Integrability in supersymmetric gauge theories. I

A. V. Marshakov
Lebedev Physics Institute, Moscow and Institute for Theoretical and Experimental Physics, Moscow

Fiz. Élem. Chastits At. Yadra **30**, 1120–1210 (September–October 1999)

This is the first part of a review devoted to the exact, nonperturbative solutions of supersymmetric gauge theories and their formulation in terms of integrable systems. The general phenomenon of integrability as it appears in the formulation of effective actions for various models of topological, low-dimensional string theories and almost realistic supersymmetric gauge field theories is discussed. First the basic features of the string-theory path integral are discussed in order to understand better the nonperturbative properties of the theory. Then a formulation of the exact effective actions based on systems of nonlinear differential equations is proposed. It is demonstrated that the resulting nonlinear differential equations belong to the class of integrable models of the Kadomtsev–Petviashvili and Toda type. Particular models of this class are discussed, with special focus on the integrable systems appearing in the context of multidimensional supersymmetric gauge theories. Their Lax representations and spectral curves are studied in detail, and a classification of the exact solutions of $N=2$ supersymmetric gauge theories along these lines is proposed. © *1999 American Institute of Physics.* [S1063-7796(99)00205-3]

1. INTRODUCTION

In the last few decades, the two problems of greatest interest in theoretical physics (elementary-particle physics) have been the fundamental problems of the confinement of quarks inside hadrons and the quantum theory of gravity. An essential feature of these problems is the fact that progress in understanding and solving them is impossible without studying the properties of non-Abelian gauge field theory—quantum chromodynamics—and the general theory of relativity with strong coupling, where essentially all the standard field-theoretic methods used in quantum electrodynamics and the Weinberg–Salam model of electroweak interactions and based on perturbation theory are inapplicable. Progress in understanding the key current problems in elementary-particle physics and the theory of gravitation is therefore impossible unless some essentially new methods are developed which allow physical theories to be studied in the nonperturbative regime.

A historically important event was the recognition of the fact that the complex nonlinear equations on which the classical limit of the corresponding quantum theories is based are integrable, and, moreover, at least some of their solutions can be constructed explicitly. For example, the instanton solutions of the classical equations discovered by Polyakov[1,2] (and, in particular, the instanton solutions of the Yang–Mills equations found beginning with the study by Belavin, Polyakov, Schwarz, and Tyupkin[3]) turned out to influence significantly our understanding of not only the classical but also the quantum structure of sigma models and gauge field theories (Ref. 4; see also Ref. 5 and references therein). Moreover, it was instantons which first demonstrated the importance of applying the techniques of complex analysis[6] to the solution of modern nonlinear problems in theoretical physics.

The appearance of instantons and other nonperturbative solutions greatly broadened the horizons of the theory of strong interactions and clearly demonstrated that elementary-particle physics does not reduce to perturbation theory, in which only high energies can be treated in QCD (the regime of asymptotic freedom), where the standard formulation of gauge field theory works well.[7] Nevertheless, instanton calculations proved to be only the next-highest approximation in QCD and are manifestly insufficient for describing quark confinement and other effects in the strong-coupling region. As far as the quantization of the general theory of relativity is concerned, even the appearance of supersymmetry[8] as a mechanism for canceling divergences has not given any hope of constructing a consistent theory of quantum gravity within the framework of quantum field theory (see, for example, Ref. 9).

It has turned out that the construction of a consistent picture of quantum gravity requires a fundamental change in the theory at Planck scales, based on going from pointlike objects to one-dimensional, extended objects—strings. The appearance[10] and development of string theory primarily led, as noted by Scherk and Schwarz,[11] to the unification of gauge field theory and gravity, so that the two fundamental problems ceased to exist independently, because the string spectrum naturally contains *massless* vector fields and spin-2 fields. The geometric structure of string theory has been formulated by Polyakov as an integral over two-dimensional geometries:[12]

$$F_g=\int Dg_{ab}D\mathbf{x}e^{-\int_{\Sigma_g}|\partial\mathbf{x}|^2};\quad \mathcal{F}=\sum_g \Lambda_{\rm str}^g F_g, \tag{1.1}$$

where $\mathbf{x}$ are the fields of two-dimensional conformal field theory or the string coordinates, g_{ab} are two-dimensional metrics on the Riemann surface Σ_g of genus g, whose equivalence classes (with respect to coordinate transforma-

1063-7796/99/30(5)/39/$15.00

tions on the unobservable world sheets) correspond to various two-dimensional geometries, and $\Lambda_{\rm str}$ is the string coupling constant. According to the Belavin–Knizhnik theorem,[13] the integral (1.1) reduces to an integral over the moduli space of complex structures of Riemann surfaces of the special form

$$F_g=\int_{\mathcal{M}_g} d\mu(y)|f(y)|^2, \qquad (1.2)$$

where $\mathcal{M}_g$ is the (finite-dimensional) moduli space of complex structures of the Riemann surface Σ_g, and the specific form of the integration measure depends on the choice of string model. As will be seen below (Sec. 2), the formulation (1.1) and (1.2) in principle allows the use of symmetry considerations to obtain nonperturbative information, although according to its actual definition it is a perturbative expansion about some vacuum, and the integral (1.1) gives the g-loop correction in the theory.

The development of string theory led to the appearance of more or less realistic string models (see, for example, Ref. 14 and references therein), based on representations of the real world as a part of a multidimensional space-time (10-dimensional, in most cases) whose unphysical dimensions are compactified and represent a manifold of a special type (for example, a Calabi–Yau manifold). The structure of compactified string models allows the assumption that there exists a completely nontrivial symmetry (duality)[15–17] relating the various string theories to each other, in particular, in such a way that the perturbative regime in one gives some information about nonperturbative effects in another. In other words, duality transformations allow the various string models to be treated as perturbative expansions (1.1) about different vacua of the same theory. A defect in this picture which remains to this day is the absence of any rigorous mathematical statements.

