{t}_3) \mathrm{d}^5\hat{t}_3 \mathrm{d}t_3 \end{aligned} \tag{9.42}$$

which is the kernel of the identity operator. The inverse metric on the Ramond sector fields is then the GSO projection

$$\mathrm{d}^5\hat{t}_1 \mathrm{d}t_1 \left[\frac{1}{2}\delta(t_1 - t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2) + \frac{1}{2}\delta(t_1 + t_2)\, \delta^5(\hat{t}_1 - \hat{t}_2)\right] \mathrm{d}^5\hat{t}_2 \mathrm{d}t_2\,. \tag{9.43}$$

The contribution of the Ramond sector fields to the graded trace is

$$-\int \mathrm{d}^5\hat{t}_1 \mathrm{d}\hat{t}_1 \left[\frac{1}{2}\delta(t_1 - t_2)\, \delta^5(\hat{t}_1 - \hat{t}_2) + \frac{1}{2}\delta(t_1 + t_2)\, \delta^5(\hat{t}_1 + \hat{t}_2)\right]_{/\hat{t}_2 = \hat{t}_1, \hat{t}_2 = \hat{t}_1} \tag{9.44}$$

where the overall minus sign comes from the antisymmetry of the metric. The first term inside the integral needs to be regularized

$$\begin{aligned} \left[\delta(t_1 - t_2)\, \delta^5(\hat{t}_1 - \hat{t}_2)\right]_{/\hat{t}_2 = \hat{t}_1, \hat{t}_2 = \hat{t}_1} &= \lim_{y \to 1} \left[\delta(t_1 - y t_1)\, \delta^5(\hat{t}_1 - y \hat{t}_2)\right] \\ &= \lim_{y \to 1} \left[(1 - y)^{-1} \delta(t_1)\, (1 - y)^5 \delta^5(\hat{t}_1)\right] \\ &= 0\,. \end{aligned} \tag{9.45}$$

The second term contributes

$$\begin{aligned} -\int \mathrm{d}^5\hat{t}_1 \mathrm{d}\hat{t}_1\, \frac{1}{2}\delta(2t_1)\, \delta^5(2\hat{t}_1) &= -\int \mathrm{d}^5\hat{t}_1 \mathrm{d}\hat{t}_1\, \frac{1}{2}\, 2^{-1}\delta(t_1)\, 2^5 \delta^5(\hat{t}_1) \\ &= -8 \end{aligned} \tag{9.46}$$

to the graded trace, as was to be shown. The first factor $1/2$ is from the GSO projection. The factor 2^5 is from the trace over spacetime spinors. The extra factor of $1/2$ comes from the trace over the states of the bosonic ghost zero modes β_0, γ_0.

9.4 Second order wave equation

In flat spacetime, the almost marginal fermionic scaling fields take the form

$$\phi_i(z,\bar{z}) = S_+(t,\hat{t},z)\,\bar{\partial}x^\mu(\bar{z})\,\mathrm{e}^{ipx} \tag{9.47}$$

indexed by $i = (t,\hat{t},\mu,p_\mu)$. The anomalous scaling dimension is $\gamma(i) = p^2$. The spacetime equation of motion $\beta^i(\lambda) = 0$ linearizes to $\gamma(i)\lambda^i = 0$ which is the second order differential equation in spacetime $p^2\lambda^i = 0$. In a curved spacetime, the linearized equation of motion on the fermionic lambda modes becomes a covariant second order differential operator.

On the on-shell states, which satisfy $p^2 = 0$, the worldsurface BRS operator is $t\,\not{p}$, where $\not{p}$ is the spacetime Dirac operator. The physical states, in either of the two conjugate Ramond sector pictures, are the BRS cohomology classes. In either picture, the BRS cohomology classes are the solutions of the first order spacetime Dirac equation. The infinite multiplicity of the ghost zero modes is eliminated.

The *a priori* measure of the lambda model, interpreted as a spacetime quantum field theory, uses a second order differential wave equation on the fermionic fields, not the traditional Dirac equation. But the physical content is the same.

9.5 Quantizing the β, γ ghost fields using the formal delta function

The β,γ ghost fields are expanded in modes

$$\beta(z) = \sum_n z^{-n-3/2}\beta_n \qquad \gamma(z) = \sum_n z^{-n+1/2}\gamma_n \tag{9.48}$$

where the index n is integer in the Ramond sector, integer plus half in the NS sector. The modes satisfy canonical commutation relations

$$[\gamma_m,\beta_n] = \delta_{m+n}\,. \tag{9.49}$$

The ground state of picture charge q is the state $|q\rangle$, satisfying

$$\gamma_n\,|q\rangle = 0 \qquad n - q = \frac{3}{2},\frac{5}{2},\cdots \tag{9.50}$$

$$\beta_n\,|q\rangle = 0 \qquad n + q = -\frac{1}{2},\frac{1}{2},\cdots\,. \tag{9.51}$$

The ground state $|0\rangle$ is the SL_2 invariant state.

In the bosonization formalism for the β,γ ghosts,

$$|q\rangle = e^{q\phi}(0)\;|0\rangle \tag{9.52}$$

The states are represented as distributional wave functions [16], say of the γ_n. The β_n act as derivative operators

$$\beta_n = -\frac{\partial}{\partial\gamma_{-n}}\,. \tag{9.53}$$

The ground state of picture charge q is

$$|q\rangle = \delta(\gamma_{3/2+q})\,\delta(\gamma_{5/2+q})\,\delta(\gamma_{7/2+q})\,\cdots\,. \tag{9.54}$$

The dual states are

$$\langle q| = \cdots\ \delta(\gamma_{-7/2-q})\,\delta(\gamma_{-5/2-q})\,\delta(\gamma_{-3/2-q}) \tag{9.55}$$

so

$$\langle -q-2\,|\,q\rangle = 1 \tag{9.56}$$

The states and dual states can be thought of as two classes of analytic subvarieties in the infinite dimensional analytic manifold whose coordinates are the γ_n. The inner product is the intersection number of the subvarieties.

The field $\delta(\gamma(z))$ acts on the ground state $|q\rangle$ by

$$\delta(\gamma(z))\ |q\rangle = \delta(z^{-q}\gamma_{1/2+q}+\cdots)\ |q\rangle = z^{q}\ |q-1\rangle + \cdots \tag{9.57}$$

where the formal delta function must be used in order that the operator product expansion will be analytic.

The inner product on the zero mode wave functions is obtained by noting that

$$\left\langle -\frac{1}{2}\right| F(\gamma_0) \left| -\frac{1}{2}\right\rangle = \int \mathrm{d}\gamma_0\ F(\gamma_0) \tag{9.58}$$

then calculating

$$\begin{aligned}\left\langle -\frac{1}{2}\right|\ \delta(t_1-\gamma_0)\,\delta(t_2-\gamma_0)\ \left| -\frac{1}{2}\right\rangle &= \int \mathrm{d}\gamma_0\,\delta(t_1-\gamma_0)\,\delta(t_2-\gamma_0)\\ &= \delta(t_1-t_2)\,.\end{aligned} \tag{9.59}$$

A formal integral representation of the formal delta function,

$$\delta(\gamma) = \int \mathrm{d}t\, e^{t\gamma} \tag{9.60}$$

allows such calculations as

$$\begin{aligned}\delta(\beta(z))\ |q\rangle &= \delta(z^{q}\beta_{-q-3/2}+\cdots)\ |q\rangle\\ &= z^{-q}\,\delta(\beta_{-q-3/2})\,\delta(\gamma_{3/2+q})\ |q+1\rangle + \cdots\\ &= z^{-q}\int \mathrm{d}t\ \exp\left(-t\frac{\partial}{\partial\gamma_{3/2+q}}\right)\ \delta(\gamma_{3/2+q})\ |q+1\rangle + \cdots\\ &= z^{-q}\int \mathrm{d}t\,\delta(-t+\gamma_{3/2+q})\ |q+1\rangle + \cdots\\ &= -z^{-q}\ |q+1\rangle + \cdots\end{aligned} \tag{9.61}$$

which is the operator product needed to make the identification

$$e^{-\phi}(z) = -\delta(\beta(z))\,. \tag{9.62}$$

The identities

$$1 = \int \mathrm{d}\gamma\,\delta(\gamma) = \int \mathrm{d}\gamma \int \mathrm{d}t\, e^{t\gamma} = -\int \mathrm{d}t \int \mathrm{d}\gamma\, e^{t\gamma} = \int \mathrm{d}t\,\delta(t) \tag{9.63}$$

are justified by the fact that the formal expression $e^{t\gamma}$ is odd under exchange of t and γ, because it implicitly contains the factor $\mathrm{d}\bar{t}\mathrm{d}\bar{\gamma}$.