On the other hand, it is in string theory (more precisely, in some of its simplest models) that it first turned out to be possible to obtain nonperturbative information about the (exact) correlation functions. Owing to the limitations of direct methods of calculation, i.e., directly using the Polyakov functional integral (1.1), which can be used only to calculate critical exponents and the simplest correlation functions on the sphere,[19,20] progress has been made with an elegant discretization of $c<1$ string models or matrix models of two-dimensional gravity:[22]

$$Z[V]=\int DM_1\ldots DM_k e^{-v(M_1,\ldots,M_k)}, \qquad (1.3)$$

where $DM_\alpha \sim \Pi_{i,j} dM_{\alpha,ij}$ denotes the integral over a finite matrix, and $V(M_1,\ldots,M_k)$ is usually a polynomial potential. Here the loop expansion reproduces a discretized version of the loop expansion of $c\leq 1$ string models (c is the central charge of the Virasoro algebra, the set of degrees of freedom of the corresponding string theory). The double scaling limit of (1.3) proposed by Kazakov[23] has permitted study of the nonperturbative partition function (or, more generally, the generating function of the string correlators) $\mathcal{F} \sim \log Z$, information on which can be encoded in nonlinear integrable equations.[24–28] One method of obtaining differential equations for the generating function (1.3) is by studying the loop equations or Ward identities for the integral (1.3) (Refs. 25–27), $\langle \delta V\rangle=0$ [the average is understood in the sense of the partition function (1.3)], which are essentially a very simple analog of the Ward identities in gauge field theory.[30]

Another example (which, however, turns out to be related to the previous one) of the direct manifestation of integrable structures is that of *topological* string models,[31,32,35–37] in which the theory is actually *defined* by the fact that its correlation functions "count" the topological characteristics of the moduli spaces of complex curves—the intersection indices, Euler characteristics, and so on. Here the integral over the moduli space of the complex structures (1.2) (more precisely, over its compactification $\bar{\mathcal{M}}_g$, which is usually taken to be the Deligne–Mumford compactification) counts the cohomology classes $\bar{\mathcal{M}}_g$, and from the viewpoint of field theory this implies that the physical degrees of freedom correspond to only a finite number of topological operators which can be constructed, for example, in the language of the cohomologies of the BRST operator.[38,39] The effective construction of the generating function of the intersection indices in the form of a matrix model allowed Kontsevich[40] to come close to formulating nonperturbative topological string models as solutions of integrable equations of a special type.[33,34]

Further developments showed that it is integrable systems which provide a suitable language in which to describe the nonperturbative solutions of quantum theories. Moreover, the formulation in terms of integrable systems is universal, i.e., it is common to many string models which *a priori* appear to be completely different from each other. For example, the important advance made by Witten and Seiberg[41,42] in understanding the nonperturbative structure of four-dimensional $\mathcal{N}=2$ supersymmetric non-Abelian gauge theory allowed these exact nonperturbative results to be stated in the language familiar from $c<1$ string theory, i.e., to be described as a deformation of the solutions of a hierarchy of integrable equations of the Kadomtsev–Petviashvili or Toda type,[43] and it indicated the essentially stringy nature of the exact solutions of quantum field theory.[44–47]

It should be particularly noted that, although the development of this area of string theory is still far from any direct relation to the problems of confinement and real quantum gravity, the universality which characterizes this approach to solving problems in non-Abelian gauge field theory and the general theory of relativity is worthy of attention. Conceptually, the approach based on integrable systems currently encompasses a wide spectrum of problems from QCD at high energies (the Lipatov approach[48]) to exactly solvable models of two-dimensional[23,34,40] and three-dimensional[49] gravity.

An advantage of the language of integrable systems is its relative simplicity: instead of string and field systems with an infinite number of degrees of freedom, it is possible to work with essentially finite-dimensional (because one often deals with reductions) dynamical systems, for the study of which a well developed formalism exists. The most convenient approaches for working with nonperturbative string and field theories have proved to be the algebraic–geometric approach of Novikov, Krichever, Dubrovin *et al.*,[50–53] and also the Hamiltonian approach of the Leningrad school of Faddeev[54,55] and the Japanese fermionic formalism,[56,57] which leads to a formulation of the solutions of integrable systems in terms of an infinite-dimensional Grassmannian—a universal moduli space.[58–60]

It should also be noted that the development of nonperturbative quantum field theory and string theory in turn has had a significant effect on the theory of integrable systems, as it has made it necessary to study new, in general, singular solutions of integrable equations in more detail.[34,40,63] The singular properties of this class of solutions (which henceforth will be called *string* solutions) are directly related to their physical meaning—identification with the sum of the perturbation series, which in general is an asymptotic series. Nevertheless, it has turned out that in the zeroth-order approximation these solutions correspond to well known problems in finite-gap integration, and the construction of exact solutions is related to Whitham deformations of finite-gap solutions.[36,51,61,62]

Studies from the last few years (see, for example, Refs. 23, 25, 34, 40–43, 64, and 65) have shown that it is possible to calculate *exactly* nonperturbative results (the spectrum, correlation functions, and effective actions) in theories which are not *quantum-integrable* models[54,66] in the canonical sense of this term. In contrast to "canonical" quantum integrable systems, where there is usually an (infinite-dimensional) quantum symmetry algebra allowing the Hilbert space of the theory to be identified with its representation space, and a sufficient number of conditions to be imposed on the correlation functions, the theories which we shall discuss below are not quantum-integrable in this sense. However, they are certainly interesting from the viewpoint of being close to realistic models of elementary-particle theory and are exactly solvable in the nonperturbative regime in the following sense.