10. Geometric identities on the manifold of spacetimes

Scale invariance in $2+\epsilon$ dimensions in a nonlinear model such as the lambda model is expressed by the vanishing of the beta function, which is a geometric identity on the target manifold of the model [1, 2, 3]. To one loop, the geometric identity expressing ordinary scale invariance is

$$0 = -\epsilon T^{-1} g_{ij} + 2R_{ij} \, . \tag{10.1}$$

The numerical coefficient 2 multiplying the Ricci tensor is due to the normalization of the action $S(\lambda)$, which is the same as the normalization of the general nonlinear model, which is designed to give anomalous dimensions of the form $p(i)^2$, with numerical coefficient 1.

Scale invariance of the generalized kind is expressed by a somewhat more elaborate geometric identity involving the potential function $T^{-1}a(\lambda)$ and whatever other couplings occur in the nonlinear model [1, 2, 3, 10].

The effective metric coupling $T^{-1}g^e_{ij}(\lambda_e)$ and the effective potential function $T^{-1}a_e(\lambda_e)$ and whatever other effective couplings might arise will satisfy the geometric identities expressing generalized scale invariance.

I will not write these *meta* Einstein equations here. The quantum corrections to the metric coupling and the other couplings will enter at each order, so the full import is in the exact equations, not in their truncation to one loop or to any finite number of loops. The geometric identities will involve the effective potential function $T^{-1}a_e(\lambda_e)$, which is the effective action in the spacetime quantum field theory. It will eventually be interesting to ask what significance the meta Einstein equations might have in the special spacetime quantum field theories produced by the lambda model.

10.1 Geometric identites from perturbative spacetime supersymmetry

Perturbative spacetime supersymmetry suppresses string loop corrections. The lambda model is formulated to cancel the effects of the string loop corrections at large distance in spacetime. Perturbative spacetime supersymmetry of the string theory in a given spacetime λ will be mirrored as a perturbative symmetry of the lambda fluctuations around the point λ in the target manifold. Perturbative spacetime supersymmetry will supress perturbative quantum corrections to the couplings of the lambda model at the point λ in its target manifold.

In particular, perturbative spacetime supersymmetry preserves the degeneracy of the manifold of spacetimes $M(\infty)$ against perturbative quantum corrections. $M(\infty)$ is the manifold of solutions of $\beta(\lambda)=0$. The restricted lambda model is the formal, perturbative nonlinear model whose target manifold is $M(\infty)$. In the restricted lambda model, $\beta_e(\lambda_e) = 0$ perturbatively on $M(\infty)$. So the restricted lambda model will be scale invariant order by order in the loop expansion, in the ordinary sense of scale invariance. The vanishing of the perturbative beta function for the metric coupling of the restricted lambda model means that the metric $T^{-1}g_{ij}(\lambda)$ on $M(\infty)$ will satisfy a series of geometric identities, indicative of a very special geometry. The fermionic directions in the manifold of spacetimes are essential for these identities, since the identities arise from cancellations between the bosonic and fermionic directions in the manifold of spacetimes.

JHEP10(2003)063

The first cancellation is the vanishing of the (graded) trace

$$\delta^i_i = T^{-1} g_{ij}\, T\, g^{ji} = 0 \tag{10.2}$$

which states that the graded dimension of the manifold of spacetimes is zero. The equation $\delta^i_i = 0$ is easily recognized from the string loop expansion. It is the condition that the one loop correction to the vacuum string amplitude is finite. The one loop correction to the vacuum amplitude is

$$\int Z_1(q, \bar{q}) \tag{10.3}$$

where $Z_1(q, \bar{q})$ is the partition function of the genus 1 worldsurface, the complex one-torus parametrized by $q = e^{2\pi i\tau}$. The integral is over the modular domain of the upper half complex τ plane. The partition function is nonsingular everywhere except possibly at $q = 0$, where the torus degenerates. The only place where the integral might diverge is at $q = 0$. The complex one-torus near $q = 0$ is an almost degenerate handle connected to a 2-sphere. The integral is cut off at $|q|^{1/2} > \mu\Lambda_0^{-1}$. The divergent part is

$$\Lambda_0 \frac{\partial}{\partial \Lambda_0} \int \mathrm{d}^2 q\, \frac{1}{2\pi}\, |q|^{-4}\, |q|^{2+\gamma(i)}\, T\, g^{ij}\, T^{-1} g_{ij} = 2(\mu\Lambda_0^{-1})^{2\gamma(i)}\, T\, g^{ij}\, T^{-1} g_{ij}\,. \tag{10.4}$$

Finiteness follows from the existence of a conserved, holomorphic supersymmetry current $Q_S(z)$ on the worldsurface, with charge operator Q_S, and the existence of a conjugate operator Q'_S such that [15]

$$[Q_S,\, Q'_S] = 1\,. \tag{10.5}$$

Finiteness in the limit $\Lambda_0^{-1} \to 0$ is precisely the condition $\delta^i_i = 0$, where the graded trace is taken over the marginal coupling constants, those having $\gamma(i) = 0$.

A more subtle version of this argument should work locally in a spacetime λ that lies in $M(L)$, at nonzero short two dimensional distance Λ^{-1}. The argument should give a version of the vanishing of the graded trace, $\delta^i_i = 0$, that applies locally in spacetime.

10.2 The meta Einstein equation on $M(\infty)$

The second geometric identity is a meta Einstein equation on $M(\infty)$, expressing one loop scale invariance of the metric coupling of the restricted lambda model in $d = 2 + \epsilon$ dimensions, with $\epsilon = T/2$,

$$0 = -\frac{1}{2} g_{ij} + 2\, R_{ij}\,. \tag{10.6}$$

The term $2\, R_{ij}$ is the usual one loop beta function of the nonlinear model. The term $-\frac{1}{2} g_{ij}$ is the contribution from the scale variation of the general nonlinear model, equation (5.1).

It should be possible to derive the meta Einstein equation (10.6) directly from one loop finiteness of the string loop corrections. Differentiating the finite one string loop vacuum correction, equation (10.3), with respect to the marginal coupling constants λ^i, gives the finiteness of the one string loop correction to the one point function. This is the vanishing of the one loop correction to $\beta^i(\lambda) = 0$. Differentiating the finite one loop vacuum correction, equation (10.3), twice with respect to the marginal coupling constants

λ^i, gives the finiteness of the one string loop correction to the two point function. As before, the scale variation must then vanish. The scale variation extracts the contribution of a degenerating handle attached to a 2-sphere in which there are two scaling fields ϕ_i, ϕ_j. This contribution is an integral of the four point expectation value on the 2-sphere, contracted with a handle gluing matrix, of the form

$$\int g^{kl} \, \langle \, \phi_i(1), \phi_j(2), \phi_k(3), \phi_l(4) \, \rangle \; . \tag{10.7}$$

The Ricci tensor of the metric g_{ij} can be calculated from the scale variation of the general nonlinear model with sources, equation (5.1). The metric is a two point expectation value of scaling fields. The curvature tensor is made from two derivatives of the metric, so the curvature is given by an integral of an expectation value of four scaling fields. The Ricci tensor is then obtained by contracting the curvature tensor with the inverse metric, g^{ij}.

These results of these two calculations have the same form, so it is plausible that the meta Einstein equation can be derived explicitly from one loop string finiteness. Heuristically, the one loop finiteness of the string loop corrections gives rise to an identity on the metric which involves two derivatives of the metric. By covariance in $M(\infty)$, this identity should be of the form of the meta Einstein equation. Only the relative numerical coefficient 1/4 between the two terms needs to be verified.

It should also be possible to verify the meta Einstein equation (10.6) by explicit calculation of the Ricci tensor of the metric $T^{-1}g_{ij}(\lambda)$ on $M(\infty)$, at least in simple cases such as the manifold of toroidal spacetimes.

The restricted lambda model is perturbatively finite because of its generalized scale invariance, which is a basic property of the lambda model. Perturbative spacetime supersymmetry is only an accidental property of individual spacetimes. Perturbative spacetime supersymmetry simplifies the realization of generalized scale invariance in the lambda model, by maintaining the degeneracy of the manifold of spacetimes $M(\infty)$ against perturbative corrections. As a consequence, there are strong identities on the geometry of the manifold of perturbatively supersymmetric spacetimes. For physics, perturbative spacetime supersymmetry is useful because, by maintaining the degeneracy against perturbative corrections, it guarantees that any effects that lift the degeneracy will be nonperturbatively small.

11. Lambda instantons

The dominant nonperturbative effects in the lambda model will be produced by harmonic surfaces in the space of string backgrounds, the *lambda instantons*. A lambda instanton is a classical field configuration $\lambda_H(z,\bar{z})$ which is a local minimum of the lambda model action, $S(\lambda)$. These are the harmonic surfaces in the manifold of spacetimes.

There are at least two kinds of lambda instanton. The *global* lambda instantons, are the harmonic surfaces in the manifold $M(\infty)$. The action $S(\lambda_H)$ of a global lambda instanton is on the order of T^{-1}, so only collective effects of global lambda instantons will be significant. I will describe here one elementary example of a global lambda instanton, in the manifold of toroidal spacetimes, and speculate on possible collective effects that might single out a macroscopic spacetime.

At a macroscopic spacetime λ in $M(L)$, there are *localized* lambda instantons, which are harmonic surfaces in the manifold $M(L)$, localized in the macroscopic spacetime at spacetime distances on the order of L. These are harmonic surfaces in the manifold of spacetime fields. The action $S(\lambda_H)$ of a localized lambda instanton is on the order of $g_s^{-2} = (VT)^{-1}$, so the localized lambda instantons could possibly produce interesting characteristic nonperturbative spacetime distances. I only point out here that local lambda instantons exist.

11.1 Example of a global lambda instanton

Consider a family of spacetimes in $M(\infty)$ with two toroidal dimensions. Each spacetime is the product of a two dimensional real torus with a fixed eight dimensional manifold. The two dimensional real torus is a complex 1-torus with a Kahler form proportional to a complex number σ in the upper half complex plane. The volume of the torus is the imaginary part, $\mathrm{Im}(\sigma)$. All of the other parameters describing the spacetime are held fixed, including the parameter describing the complex structure of the complex 1-torus.

The modular group is the group of fractional linear transformations of the upper half plane with integer coefficients, $\sigma \to (a\sigma+b)/(c\sigma+d)$. The modular group is generated by $\sigma \to \sigma+1$ and $\sigma \to -1/\sigma$. The two dimensional quantum field theories of the worldsurface parametrized by σ, $\sigma+1$, and $-1/\sigma$ are all equivalent. So the family of spacetimes in $M(\infty)$ is parametrized by the modular domain, which is the quotient of the upper half complex σ plane by the action of the modular group. The modular domain can be parametrized by the classical modular function $j(\sigma)$ whose values range over the entire complex plane when σ ranges over the modular domain. The family of toroidal spacetimes is parametrized by the complex j plane.

The torus becomes macroscopic in the limit $\mathrm{Im}(\sigma) \to \infty$. In this limit, $j \approx e^{-2\pi i\sigma}$. The family of spacetimes can be compactified to a 2-sphere by appending the point $j=\infty$. The compactified family of spacetimes is a complex curve of genus 0, parametrized by the complex projective j plane.

The *j-instanton* is the three parameter family of maps from the worldsurface to $M(\infty)$

$$j(z,\bar{z}) = \frac{az+b}{cz+d} \tag{11.1}$$

parametrized by complex numbers a,b,c,d satisfying $ad-bc=1$. The $\bar{j}$-instanton is the complex conjugate map. The three complex parameters are just the paramters of the group $SL_2(C)$, the conformal group of the instanton. The j-instanton and the $\bar{j}$-instanton are each three parameter families of global lambda instantons. They depend implicitly on all the other parameters of the spacetime, the parameters describing the fixed eight dimensional manifold and the complex structure of the torus.

Compactifying the family of 2-tori with the point $j=\infty$ adds a submanifold to $M(\infty)$, described by all the other parameters of the spacetime besides j. Near $j=\infty$, the two spacetime dimensions of the 2-torus become macroscopic. The volume of the macroscopic spacetime is $V = \mathrm{Im}(\sigma) \approx (2\pi)^{-1}\ln|j|$. The $j=\infty$ submanifold is part of the decompactification locus.

JHEP10(2003)063

There are three distinguished points in the modular domain, the decompactification point $j = \infty$, and two orbifold points at $j = 0$ and $j = 1728$. The point $j = 0$ corresponds to $\sigma = e^{i\pi/3}$, which is the fixed point of the Z_3 subgroup of modular transformations generated by $\sigma \to 1 - \sigma^{-1}$. The point $j = 1728$ corresponds to $\sigma = i$, the fixed point of the Z_2 subgroup of modular transformations generated by $\sigma \to -\sigma^{-1}$. The decompactification point $j = \infty$ can also be regarded as an orbifold point, left fixed by the full integer subgroup Z generated by $\sigma \to \sigma + 1$.

The three complex parameters of the j-instanton can be taken to be the three points on the worldsurface, z_3, z_2, z_∞, where the j-instanton passes through the three orbifold points, $j(z_3) = 0$, $j(z_2) = 1728$ and $j(z_\infty) = \infty$. The j-instanton can be described as a configuration of three lambda defect operators, τ_3, τ_2, τ_∞,

$$\tau_3(z_3, \bar{z}_3)\ \tau_2(z_2, \bar{z}_2)\ \tau_\infty(z_\infty, \bar{z}_\infty)\,. \tag{11.2}$$

Similarly, the $\bar{j}$-instanton is described as a configuration of three complex conjugate defect operators $\bar{\tau}_3$, $\bar{\tau}_2$, $\bar{\tau}_\infty$.

At each of the orbifold points $j = 0$, $j = 0$ or $j = 1728$, the orbifold group, or defect group, Z_3, Z_2 or Z, acts as a group of internal symmetries of the worldsurface. The worldsurface in the spacetime $j = 0$ has a Z_3 symmetry; the worldsurface in the spacetime $j = 1728$ has an internal Z_2 symmetry. There is no actual spacetime at the decompactification point $j = \infty$, so the action of the orbifold group, the integers Z, has to be defined in the limit $j \to \infty$ as an internal symmetry group of the worldsurface.