For each of the examples discussed, *there exists* an effective description in terms of the generating functional of the exact correlation functions in the theory,

$$\langle \mathcal{O}_{i_1} \ldots \mathcal{O}_{i_n} \rangle = \frac{\delta^n \mathcal{F}}{\delta_{a_1} \ldots \delta_{a_n}}, \tag{1.4}$$

and/or the effective action. It should be especially noted that the method of the effective action[67] is natural for formulating string theory, where all string effects are important only at short distances. The effective theory can be formulated in terms of a (classical) integrable system. Moreover, in all the cases considered below, this integrable system turns out to be a hierarchy of integrable equations of the Kadomtsev–Petviashvili (KP) type defined by reduction. The first equation of the hierarchy is

$$3\frac{\partial^2 U}{\partial T_2^2} = \frac{\partial}{\partial T_1}\left(4\frac{\partial U}{\partial T_3} - 12U\frac{\partial U}{\partial T_1} - \frac{\partial^3 U}{\partial T_1^3}\right) \tag{1.5}$$

or the two-dimensional Toda lattice, for which the first equation has the form

$$\frac{\partial^2 \phi_n}{\partial T_1 \partial \bar{T}_1} = e^{\phi_{n+1}-\phi_n} - e^{\phi_n - \phi_{n-1}}, \tag{1.6}$$

although, in general, the restriction of the class of integrable models is due solely to the fact that we are dealing with the string theories which are simplest from this point of view: two-dimensional (2D) topological string models and $c \leq 1$ theories interacting with two-dimensional quantum gravity (Secs. 2 and 3), and also the $\mathcal{N}=2$ supersymmetric non-Abelian gauge field theories arising in the point (field) limit of realistic string theories (Sec. 4). The effective formulation is universal in the sense that it is independent of very many of the properties of the "bare" theory, such as the dimension of space-time: two-, four-, and even five-dimensional theories look practically the same from this point of view. Moreover, the resulting effective theories are in many respects reminiscent of *topological* field theories, as they possess a number of properties characteristic of two-dimensional topological theories, even though the bare theories are certainly multidimensional. In particular, the spectrum of the effective theories contains massless propagating particles.

To be more precise, we shall understand an effective nonperturbative formulation as the construction of an explicit dependence of the spectrum, the correlation functions, and the effective actions (i.e., the effective coupling constants) as functions of the parameters (or moduli) of the theory, which, as a rule, are the low-energy values of the background fields in physical space-time. For example, in 4D supersymmetric gauge theories, these are the vacuum expectation values of the Higgs fields, $h_k = (1/k)\langle \mathrm{Tr}\phi^k \rangle$, while in more general string theories they are the moduli of the space-time metric (for example, the parameters of a complex or Kähler structure), the gauge fields (the moduli of flat connections or instantons), and so on. The problem amounts to finding the exact nonperturbative dependence of the physical quantities on these parameters.[1)] The solution of this problem is considerably simplified by the fact that the space-time moduli in string theories can often be identified with the moduli of complex manifolds, the space of which is usually a factor of a topologically trivial manifold under the action of a discrete group. The action of this discrete group is hypothetically identified with duality transformations[15–17] relating the various perturbative expansions of the string theory to each other.

It is primarily the complex-analytic structure which distinguishes the class of theories for which exact nonperturbative results will be formulated below. Here the possibility immediately arises of posing problems which are technically solvable, because many results can be formulated in terms of *holomorphic* (or meromorphic) functions. The idea of working with holomorphic functions originated with the use of complex analysis in instanton theory[6] and the Belavin–Knizhnik theorem[13] in perturbative string theory. Secondly, the class of problems which can be studied is restricted even more by the fact that in it the moduli of physical theories can be identified with the moduli of *one-dimensional* (1D) complex manifolds—(space-time!) complex curves Σ (or two-dimensional real manifolds—Riemann surfaces). Here two explanatory remarks should be made. First, space-time Riemann surfaces *a priori* do not have anything to do with world sheets in string theory; this does not at all hinder the use of the same technical tools as in perturbative string theory when working with them. Second, in principle one should expect the same picture for theories in which the moduli spaces are identified with the moduli spaces of complex manifolds of higher dimension ($K3$, $\dim_{\mathbf{C}}=3$ Calabi–Yau manifolds, and so on). Moreover, in the unified picture of string theory the space-time complex curves considered below must often be assumed to be degenerate cases of string compactification manifolds: when a Calabi–Yau manifold effectively degenerates into a $1_{\mathbf{C}}$D curve Σ (Ref. 47). The nontrivial topological structure of the spectral curve Σ is essentially nonperturbative information; in perturbation theory the spectral curve is manifested only "locally," as a scale parameter. This means that it is string effects which play the essential role in the structure of the exact nonperturbative solutions of gauge field theory, while the topological degrees of freedom important for constructing the effective theory are directly related to the "windings" of the strings (or, in general, D-branes) on nontrivial cycles of the string compactification manifolds.

The relation between nonperturbative solutions of quantum theory and integrable systems has been studied in detail only for a few 2D topological theories and theories of quantum gravity, and also for the Seiberg–Witten solutions of $\mathcal{N}=2$ supersymmetric non-Abelian gauge theories: pure gluodynamics[41,68] including interactions with the ($N_a=1$) $\mathcal{N}=2$ matter hypermultiplet in the adjoint representation of the gauge group, described in terms of the family of Calogero–Moser integrable systems.[69–73] When the hypermultiplet mass becomes infinitely large, it is effectively decoupled from the Yang–Mills theory, and dimensional transmutation causes the elliptical Calogero–Moser model to degenerate into a periodic Toda chain describing pure 4D $\mathcal{N}=2$ supersymmetric gluodynamics.[43] It is also known[74,75] that $N_c=3$, $N_f=2$ (where N_c and N_f are the numbers of colors and flavors) curves correspond to the integrable Goryachev–Chaplygin top, while the natural hypothesis concerning families of models of $\mathcal{N}=2$ supersymmetric QCD[42,76] is formulation in terms of integrable (in general, inhomogeneous) $sl(2)$ spin chains, for which the Toda chain is again the limiting case. The idea of such an identification is based[74] on the special quadratic form of the algebraic equations describing the embedding of the spectral curve in $\mathbf{C}^2$ (Refs. 42 and 76). Another possibility,[78] that of remaining within the dynamics of the Toda chain while changing the boundary conditions, encounters difficulties when $N_f>N_c$.