Each defect operator τ or $\bar{\tau}$ pins the worldsurface to an orbifold point in the manifold of spacetimes. The defect $\tau_3(z, \bar{z})$ pins the point z to the torus $j = 0$. The defect $\tau_3(z, \bar{z})$ pins z to the torus $j = 0$. The decompactifying defect, τ_∞ pins the point z to the decompactification locus at $j = \infty$.

Each lambda defect operator is associated to an element in the corresponding orbifold or defect group. The group element is the monodromy of the coupling constants λ^i circling the defect operator on the worldsurface. Away from the lambda defects, the scaling fields $\phi_i(z, \bar{z})$ vary adiabatically over the lambda instanton $\lambda_H(z, \bar{z})$, in a path independent fashion, because nearby general nonlinear models have the same degrees of freedom. But when a path on the worldsurface circles around one of the lambda defect operators, the scaling fields ϕ_i are transformed among themselves by the element of the orbifold group carried by the defect operator. The lambda defect acts on the worldsurface as the twist operator of the orbifolded spacetime. The lambda defect operator twists locally by its orbifold group element, projecting on the invariant degrees of freedom, removing the non-invariant degrees of freedom, and adding twist fields as new effective degrees of freedom on the worldsurface.

The nonperturbative lambda model is a two dimensional gas of lambda defect operators. At issue is the detailed dynamics of the defect gas. Is it a plasma? Or a neutral gas, with the defects all bound together? Or a combination, a plasma of some defects and some bound systems of defects?

11.2 Existence of localized lambda instantons

Near a macroscopic spacetime, there will exist lambda instantons in the target manifold $M(L)$ which are localized in bounded regions of the macroscopic spacetime. These localized lambda instantons are the harmonic surfaces in the manifold of spacetime fields.

There is a standard topological argument for the existence of instantons [17]. The localized lambda instantons are indexed by the second homotopy group, π_2, of the manifold of the target manifold $M(L)$. Every homotopy class of 2-spheres in $M(L)$ should contain local minima of $S(\lambda)$.

The target manifold $M(L)$ is the manifold of fields of the effective spacetime field theory. The spacetime fields are localized, which means that they go to zero outside a bounded region of the macroscopic spacetime, or more generally become trivial there. So the spacetime fields can be regarded as defined on a ball in n-dimensional euclidean space, where n is the dimension of the macroscopic spacetime, and the boundary of the ball can be identified to a point. Topologically, the spacetime fields can be regarded as defined on the n-sphere.

The manifold of spacetime fields is actually the manifold of gauge equivalence classes of spacetime tensor fields, including the metric tensor and the gauge fields. The manifold of localized spacetime fields is the quotient manifold F_n/G_n, where F_n is the space of tensor fields on the n-sphere and G_n is the group of local gauge transformations on the n-sphere.

The second homotopy group $\pi_2(F_n/G_n)$ is calculated using the long exact sequence:

$$\cdots \to \pi_k(F_n) \to \pi_k(F_n/G_n) \to \pi_{k-1}(G_n) \to \pi_{k-1}(F_n) \to \cdots \tag{11.3}$$

the relevant part of which is

$$\cdots \to \pi_2(F_n) \to \pi_2(F_n/G_n) \to \pi_1(G_n) \to \cdots . \tag{11.4}$$

Nontrivial topology in the manifold F_n of tensor spacetime fields comes only from the spacetime scalar fields. The space of metrics and gauge fields is topologically trivial before gauge equivalence is taken into account. The spacetime scalar fields take their values the parameters that describe the non-macroscopic dimensions of the spacetime. These are the coupling constants that parametrize the decompactification locus $M(L)_d$. The scalar fields form a map from the n-sphere to the decompactification locus. So

$$\pi_2(F_n) = \pi_{n+2}(M(L)_d) . \tag{11.5}$$

When $n = 0$, this is the homotopy group that classifies the global lambda instantons.

Localized lambda instantons formed from the spacetime scalar fields might have interesting physical effects. Locally in spacetime, they might pin to submanifolds of the decompactification locus where additional spacetime dimensions become macroscopic.

The localized lambda instantons formed from the spacetime metric and the spacetime gauge fields are indexed by the first homotopy group of the local gauge group, $\pi_1(G_n)$. If the global internal gauge group is G, then the local gauge transformations are maps from the n-sphere to G. They contribution $\pi_{n+1}(G)$ to $\pi_1(G_n)$. The local gauge transformations of the spacetime metric are the maps from the n-sphere to itself, so they contribute $\pi_{n+1}(S^n)$. For $n = 4$, these homotopy groups are typically nontrivial, so localized lambda instantons do exist.

It is not clear to me that this is a complete classification of the localized lambda instantons. In order to find the example of a global lambda instanton described above, the j-instanton, it was necessary to complete the manifold of spacetimes by adding the decompactification locus at $j = \infty$. Is there an analogous process of completion for the space of localized spacetime metrics and gauge fields modulo gauge equivalence, which would give rise to additional localized lambda instantons?

It seems quite possible that localized lambda defects will exist. A localized lambda defect would occur at a point z on the worldsurface where a localized lambda instanton $\lambda_H(z,\bar{z})$ passes through a spacetime field configuration with symmetry. The symmetry subgroup of the local spacetime gauge group would be the defect group of the localized lambda defect. The nontrivial closed path in the local gauge group G_n associated with the localized lambda instanton would then be composed of a sequence of path segments, each path segment implementing a defect twist. The homotopy argument shows the existence of local lambda instantons. They still need to be constructed explicitly. Then it can be determined whether they are smooth objects or composed of local lambda defects.

The localized lambda instantons in $M(L)$ are made from the spacetime wave modes at spacetime distances greater than L. The spacetime physics at distance L will be affected by those localized lambda instantons that are made from the spacetime wave modes λ^i at distances of the order of L. Calculations of their effects will be done locally in spacetime, in local spacetime regions at distances of the order of L.

Taking $L \to \infty$ contracts $M(L)$ to $M(\infty)$, formally. The localized lambda instantons in $M(L)$ are pushed closer and closer to the decompactification locus $M(\infty)_d$. There should be an interpretation of the limit defining a completion of $M(\infty)$ that can stand for the target manifold of the lambda model at $L = \infty$. The limit $L \to \infty$ will be a practical issue in calculations of the properties of decompactifying lambda defects, such as τ_∞. Spacetime is macroscopic in the core of a decompactifying defect. The core of the defect is dressed with localized lambda instantons in the macroscopic spacetime. The limit $\Lambda^{-1} \to 0$ will see the center of the decompactifying defect, where the difficulties of the $L \to \infty$ limit will have to be resolved.

11.3 Lambda instanton calculations

To calculate the quantum corrections to lambda instanton configurations, some way is needed to calculate the contribution of the general nonlinear model in the presence of a nontrivial lambda field $\lambda_H(z,\bar{z})$. The general nonlinear model contributes at order T^0, the same order as the one loop corrections in the lambda model. Each contributes a pre-factor multiplying the classical instanton contribution $e^{-S(\lambda_H)}$. Neither pre-factor is scale invariant separately, but only the combination.

In principle, the general nonlinear model in the presence of a lambda instanton can be made out of local two dimensional patches, the sources $\lambda^i(z,\bar{z})$ being almost constant within each patch. But I have no practical method of putting together the patches that could be used for calculation. A possibly effective method of calculation might be to treat the general nonlinear model in the presence of a lambda instanton as a correlation function of lambda defects, then calculate using the monodromy properties of the defects.