Since the solutions are formulated in terms of the periods of meromorphic differentials on complex curves (one of the possible choices of coordinates on the moduli space), integrable systems [moreover, integrable systems of the KP or Toda type, by which we mean those in which the Liouville torus (angular variables) is the real cross section of a special complex torus—the Jacobian of a complex curve] arise almost by definition, owing to the Krichever construction.[50]

At the present time there is no mechanism which allows exact nonperturbative results to be obtained when starting directly from the first principles of string theory and quantum field theory [possibilities of obtaining some nonperturbative results by directly studying the Polyakov functional integral (1.1) will be one of the topics discussed in Sec. 2]. However, the striking analogy between the equations arising in the description of the deformation of finite-gap solutions of the Whitham equations—equations of the hydrodynamical type—and the renormalization-group method in standard perturbation theory[79] should be noted. Actually, the scale invariance of the correlation functions is related to their dependence on the coupling constant by a first-order equation: $[d/(d\log\Lambda)-\Sigma\beta_i(g)(\partial/\partial g_i)]F(g;\Lambda)=0$. Naively, there is no scale dependence in the exact solution, and it can be assumed that the derivatives with respect to Λ can be replaced by derivatives with respect to the moduli. However, even with this hypothesis it should be noted that there is no way of determining the β function outside the framework of perturbation theory. In addition, the associativity equations for the effective action suggest that some idea of the mechanism by which integrable equations arise must be sought in string theory.

The basic concepts used in this review. String perturbation theory (1.1) is defined by a path integral[12] over mappings $\mathbf{x}:\Sigma_g\to X$ of the world sheet Σ_g onto physical "space-time" X and over two-dimensional geometries, the metrics g_{ab} in (1.1) or their equivalence classes up to reparametrizations—moduli of the complex structure y on the space $\mathcal{M}_g$ in (1.2). In general, X need not be a four-dimensional Minkowski space or flat space $\mathbf{R}^4$; it may have a nontrivial metric (which, owing to the two-dimensional symmetries, satisfies the Einstein equations[67]) or even be a nontrivial compact manifold (more precisely, have a compact component), which physically corresponds to the origin of internal (gauge) degrees of freedom in the Kaluza–Klein scheme. Here the Polyakov integral (1.1) must be understood in a *generalized* sense, where instead of a free two-dimensional theory of the field $\mathbf{x}$ corresponding to flat space-time X, it is necessary to consider some general *two-dimensional conformal field theory*,[80] the fields of which, in general, implicitly correspond to a σ model on a nontrivial manifold. The anomaly of the measure for the integration over two-dimensional metrics Dg_{ab} can also be attributed to the gravitational contribution to the action of the two-dimensional conformal theory. When $c_{\text{matter}}+c_{\text{gravity}}-26=0$, two-dimensional gravity cancels the conformal anomaly, and the integral (1.1) reduces to (1.2), as before. It follows, in particular, that even in the absence of matter, (pure) two-dimensional gravity is, in general, a nontrivial string theory, and it is just such theories (for which the matter contribution is small, $c_{\text{matter}}<1$, compared with the contribution of gravity) which are amenable to nonperturbative analysis.

Two features of a conformal theory of a general type should particularly be noted. First, in string theory a single conformal theory may correspond to strings on *different* manifolds X_1 and X_2. Such manifolds are called mirror manifolds,[18] and the simplest example of this phenomenon is the theory of a free field taking values on a circle: the theories on $X_1=S_R$ and $X_2=S_{1/R}$ are equivalent. Second, despite the fact that conformal theories corresponding to nontrivial manifolds naively do not appear to be free, for any 2D conformal field theory there exists a free-field or bosonization technique, i.e., the string theories are basically defined perturbatively, and the integrals (1.1) and (1.2) can be calculated. Naturally, for a conformal theory of a general type this problem is very complicated technically, but there exist examples of conformal theories in which the integral over the matter fields, and even the resulting integral for the induced gravitation, can be fully calculated. Examples of such theories are free theories

$$S_{\rm CFT}=\int_{\Sigma_g}\bar{\partial}\mathbf{x}\partial\mathbf{x}\equiv\int_{\Sigma_g}\sum_{\mu=1}^{c_{\rm matter}}\bar{\partial}x^{\mu}\partial x^{\mu} \tag{1.7}$$

and certain theories with $c<0$. The presence of "negative" matter leads to additional cancellations in the integration measure (1.2), and this often makes it possible to calculate the Polyakov functional integral. The bosonization technique effectively reduces the calculations in nontrivial conformal theories to calculations (of fairly complex correlation functions) in theories of the form (1.7), or more precisely, theories with "extended" free action $S_{\rm CFT}(\varphi)=\int\bar{\partial}\varphi\partial\varphi+\alpha_0R\varphi$, where the constant α_0 (which in the case of many fields is a vector) is related to the central charge of the theory: $c_{\rm CFT}=1-12\alpha_0^2$. The best known examples of conformal theories of this type are the pq models[80] with central charge $c=1-6(p-q)^2/pq$. Because of this, the integral (1.1) can sometimes be calculated explicitly, and we shall discuss examples of such theories in the first section of this review. They all have $c_{\rm matter}<1$ (including negative values) or an integer, and the conclusions about the properties of the theory which are preserved in the nonperturbative regime (the correlation functions, for which there are no higher-order corrections, and the operator algebra) are based on the algebraic structure of the two-dimensional conformal theory (the Virasoro and Kac–Moody algebras[81,82]) and the representation of 2D conformal field theories by free fields.[83–87]

As expected, the information obtained directly from study of the integral (1.1) is very limited.[19,20,88,89] The formulation of the nonperturbative theory is rather implicitly related to the properties of the theory on the world sheet, and it can be described as follows. The central object is the generating function (1.4) of the exact physical correlators (scattering amplitudes), the calculation of which is the main problem of the theory. In the case of two-dimensional gravity and topological string models, the functional $\mathcal{F}$ in (1.4) is literally a function which, in general, is an infinite (although discrete in the present case) set of numerical variables, and the variational derivatives in (1.4) are transformed into ordinary partial derivatives.