It is this difficulty of calculation that motivates the proposal of section 5 to account for the general nonlinear model contribution to the lambda model by continuing the dimension from $d = 2$ to $d = 2 + \epsilon$, dropping the general nonlinear model entirely, and determining the quantum corrections by finding the scale invariant fixed point of the effective lambda model in $d = 2 + \epsilon$ dimensions.

11.4 Speculation about the nonperturbative structure

I cannot resist indulging in some premature idle speculation about the nonperturbative lambda model. Lambda instantons will make nonperturbative corrections to the beta function $\beta^i(\lambda)$ of the general nonlinear model. It seems possible that these corrections will disturb the degeneracy of the manifold of spacetimes. I see two ways this might happen.

In the first type of scenario, nonperturbative corrections to $\beta^i(\lambda)$ simply single out some particular spacetimes from the manifold of spacetimes. These become the local minima of the effective potential function $T^{-1}a_e$. The *a priori* measure concentrates at these particular spacetimes, breaking the degeneracy. Global lambda instantons might concentrate the *a priori* measure at a particular macroscopic spacetime, at a particular point near the locus of decompactification. In that macroscopic spacetime, local lambda instantons might contribute terms to the local spacetime action $g_s^{-2}\, V\, a_e$, violating perturbative spacetime supersymmetry and giving the perturbatively massless spacetime fields definite vacuum expectation values and small masses. The original perturbative degeneracy of the manifold of spacetimes would come to be seen as merely accidental.

In the second type of scenario, the lambda instantons disorder the system. A plasma of lambda defects would accomplish this. The lambda defects would act as twist operators, projecting on the singlets of the defect group, removing the non-singlet degrees of freedom, and adding the twist degrees of freedom. The degrees of freedom λ^i would take entirely different effective forms. An effective target manifold $M(L)_e$ would replace the original target manifold $M(L)$. The effective *a priori* measure might concentrate at particular places in the effective target manifold $M(L)_e$. Or something more complicated might happen, perhaps a hierarchy of disordered systems.

A lambda instanton makes logarithmically divergent corrections to the general nonlinear model when it is configured as a 2-sphere connected to the worldsurface by an almost degenerate handle. The lambda instanton is a complex analytic curve of genus 0, so three complex parameters describe its configuration in the worldsurface. In the j-instanton, for example, the three complex parameters are the locations of the three lambda defects τ_2, τ_3, τ_∞. In the logarithmically divergent configuration, one parameter becomes the point on the worldsurface where the instanton is attached. One parameter is the thickness of the handle by which the instanton is attached. The third parameter is the point on the lambda instanton where it is attached to the worldsurface. The effective measure on the third parameter determines what states flow through the handle to appear on the worldsurface as corrections to $\beta^i(\lambda)$. If the measure on the third parameter is concentrated at a smooth point on the lambda instanton, the first scenario will apply. That point in the manifold of spacetimes will be singled out. The lambda defects will appear entirely bound. If the measure on the third parameter is concentrated at one or more of the lambda defects, the worldsurface will see a plasma of lambda defects. Intermediate possibilities might better be described as an interacting gas of lambda defects on the worldsurface. The effective measure on the configuration of the lambda instanton might depend on the spacetime distance L. The interactions among the lambda defects might depend on L, so the effective form of the degrees of freedom might change with L.

For a local lambda instanton, in the first kind of scenario, where a particular spacetime is singled out on the lambda instanton, the lambda instanton will insert local fields with logarithmically divergent coefficients into the general nonlinear model. If no spacetime supersymmetry generator can be globally defined over the lambda instanton, then the divergent insertions can violate spacetime supersymmetry. Likewise, any other spacetime symmetry can be removed, if the symmetry generator cannot be defined as a single-valued object over the lambda instanton. Perhaps even local spacetime gauge symmetry might be removed in this fashion.

In the disordered scenario, a plasma of lambda defects could distribute the *a priori* measure over the lambda instanton. Alternatively, degeneracy could be broken by pinning to the orbifold spacetimes. For example, a plasma of global decompactifying lambda defects, like the defect τ_∞ of the j-instanton, would pin the system to the locus of decompactification. The τ_2 and τ_3 defects would appear bound. This would be a novel form of decompactification, described by the orbifolded general nonlinear model at the decompactification locus. Such virtual orbifold models still need to be analyzed. The simplest case to examine is the Z orbifold of the 2-torus at $j = \infty$. There would presumably be no definite global spacetime geometry. Twisting by the defect τ_∞ would remove the angular parameter of the global spacetime geometry, the real part of σ, as a degree of freedom. In general, lambda defects will disorder angular parameters in the neighborhood of the orbifold point in the manifold of spacetimes. This is a tantalizing possibility. Mechanisms that might remove angular degrees of freedom are especially interesting because of the problem of the θ parameter in QCD.

Symmetries such as spacetime supersymmetry might also be removed as a result of twisting in a plasma of local lambda defects. If a spacetime supersymmetry generator winds nontrivially around a lambda defect, then the generator would be removed by the plasma of defects.

A form of spacetime gauge confinement could conceivably be produced by a plasma of local lambda defects twisting by elements of the local gauge group. The plasma would disorder the local spacetime gauge group, projecting on gauge singlets. The most interesting case to investigate is of course the SU(3) local gauge group in four spacetime dimensions. Perhaps this could provide a viable alternative to the hypothetical quantum field theoretic confinement of QCD. It might even be possible to find effective methods of calculation, so that a confinement mechanism in the lambda model could be checked against the experi-

JHEP10(2003)063

mental data. It would be essential that the dynamics of the lambda defects depend on L. When L drops below a characteristic confinement distance, the lambda defects would have to bind, so that the perturbative spacetime gauge theory would become visible.

I would even guess at a general principle, that the lambda model always disorders at large enough values of L. In the limit $L \rightarrow \infty$, I would expect the lambda model to explore the entire manifold of spacetimes. The effective degrees of freedom at $L = \infty$ will not be those associated with any particular spacetime, but will be constructed from the entire manifold of spacetimes by the nonperturbative fluctuations in the lambda model. Physics in any individual spacetime will give only a partial view of the large distance physics.

Undoubtedly, these speculations are far too naive, and far too much influenced by the simple-looking example of the j-instanton. The nonperturbative lambda model is likely to be a hugely complicated gas of interacting lambda defects and smooth lambda instantons. The hope is that there are relatively simple regimes at spacetime distances L which correspond to the distances in nature where relatively simple theoretical descriptions of physics have been found to apply. My speculations are offered only as suggestions of a possible complexity and richness in the nonperturbative lambda model that will be a challenge to calculation, but might yield interesting physics.

12. Spacetime gauge invariance

The lambda model needs a practical implementation of spacetime gauge invariance, including spacetime general covariance. The manifold of spacetimes is the manifold of spacetime tensor fields modulo equivalence under spacetime gauge transformations. In principle, the target manifold of lambda model is the manifold of gauge equivalence classes. But the fields $\phi_i(z,\bar{z})$ of the general nonlinear model couple to the wave modes of the spacetime tensor fields, not to the gauge equivalence classes. The coupling constants λ^i are the wave modes of the spacetime tensor fields.

Some of the fields $\phi_i(z,\bar{z})$ make no difference when they perturb the action of the general nonlinear model. These are the *redundant* fields. The redundant fields are the derivatives of spin 1 fields in the general nonlinear model. For every spin 1 field $(\chi^z_a(z,\bar{z}), \chi^{\bar{z}}_a(z,\bar{z}))$ there is a redundant spin 0 field

$$\phi^{red}_a = \partial\chi^z_a + \bar{\partial}\chi^{\bar{z}}_a\,. \tag{12.1}$$

The redundant fields in a general nonlinear model λ are certain linear combinations

$$\phi^{red}_a = G^i_a(\lambda)\phi_i \tag{12.2}$$

of the fields ϕ_i. The redundant coupling constants are the coupling constants λ^i that couple to the redundant fields.

In a macroscopic spacetime, the redundant coupling constants of the general nonlinear model are the gauge variations of the spacetime tensor fields. For example, in a macroscopic spacetime with spacetime metric $h_{\mu\nu}(x)$, each vector field $v^\mu(x)$ on the macroscopic spacetime gives a spin 1 field

$$
\begin{aligned}
v^a \chi_a^z(z,\bar z) &= v^\sigma(x)\, h_{\sigma\nu}(x)\, \bar\partial x^\nu \\
v^a \chi_a^{\bar z}(z,\bar z) &= v^\sigma(x)\, h_{\mu\sigma}(x)\, \partial x^\mu
\end{aligned}
\tag{12.3}
$$

whose derivatives give a redundant field spin 0 field

$$
v^a G_a^i \phi_i = v_* h_{\mu\nu}(x) \partial x^\mu \bar\partial x^\nu \tag{12.4}
$$

which represents the infinitesimal gauge transformation of the spacetime metric produced by the vector field $v^\mu(x)$.

The $G_a^i(\lambda)$ form a Lie algebra of vector fields on the manifold of spacetimes

$$
G_a^j \partial_j G_b^i(\lambda) - G_b^j \partial_j G_a^i(\lambda) = F_{ab}^c G_c^i(\lambda)\,. \tag{12.5}
$$

The F_{ab}^c are the structure constants of the Lie algebra of redundancy transformations in the general nonlinear model, which is the Lie algebra of local gauge transformations in spacetime.

The manifold of spacetimes is parametrized by the coupling constants λ^i modulo the redundant coupling constants. The lambda model must respect the equivalence relations given by redundancy in the general nonlinear model. If the lambda model respects redundancy, then gauge invariance in any macroscopic spacetime will follow automatically. In particular, the *a priori* measure will respect equivalence under redundancy. The effective spacetime quantum field produced by the lambda model in a macroscopic spacetime will be gauge invariant.

The lambda field $\lambda(z,\bar z)$ can be pictured as a map to the manifold of redundancy equivalence classes, but the component lambda fields $\lambda^i(z,\bar z)$ would then couple ambiguously to the fields $\phi_i(z,\bar z)$, up to arbitrary admixtures of redundant fields.

Instead, let there be a lambda field $\lambda^i(z,\bar z)$ for each spin 0 field $\phi_i(z,\bar z)$ in the general nonlinear model, including the redundant fields. Then introduce auxiliary spin 1 sources $(\xi_z^a, \xi_{\bar z}^a)$ to implement spacetime gauge invariance. Couple each spin 1 field $(\chi_a^z, \chi_a^{\bar z})$ in the general nonlinear model to a spin 1 source $(\xi_z^a, \xi_{\bar z}^a)$, adding

$$
\int \mathrm{d}^2 z\, \mu^2 \frac{1}{2\pi}\, \left[\xi_z^a(z,\bar z)\, \chi_a^z(z,\bar z) + \xi_{\bar z}^a(z,\bar z)\, \chi_a^{\bar z}(z,\bar z)\right] \tag{12.6}
$$

to the action of the general nonlinear model

The general nonlinear model is now locally invariant under infinitesimal gauge transformations

$$
\lambda^i \to \lambda^i + \epsilon^a(z,\bar z)\, G_a^i(\lambda) \tag{12.7}
$$

if at the same time the auxiliary lambda fields are transformed by

$$
\begin{aligned}
\xi_z^a &\to \xi_z^a + \partial\epsilon^a \\
\xi_{\bar z}^a &\to \xi_{\bar z}^a + \bar\partial\epsilon^a\,.
\end{aligned}
\tag{12.8}
$$

The action density of the general nonlinear model changes by the total derivative

$$\partial(\epsilon^a \chi_a^z) + \bar{\partial}(\epsilon^a \chi_a^{\bar{z}}))\,. \tag{12.9}$$

A locally gauge invariant action $S(\lambda, \xi)$ for the lambda model will be determined by the scale variation of the general nonlinear model, as in section 6. But the form it will take is obvious. The two dimensional derivatives $\partial\lambda^i$ and $\bar{\partial}\lambda^i$ are simply replaced in $S(\lambda)$ by the covariant derivatives

$$\begin{aligned} \mathrm{D}\lambda^i &= \partial\lambda^i - G_a^i(\lambda)\xi_z^a \\ \bar{\mathrm{D}}\lambda^j &= \bar{\partial}\lambda^j - G_a^j(\lambda)\xi_{\bar{z}}^a\,. \end{aligned} \tag{12.10}$$

A localized lambda instanton is now a local minimum of $S(\lambda, \xi)$. The covariant derivatives $\mathrm{D}\lambda^i$ and $\bar{\mathrm{D}}\lambda^i$ must go to zero as $z \to \infty$. The auxiliary lambda field $(\xi_z^a, \xi_{\bar{z}}^a)$ is a 1-form on the complex plane with values in the Lie algebra of infinitesimal spacetime gauge transformations. Its path-ordered integrals are group elements in the spacetime gauge group. Along a closed contour around $z = \infty$, the indefinite path ordered integral from a fixed starting point gives a closed loop in the group G_n of spacetime gauge transformations, representing the element in $\pi_1(G_n)$ that indexes the localized lambda instanton.

The renormalization of the general nonlinear model respected general covariance in the target manifold [3]. No particular symmetry of the target manifold was assumed. Renormalization of target manifold symmetry was subsumed in renormalization of target manifold general covariance. The renormalization of general covariance in the target manifold was subject to possible obstructions which were cohomology classes on the target manifold, the nonlinear model anomalies.

The target manifold of the lambda model is the manifold of spacetimes. The spacetime gauge symmetries are internal symmetries of the lambda model, analogous to spacetime symmetries in the general nonlinear model. The renormalization of spacetime gauge symmetry in the lambda model is subsumed into the renormalization of reparametrization invariance in the manifold of spacetimes. Potential nonlinear model anomalies in the lambda model would obstruct renormalization of reparametrization invariance in the manifold of spacetimes, and might show themselves in spacetime quantum field theory as gauge anomalies. It will have to be shown that the lambda model is free from anomalies.

13. What needs to be done

The most urgent task now is to find all the local lambda instantons in explicit form, and develop concrete methods for calculating their contributions to the effective beta function of the general nonlinear model. Temporarily assume a particular macroscopic spacetime and assume a fixed small value for the spacetime coupling constant g_s, in order to find out if the lambda model actually does remove spacetime supersymmetry, produce small nonzero masses, and lift the degeneracies that are local in the macroscopic spacetime.

Developing effective methods of calculation will require filling in details of my arguments for the structure of the theory, or finding better arguments. The most essential elements include the principle of tandem renormalization and the effective renormalization group invariance of the effective general nonlinear model, which together imply the tautological scale invariance of the effective lambda model. Also crucial is the identification the action $S(\lambda)$ with the scale variation of the general nonlinear model, which is used to establish the gradient property and the spacetime action principle in macroscopic spacetimes.

Details of the action $S(\lambda)$ also need to be filled in. This should be straightforward, since $S(\lambda)$ is completely determined by the scale variation formula, equation (6.14). Lambda fields $\lambda^i(z,\bar{z})$ are introduced as sources for all the scaling fields $\phi_i(z,\bar{z})$ that occur in the general nonlinear model of the worldsurface. The action $S(\lambda)$ is read off from the scale variation of the general nonlinear model in the presence of those sources.