The generating function depends on two types of variable. Those of the first type are sources for physical operators

$$F_g\to F_g(\mathbf{T})=\int Dg_{ab}D\mathbf{x}e^{-S_{\rm CFT}(\mathbf{x},g_{ab})+\sum T_k\mathcal{O}_k},$$

$$\mathcal{F}\to\mathcal{F}(\Lambda_{\rm str},\mathbf{T})=\sum_g\Lambda_{\rm str}^gF_g(\mathbf{T}), \tag{1.8}$$

the derivatives with respect to which determine the correlation functions in the theory. Equation (1.8) certainly depends on the choice of basis $\mathcal{O}_k$ or T_k, and only in a special basis (which is not necessarily convenient from the viewpoint of formulating the theory on the world sheet) can it be elegantly described in the language of nonlinear differential equations or relations of the unitarity type for the correlators.[25,26,28,32] In general, such relations are well known in traditional quantum field theory (the Ward identities,[30] the Schwinger–Dyson equations, and so on), but the situation in string theory is different because these equations can be written as a *complete system of integrable* differential equations completely determining the generating function (1.8). Like the function $\mathbf{T}$, the generating function (1.4), (1.8) can be defined only as a formal series whose coefficients are identified with correlation functions. The series, in general, has zero radius of convergence. This fact certainly reflects the well known properties of perturbation series in string theory and in quantum field theory, and, moreover, is consistent with the existing explicit expressions for exact nonperturbative solutions, which, if they exist at all, usually have an integral form and can sometimes be reduced to matrix integrals (1.3), i.e., to simple analogs of gauge field theory.

The other parameters on which the partition function or the generating function depends are the physical or space-time moduli of the theory. The space of these parameters is usually finite-dimensional, and complex in the cases considered, and it can often be interpreted as the moduli space of complex curves. It should be particularly noted that the complex curves or Riemann surfaces which arise here have a "space-time" origin (for example, they originate from string compactification) and are not at all related to the world sheets in string theory! Like a function of the moduli, the generating function is an ordinary (for example, meromorphic) function of many complex variables, and it can often be calculated more or less explicitly. The moduli themselves can be interpreted as the low-energy values of the background fields (the Higgs expectation values of scalars, the moduli of the physical metric—complex and Kähler structures, and so on), and, as a function of the moduli, $\mathcal{F}$ usually has the meaning of the effective action.

In topological 2D gravity and some topological string models (of the A_p series), the dependences on the moduli t and the sources T practically coincide (the $t+T$ formula[64]). The problem is to find the explicit form of the function $\mathcal{F}(t,\mathbf{T})$, or at least the equations which it satisfies. In the case of topological theories this problem can be solved explicitly, and the answer is expressed in terms of an integral of the form $\int DX\exp[-\mathrm{Tr}V(X)+\mathrm{Tr}\Lambda X]$, where the moduli t are related to the coefficients of the potential $V(X)$, and the external sources to the traces of powers of the matrix Λ. The proof of the $t+T$ formula is a nontrivial problem (see Sec. 3).

Naturally, this dependence is, in general, different, and both problems are of independent interest. In the case of an effective $\mathcal{N}=2$ supersymmetric gauge theory in four dimensions, as yet only the first question has an answer, and an extremely important fact is that the Wilson effective action in the massless sector, being a functional of the fields, can be expressed in terms of a *function* of several complex variables (see Refs. 41 and 42 and references therein). This effect is easily understood as follows.

For an $\mathcal{N}=2$ supersymmetric gauge theory with the group $SU(N_c)$, the scalar potential has the form $V(\phi)=\mathrm{Tr}[\phi,\phi^\dagger]^2$, and its minimum corresponds, apart from gauge transformations, to diagonal, traceless matrices $\phi=\mathrm{diag}(A_1,\dots,A_{N_c})$, whose invariants

$$\det(\lambda-\phi)=P_{N_c}(\lambda)=\sum_{k=0}^{N_c} S_{N_c-k}\lambda^k, \tag{1.9}$$

equal in number to the rank of the group, rank $SU(N_c)=N_c-1$ [or any other set of algebraically independent invariants, for example, $h_k=(1/k)\mathrm{Tr}\phi^k$], parametrize the space of physical moduli. The Higgs effect leads to the appearance of a mass for the off-diagonal part of the gauge field $\mathbf{A}_\mu$, since $[\phi,\mathbf{A}_\mu]=(A_i-A_j)\mathbf{A}_\mu^{ij}$, while the diagonal part remains massless, and the gauge group is broken from $G=SU(N_c)$ down to $U(1)^{\mathrm{rank}G}=U(1)^{N_c-1}$. Therefore, the massless sector can be represented as an $\mathcal{N}=2$ Abelian gauge theory whose effective Lagrangian is defined in terms of the superfields $\Phi_i=\varphi^i+\vartheta\sigma_{\mu\nu}\bar{\vartheta}G^i_{\mu\nu}+\dots$, whose vacuum values coincide with the diagonal elements of the matrix ϕ. Therefore, a function of complex variables $\mathcal{F}(a)=\mathcal{F}(A)|_{\Sigma A_i=0}$ actually determines the Wilson effective action of the massless fields, which is obtained from it by the substitution

$$\mathcal{L}_{\mathrm{eff}}\sim\int d^4\vartheta\mathcal{F}(A_i\to\Phi_i)=\dots\frac{\partial^2\mathcal{F}}{\partial a_i\partial a_j}G^i_{\mu\nu}G^j_{\mu\nu}+\dots. \tag{1.10}$$

As far as the massive excitations in $\mathcal{N}=2$ non-Abelian gauge theory are concerned, it turns out[41,42] that at least the spectrum of BPS states[2)] is related to the function $\mathcal{F}$ as $M\sim|\mathbf{na}+\mathbf{ma}_D|$, where $\mathbf{a}_D=\partial\mathcal{F}/\partial\mathbf{a}$.