In particular, several special scaling fields $\phi_i(z,\bar{z})$ occur in the string worldsurface, made entirely from worldsurface ghost fields. The coupling constants λ^i that couple to these special scaling fields play distinguished roles [18, 19]. One special bosonic coupling constant λ_D has the effect of shifting the value of the number $\ln(T)$. It is conjugate in the metric $T^{-1}g_{ij}(\lambda)$ to a second special bosonic coupling constant λ'_D, which is redundant, at least in a scale invariant worldsurface. When these special coupling constants are made into lambda fields $\lambda_D(z,\bar{z})$ and $\lambda'_D(z,\bar{z})$, it appears that $\lambda'_D(z,\bar{z})$ can be interpreted as the logarithm of the local two dimensional scale factor $\Lambda(z,\bar{z})$ and acts as a Lagrange multiplier, enforcing local two dimensional scale invariance. The combined coefficient of the two dimensional curvature density $\Lambda^2 R_2(\Lambda)$ from the combined local lagrangians of the general nonlinear model and the lambda model, is the sum of the special coupling constant λ_D, the number $\ln(T)$, and the potential function $T^{-1}a(\lambda)$. This seems worth pursuing. The details of the system of special lambda fields need to be worked out. It might be that they play only a formal role in the lambda model. But it is also possible that there will be some indication of how the number T might be determined.

A second basic detail that needs filling in is the possible antisymmetric tensor coupling in the lambda action. The heterotic worldsurface is chirally asymmetric. The scale variation of the general nonlinear model of the heterotic worldsurface can contain a graded antisymmetric tensor coupling $\Theta b_{ij}(\lambda)$ in addition to the graded symmetric metric coupling $T^{-1}g_{ij}$. It would be surprising if an antisymmetric coupling did not appear.

If calculation shows that the lambda model can in fact produce the needed local effects in a macroscopic spacetime, then there will be two obvious directions to take. One will be a renewed search among the possible macroscopic spacetimes for a match to the standard model. The lambda model will produce a local quantum field theory in each macroscopic spacetime. Methods will be needed to winnow the macroscopic spacetimes for promising candidates to compare in detail with the standard model.

It will also become promising to investigate basic issues, including decompactification mechanisms, mechanisms that could determine the spacetime coupling constant $g_s^2 = V\,T$, cosmological interpretation, the construction of real time, a mechanism that could fix the number T or the dimension $d = 2 + \epsilon$, topology change, and the issue of security in the limit $\Lambda^{-1} \to 0$, $L \to \infty$.

The limit $L \to \infty$ raises two questions. First is simply the existence of a scale invariant limit, without which the lambda model would have no foundation on which to build the large distance physics. If the lambda model does have a scale invariant limit at $\Lambda^{-1} = 0$, the question becomes, does degeneracy remain in the limit? Whatever form the effective degrees of freedom λ_e^i take, do any of them have vanishing effective anomalous dimension $\gamma_e(i) = 0$? Are there marginal effective coupling constants λ_e^i in the short distance limit? If not, if all degeneracy is lifted at $L = \infty$, then the logarithmic divergence will be removed. The original purpose of the lambda model will be realized. This last question evokes the historical roots of the two dimensional nonlinear model and the lambda model in the ideas of Bloch, Hohenburg, Mermin, Wagner and Coleman about the logarithmic divergences of spin waves in $d = 2$ dimensions, and the physical consequences of their impossibility for two dimensional physics.

14. Discussion

The lambda model is a theory of physics which has a fundamental unit of spacetime distance and works entirely at large distance compared to that fundamental unit. The lambda mode appears capable of explaining some of the most basic principles of physics. It appears capable of constructing quantum mechanics in spacetime, and determining the hamiltonian. The lambda model appears capable of doing this without assumptions about physics at experimentally inaccessible spacetime distances near the Planck length.

It seems futile to speculate about small distance physics without having in hand a coherent and testable theory of large distance physics, given the enormous gulf between the Planck length and the length scales of practical experiments. Without a means of reliably predicting observable large distance physics, how can a speculative theory of small distance physics be checked against the real world? There is considerable room for surprise in the roughly 14 or 15 orders of magnitude between the Planck length and the smallest distances where theories can be checked. It might be worth remembering that past explorations over 14 or 15 orders of magnitude in distance discovered such surprises as quantum mechanics and the elementary particles. What could possibly justify theoretical assumptions about physics across such an enormous gulf of spacetime distances, if those theoretical assumptions cannot lead to definite statements that can be checked in the real world?

The lambda model presents the possibility of exceptions to the well-supported principle that the physics of the large is completely explained by the physics of the small. Under conditions of degeneracy, the lambda model may produce nonperturbative effects in spacetime which are not explicable on atomistic principles. If such effects can be derived from the lambda model, and confirmed by experiment, it will be a salutary reminder that knowledge in physics is always incomplete, no matter how striking the success of existing theory. There is a temptation to extrapolate successful theories far beyond the extent of their demonstrated reliability, especially after the past successes of atomistic physics. A theory which succeeds at describing all available experimental results in a certain regime of distances, such as the standard model of particle physics does now, is assumed to explain *in principle* all the complicated phenomena observed at larger distances, if only the necessary difficult calculations could be carried out. Even when many such complicated phenomena are successfully explained, there is no guarantee that all large distance phenomena will be explained. There still remains a remote possibility that subtle unexpected effects are yet to be observed. To search at random for such effects is unlikely to be useful. Guidance is needed from a highly credible theory. The lambda model is proposed as a theory that might be capable of acquiring such credibility and also predicting unexpected phenomena.

The crucial advantage that the lambda model might have over a fundamentally atomistic model of physics is the security that the lambda model could give at large distance in spacetime by building physics from the limit $L = \infty$ downwards in L. Infrared security in the lambda model would eliminate the need to guess at the nature of microscopic physics at unobservably small distances in spacetime. The need for some such infrared security is suggested by the miniscule value of the observed cosmological constant, which seems inexplicable in any atomistic version of spacetime physics.

The lambda model is an attempt to make a weakly coupled theory of physics. Weak coupling means that the spacetime coupling constant g_s should be a reasonably small number, say on the order of $1/10$. The value of the number T is a separate matter. The lambda model undoubtedly needs T to be an extremely small number. The dimension $d = 2 + \epsilon$ must be very near 2, otherwise the entire analysis and physical interpretation of the lambda model would break down. The spacetime coupling constant g_s emerges only in a macroscopic spacetime of volume V, by the relation $g_s^2 = VT$. The lambda model does not seem to require that g_s be small. The lambda model might well be a strongly coupled two dimensional quantum field theory in some spacetime regimes. The lambda model might still be useful there, if it happens to be an integrable two dimensional field theory. The most obvious prospects of the theory, however, seem to call for weak coupling. For example, it is difficult to imagine how a spectrum of exponentially large spacetime distances could arise without a small spacetime coupling constant.

I retain a naive hope that a weakly coupled theory of large distance physics can succeed in explaining the standard model of the elementary particles. It is remarkable that all observed couplings of the standard model are in fact weak at the smallest distances accessible to experiment. The weakness of all the observed couplings is one of the most striking results from high energy experimental physics. It seems to me misguided to turn away from the possibility of a weakly coupled theory before having in hand a coherent method to determine large distance physics. A systematic weakly coupled theory of large distance physics would be so useful that nothing but a definitive demonstration of infeasibility should forestall the attempt. In the end, of course, the assumption of weak coupling must be justified dynamically, since the spacetime coupling constant is a parameter of the manifold of spacetimes.