Therefore, it is knowledge of the function of complex variables $\mathcal{F}$ as a function of the moduli and all its derivatives, for example, the expansion in sources $\mathbf{T}$, which gives the most complete nonperturbative information about the theory. It will be demonstrated below that, at least in one class of problems whose general features are described in this section, the main goal is to find a formulation of the properties of the generating function $\mathcal{F}$ which depends on some of its variables as on the moduli of the complex structure of some Riemann surface.

As already stated, the main idea is to identify the function $\mathcal{F}$, and also other characteristics of the physical theory, with quantities which are meaningful in systems of integrable equations of the KP/Toda type. To formulate this relationship, we also need to give some definitions from the theory of integrable models. In the class of problems studied in this review, all the equations which arise pertain to KP (1.5) or Toda (1.6) hierarchies, or, more precisely, to reductions of them. The concept of a hierarchy implies that the dynamical systems (1.5) and (1.6) possess an infinite number of integrals of the motion, which can be associated with an infinite number of mutually commuting [and commuting with the first equations in (1.5) and (1.6)] flows. The differential equations involving higher-order times T_k have a more complicated form if they are written as equations for the functions $U(\mathbf{T})$ and $\phi_n(\mathbf{T})$, but there is a more elegant method of specifying the entire hierarchy.

This method is based on the auxiliary linear problem for the hierarchy of integrable equations

$$\frac{\partial}{\partial T_k}\Psi=B_k\Psi, \tag{1.11}$$

where the $B_k=B_k[U;\phi]$ are differential operators *only* with respect to T_1 in the KP case (1.5) or difference operators with respect to the discrete time n in the Toda case (1.6), and the solution Ψ of the auxiliary linear problem is usually called the Baker–Akhiezer function. To (1.11) we can add the Lax equation

$$\mathcal{L}\Psi=\lambda\Psi, \tag{1.12}$$

which appears in the reductions as one of the equations of the chain (1.11). The hierarchy of nonlinear integrable equations here is equivalent to the Lax equations

$$\frac{\partial\mathcal{L}}{\partial T_k}=[B_k,\mathcal{L}] \tag{1.13}$$

or to the Zakharov–Shabat compatibility conditions

$$\left[\frac{\partial}{\partial T_k}-B_k,\frac{\partial}{\partial T_l}-B_l\right]=0. \tag{1.14}$$

The most universal object in this formulation of integrable problems is the Hirota τ function satisfying an infinite chain of bilinear differential (difference) equations and generating the solutions of the integrable hierarchy, the Baker–Akhiezer functions, and so on. For example, for the KP hierarchy,

$$\Psi=e^{\sum T_k\lambda^k}\frac{\tau\left(T_k-\dfrac{1}{k\lambda^k}\right)}{\tau(T)};\quad U(\mathbf{T})=\partial^2\log\tau(\mathbf{T});\quad\dots, \tag{1.15}$$

and similar expressions exist also for other hierarchies.

The Toda and KP hierarchies have an infinite number of solutions parametrized by a point of the infinite-dimensional Grassmannian[58,59] or, roughly speaking, a function of two variables. The particular solutions can be obtained by means of auxiliary conditions, which often have the form of additional (often linear) equations for the τ function.

Finite-dimensional reductions of hierarchies of integrable equations in which only a *finite* number of integrals of the motion and flows $\partial/\partial T_k$ are independent play a special role. A beautiful example of finite-dimensional reductions of hierarchies of KP/Toda equations is that of the so-called finite-gap solutions defined by the conditions

$$[\mathcal{L},\mathcal{A}]=0;\quad\mathcal{A}=\sum_k^{\mathrm{finite}}c_kB_k, \tag{1.16}$$

where $\mathcal{L}$ is the Lax operator (1.12), the B_k are the operators of the Baker evolution function (1.11), and c_k is some *finite* set of nonzero constants. The integration of finite-gap solutions is referred to as the Krichever construction[50] and reduces to the following steps:[3)]

• The combined spectrum of the commuting operators $\mathcal{L}$ and $\mathcal{A}$ (1.16) is specified by a system of equations describing a complex curve Σ; in the simplest case, $\mathcal{P}(\mathcal{L},\mathcal{A})=0$.

• A Baker–Akhiezer function is a section of some bundle over Σ. In the cases used below, this bundle is almost always linear.

• The moduli of a complex curve are integrals of the motion of the system (1.16).

• The integrating change of variable is the Abel transformation, and the Liouville torus (angular variables) is the real section of the Jacobian of the curve Σ.

• The Hamiltonian structure of a finite-gap solution is formulated by means of a generating meromorphic 1-differential dS, whose periods (integrals over nontrivial cycles on the Riemann surface) are the action variables (the canonical set of integrals of the motion) of the system.

The resulting complex curves are specified by algebraic equations of the form

$$\mathcal{P}(\lambda,w)=0 \tag{1.17}$$

[one relation of the form (1.17) for two variables, where $\mathcal{P}$ is a polynomial whose coefficients are the moduli of the complex structure, specifies a one-dimensional complex (or two-dimensional real) manifold], or by systems of equations in several complex variables. Topologically, each complex curve is characterized by a single parameter, the genus g (the number of handles), and for a surface of fixed genus Σ_g the complex structure is defined by $3g-3$ parameters—the moduli of the complex structure, i.e., $\dim_{\mathbf{C}}\mathcal{M}_g=3g-3$. Finite-gap integrable systems usually correspond to g-parameter families of complex curves (so that the dimension of the moduli space, equal to the number of independent integrals of the motion, coincides with the dimension of the Jacobian of the curve, i.e., with the number of angular variables). The dimension of the Jacobian coincides with the number of globally defined holomorphic differentials $d\omega_i$, $i=1,\dots,g$, and is equal to the genus of the surface. On a surface of genus g there are $2g$ independent incontractable cycles (two around each handle), the canonical set of which corresponds to division into the A_i, $i=1,\dots,g$, and B_i, $i=1,\dots,g$, cycles with intersection index $A_i\circ B_j=\delta_{ij}$. The holomorphic differentials are canonically chosen to be normalized on the **A** cycles, $\oint_{A_j}d\omega_i=\delta_{ij}$, and the integrals over the **B** cycles give the period matrix, $\oint_{B_j}d\omega_i=T_{ij}$. The finite-gap solutions are the simplest solutions of integrable systems, of those related to nonperturbative quantum theories. In general, they represent only the first approximation to the nonperturbative solutions, which are the main topic of the present review. They already allow some of the information on the physical characteristics of the effective theory to be described. In addition, in many cases the exact solutions can be viewed as integrable deformations of finite-gap solutions described by hierarchies of Whitham equations. When the solution of a string theory can be found exactly, it turns out to be a poorly defined solution of an integrable system (one which is not periodic and does not fall off, or, conversely, one which grows at infinity, and so on), but at the same time it corresponds to the minimal deformation of Eq. (1.16)—to the appearance of a constant on the right-hand side. Such an additional condition is called a string equation;[24] in extending it to the entire hierarchy, this condition takes the form of the condition for invariance of the τ function under the action of the linear differential operators forming the Borel $(n\geqslant -1)$ part of the Virasoro algebra.

The content of this review. We shall begin with the formulation of the topological phase of 2D gravity on the basis of the technique of two-dimensional conformal field theory and the theory of free fields on Riemann surfaces—world sheets in string theory. It is shown that the correlators in the theory of topological gravity have a representation in terms of a $c_{\rm CFT}=-2$ two-dimensional conformal theory with special matter interacting with ordinary Liouville gravity, and we shall study the generalization of two-dimensional gravity to the case of higher nonlinear two-dimensional symmetries corresponding to W algebras. We shall consider the geometrical formulation of 2D and W gravity, in which objects from the theory of integrable systems arise naturally. Studies of the algebra of the observables in 2D and W gravity, in particular, have shown that the open-string sector (or the holomorphic sector of the model) contains closed subalgebras whose structure is determined by the *model* (the sum of all unitary representations with unit weight) corresponding to a finite-dimensional group (the group G for W_G gravity).

Next, in Sec. 2 we turn directly to the exact nonperturbative formulation of quantum theories in terms of integrable systems and start by discussing the historically first example: the model of 2D quantum gravity. In Sec. 2.3 we show how the exact nonperturbative solution of models of 2D gravity of the $(2,2k+1)$ series (including pure gravity, the theory with $k=1$) is formulated in terms of the solution of the hierarchy of the Korteweg–de Vries (KdV) equation, invariant under the action of the $n\geqslant 1$ (Borel) subalgebra of the Virasoro algebra. In that section we encounter for the first time the logarithm of the τ function of an integrable hierarchy, which is the generating function for the correlators of nonperturbative string theory, and we also construct the nonperturbative effective action for the string equation allowing the possibility of interpolating between various critical points of the $(2,2k+1)$ series. In Sec. 2.4 we prove that there exists an *explicit* solution of the Virasoro conditions, i.e., a representation in the form of a matrix integral for the τ function, for the theory of *topological* gravity, whose generating function formally satisfies the same equations as the generating function of 2D quantum gravity [the $(2,2k+1)$ series].

Section 3 is devoted to the nonperturbative formulation of topological $c<1$ string models—theories in which not only the formulation of the nonperturbative regime in terms of systems of integrable equations is known, but also explicit solutions of these equations (in integral form) have been found which, in principle, permit one to obtain an exact result for the correlation function. In Sec. 3.1 we prove the integrability of topological string models of the $A_{(p-1)}$ series, namely, that the generating function is the logarithm of the τ function of the reduced KP hierarchy satisfying the *string equation* $\Sigma_k k T_k (\partial \mathcal{F}^{(p)}/\partial T_{k-p}) + \Sigma_{a+b=p} a T_a b T_b = 0$. In Sec. 3.2 we explicitly construct the solutions of topological string models and their topological deformations in the Ginzburg–Landau theory [i.e., determined by the polynomial (super)potential] in the form of matrix-integral solutions of the KP hierarchy. In Sec. 3.3 we study the generalization of the construction to the case of nontopological solutions. We show that the explicit solutions of topological theories are a consequence of a more general expression, the p,q-duality transformation $\Psi^{(p,q)}(\lambda) = \int d\mu(x) e^{S^{(p,q)}(\lambda,x)} \Psi^{(q,p)}(x)$, the formulation of which is the central topic of Sec. 3.3. Finally, in Sec. 3.4 we study the $c \to 1$ limit of nonperturbative topological solutions, and construct its generating function in terms of the τ function of the hierarchy of the two-dimensional Toda lattice. In addition, we propose an interpretation of our results from the viewpoint of string field theory.

Finally, in Sec. 4 we turn to the case of greatest physical interest: four-dimensional (4D) $\mathcal{N}=2$ supersymmetric gauge field theories, which are the point (or field) limit of the more complex $c>1$ string theories. In Sec. 4.1 we formulate the nonperturbative solution of $\mathcal{N}=2$ supersymmetric gluodynamics in terms of periodic solutions of the Toda chain. In particular, we show that the moduli of physical solutions (the parameters of the solutions of the classical equations of motion) in the $\mathcal{N}=2$ case can be identified with the integrals of the motion of the Toda chain, and that there exists an elegant formalism for this identification in terms of complex curves. In Sec. 4.2 we study the elliptic deformation of the Toda chain in the Calogero–Moser model corresponding to the inclusion of the interaction of $\mathcal{N}=2$ non-Abelian gauge theory with matter in the adjoint representation of the gauge group. In Sec. 4.3 we study an alternative deformation of the Toda chain into (classical) spin chains, and the solution of the corresponding periodic problem is identified with the nonperturbative formulation of $\mathcal{N}=2$ supersymmetric QCD.