The lambda model is mathematically universal. The target manifold, the metric coupling, the potential function are all mathematically natural objects. The couplings of the lambda model satisfy mathematically natural differential equations on the manifold of spacetimes, expressing generalized two dimensional scale invariance. No arbitrary choices are made.

The lambda model is *not* universal in the in sense of quantum field theory. As a nonlinear model, it is scale invariant in the generalized sense. Its couplings are at a fixed

JHEP10(2003)063

point of the renormalization group of the general nonlinear model whose target manifold is the manifold of spacetimes. The fixed point is not stable under the renormalization group. If a small perturbation were made, the renormalization group would drive the nonlinear model far away from the fixed point. There are infinitely many unstable directions. Perturbing the action density by any function $f(\lambda)$ on the target manifold gives a dimension 2 perturbation, which would grow quadratically in the two dimensional distance, freezing the lambda field to the minimum of the function $f(\lambda)$. A dimension 2 perturbation would freeze the system into a fixed spacetime, suppressing the fluctuations of the lambda field that are needed to cancel the divergence due to local handles in that spacetime. The logarithmic divergence would return.

The lambda model must be held at the fixed point. All relevant perturbations of the lambda model must be tuned to zero. There might be a formal apparatus in the lambda model, perhaps involving the special lambda field λ'_D, that enforces this tuning. Or there might be a deeper mechanical explanation. If some mechanical model of the string worldsurface automatically gives rise to the lambda model, it would presumably hold the lambda model precisely at the fixed point.

However the tuning is done, by hand if necessary, it is possible to carry out the task because the couplings of the lambda model do not fluctuate. The couplings of the lambda model are classical geometric quantities on the manifold of spacetimes.

The lambda model produces a probabilistic description of spacetime. It may single out a number of possible macroscopic spacetimes. In each, the spacetime fields describing geometry and matter fluctuate according to a quantum field theory produced by the lambda model. But the geometry that defines the lambda model does *not* fluctuate. The couplings of the lambda model take definite values satisfying classical differential equations on the manifold of spacetimes. Strict causality, which was renounced in spacetime when quantum mechanics was discovered, might be regained at another level of abstraction.

If the theory works, part of the *a priori* measure on the manifold of spacetimes will be found to concentrate at a spacetime that matches our spacetime in its dimension, its cosmology, and in the phenomenology of its elementary particles. Our spacetime might turn out to be only one of many where the *a priori* measure concentrates, and might carry only a small share of the total measure. It would become a challenge to devise experiments that could detect the other possibilities.

Acknowledgments

Very early in my thinking about the lambda model, circa 1988, I was stimulated by a comment from V. Periwal and a comment from S. Shenker. I am grateful for their help.

This work has been supported since 1989 by Rutgers, The State University of New Jersey through its New High Energy Theory Center. Since 1986, much of the work has been done in Reykjavík, during visits to Raunvísindastofnum Háskólans Íslands, The Natural Science Institute of the University of Iceland.

The United States Department of Energy provided support from 1978 until 2001, first through the Lawrence Berkeley Laboratory of the University of California, Berkeley during

1978-1980, then through the Fermi Institute of the University of Chicago during 1981-1989, then through the Rutgers New High Energy Theory Center during 1989-2001. The Centre dés Études Nucleaire de Saclay, France provided support during 1980-1981. A fellowship from the Sloan Foundation provided support during 1981-1983, and a fellowship from the MacArthur Foundation during 1987-1993.

I am grateful to all the people and institutions involved. I am especially grateful to I. Singer, Y. Nambu and L. Kadanoff for arranging support during the first years.

References

[1] D. Friedan, unpublished lectures given at the Mathematics and Physics Conference (Durham, England, July 1979), at the Nuffield Workshop on Quantum gravity (Cambridge, England, August 13, 1979), at the Nato Advanced Study Institute on Recent Developments in Gauge Theories (Cargese, France, August 1979), and unpublished manuscript, *Geometric models for critical systems in $2+\epsilon$ dimensions,* subtitled *Expanded version of a talk given at the Nuffield Workshop on Quantum Gravity, DAMTP, Cambridge University, August 13, 1979,* commissioned by the directors of the 1979 Nuffield Workshop, and privately circulated to them at the conclusion of the workshop. Scanned images of the manuscript are available on the World Wide Web at `http://physics.rutgers.edu/~friedan/Nuffield79/`

[2] D. Friedan, *Nonlinear models in $2+\epsilon$ dimensions*, *Phys. Rev. Lett.* **45** (1980) 1057.

[3] D. Friedan, Ph.D. thesis, University of California at Berkeley (August 1980), published as Lawrence Berkeley Laboratory Publication LBL-11517 (August 1980); *Nonlinear models in $2+\epsilon$ dimensions*, *Ann. Phys. (NY)* **163** (1985) 318.

[4] E. Calabi, *On Kähler manifolds with vanishing canonical class*, *Proc. Int. Congr. Math. Amsterdam* **2** (1954) 206.

[5] S.T. Yau, *On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation I*, *Proc. Natl. Acad. Sci. U.S.A.* **74** (1977) 1798; *On Calabi's conjecture and some new results in algebraic geometry*, *Comm. Pure Appl. Math.* **31** (1978) 339.

[6] L. Alvarez-Gaumé and D.Z. Freedman, *Kähler geometry and the renormalization of supersymmetric sigma models*, *Phys. Rev.* **D 22** (1980) 846.

[7] C. Lovelace, *Strings in curved space*, *Phys. Lett.* **B 135** (1984) 75.

[8] D.J. Gross, J.A. Harvey, E.J. Martinec and R. Rohm, *The heterotic string*, *Phys. Rev. Lett.* **54** (1985) 502.

[9] E.S. Fradkin and A.A. Tseytlin, *Effective field theory from quantized strings*, *Phys. Lett.* **B 158** (1985) 316.

[10] C.G. Callan Jr., E.J. Martinec, M.J. Perry and D. Friedan, *Strings in background fields*, *Nucl. Phys.* **B 262** (1985) 593.

[11] A.B. Zamolodchikov, *Irreversibility of the flux of the renormalization group in a 2-d field theory*, *Sov. Phys. JETP Lett.* **43** (1986) 730.

[12] A.B. Zamolodchikov, *Renormalization group and perturbation theory near fixed points in two-dimensional field theory*, *Sov. J. Nucl. Phys.* **46** (1987) 1090.

[13] C. Lovelace, *Stability of string vacua, 1. A new picture of the renormalization group*, *Nucl. Phys.* **B 273** (1986) 413.

[14] A.M. Polyakov, *Quantum geometry of bosonic strings*, *Phys. Lett.* **B 103** (1981) 207; *Quantum geometry of fermionic strings*, *Phys. Lett.* **B 103** (1981) 211.

[15] D. Friedan, E.J. Martinec and S.H. Shenker, *Conformal invariance, supersymmetry and string theory*, *Nucl. Phys.* **B 271** (1986) 93; See also, *Notes on string theory and two dimensional conformal field theory*, in *Proceedings of the workshop on unified string theories*, M.B. Green and D. Gross (eds.), World Scientific (1986).

[16] E. Verlinde and H. Verlinde, *Multiloop calculations in covariant superstring theory*, *Phys. Lett.* **B 192** (1987) 95.

[17] I.M. Singer, *Some remarks on the Gribov ambiguity*, *Commun. Math. Phys.* **60** (1978) 7.

[18] J. Polchinski, *Factorization of bosonic string amplitudes*, *Nucl. Phys.* **B 307** (1988) 61.

[19] H. La and P. Nelson, *Effective field equations for fermionic strings*, *Nucl. Phys.* **B 332** (1990) 83.