Thus, in Sec. 4 the nonperturbative solutions of $\mathcal{N}=2$ supersymmetric non-Abelian gauge theory are formulated in terms of the finite-gap solutions of equations of the KP/Toda type. The relation between gauge theories and integrable systems will be studied in more detail in the second part of this review.[90]

2. EXACT RESULTS IN 2D GRAVITY AND STRING SOLUTIONS OF INTEGRABLE SYSTEMS

2.1. 2D and *W* gravity in the formalism of conformal theories and integrable systems of the KP type

As the starting point in studying quantum gravity, we shall assume that the theory, which by definition reduces to integration over all metrics (if we are located in the spontaneously broken phase), is at least naively a topological theory. Leaving aside for now the detailed discussion of the topological nature of an arbitrary theory of gravity, in this subsection we shall solve a much more modest problem, namely, we shall show that topological correlation functions in fact arise in the original Polyakov approach[12,19–21] to perturbative 2D quantum gravity.

We shall show that observables in the topological sector[32] of two-dimensional gravity are related to operators of *zero* dimension from the viewpoint of conformal theory, and the calculation of multiloop topological correlation functions is related to the problem of the bosonization of first-order bosonic systems—so-called $\beta\gamma$ systems.[83,86,87]

The results presented below are based on the following *postulates* of the conformal approach to two-dimensional gravity:[19–21]

• The choice of conformal structure in two dimensions is equivalent to choosing the complex structure: in the coordinates z, $\bar{z}$ on a surface of fixed genus g, the metric g_{ab} can be replaced by the Liouville field $\phi_{\mathcal{L}}$: $g_{ab} = e^{\phi_{\mathcal{L}}}(g_0)_{ab}$, and the ghosts bc—anticommuting fields of spin 2 and -1 and their complex conjugates. The action also depends on the reference metric $g_0 = g_0(y)$, which is a function of the moduli of the complex structure ($\{y\}$ are the coordinates on $\mathcal{M}_g$; Refs. 12 and 13).

For a system from a conformal theory of "matter," with action $S_{\mathrm{CFT}}\{\varphi, g\}$, central charge c, and 2D gravity, the Polyakov functional integral[12,19,20] (1.1) takes the form

$$F_g\{\mathcal{K}\} = \int_{\Sigma_g} D g_{ab} e^{-S_{\mathrm{gravity}}\{g_{ab}\}} \int D\varphi e^{-S_{\mathrm{CFT}}\{\varphi, g_{ab}\}} \tilde{\mathcal{K}}\{\varphi, g_{ab}\}$$
$$= A_g \int_{\mathcal{M}_p} dy \int D\phi_{\mathcal{L}} D\varphi e^{-25-c/48\pi \int_{d^2z} \left(\frac{1}{2}|\partial\phi_{\mathcal{L}}|^2 + R_0 \phi_{\mathcal{L}}\right)}$$
$$^{-S_{\mathrm{CFT}}\{\varphi, g_0\}} \int |Db\, Dc\, e^{\int b\bar{\partial}c}|^2 \mathcal{K}\{\varphi, b, c, \phi_{\mathcal{L}}\}, \qquad (2.1)$$

where $R_0 = \partial\bar{\partial} \log g_0$, $|\partial\phi_{\mathcal{L}}|^2 \equiv \partial\phi_{\mathcal{L}} \bar{\partial}\phi_{\mathcal{L}}$, and $\tilde{\mathcal{K}}\{\varphi, g_{ab}\}$ denotes a set of vertex operators in the theory.[4)] The coefficient $(25-c)/48\pi$ in front of the Liouville action on the right-hand side of (2.1) is determined by the condition that the conformal, holomorphic, and gravitational anomalies cancel.[12,13,19,20] The constant A_g depends on the choice of normalization of the functional integral (2.1) and, in general, is not fixed in Polyakov perturbation theory. The question of fixing this normalization (more precisely, the relative normalization of the F_g corresponding to different genera) is only solved by imposing nonlinear equations on the generating function and will be studied later on.

• The fundamental postulate is that $D\phi_{\mathcal{L}}$ is the measure for a *free* scalar field, i.e., it is determined by the norm $\|\delta\phi\|^2 = \int_{d^2z} |\delta\phi|^2$, and not by $\int_{d^2z} |\delta\phi|^2 e^{\phi}$. The main (but still not completely convincing) argument in favor of this postulate is the relation of 2D gravity to the Drinfeld–Sokolov reduction of the $SL(2,\mathbf{R})$ Wess–Zumino–Novikov–Witten (WZNW) model or the geometric quantization of the Virasoro algebra,[21,87] from which the $SL(2)$ invariant integration measure follows: $D\phi_{\mathcal{L}} \sim \Pi_z (dF/F')$ $\times(z) \sim (\det \partial)^{-1} \Pi_z d\phi_{\mathcal{L}}(z)$, $F'(z) = \partial F(z) = e^{\phi_{\mathcal{L}}(z)}$, while $\Pi_z e^{\phi_{\mathcal{L}}(z)} d\phi_{\mathcal{L}}(z)$ corresponds to $\Pi_z (dF/F'^2)$, which is *not* invariant under linear-fractional transformations $F \to (aF + b)/(cF + d)$. We shall apply this argument only to the "holomorphic square root" of $D\phi_{\mathcal{L}}$ and not to the measure itself.

• Not all the observables in the conformal theory on the right-hand side of (2.1) have the meaning of observables in quantum gravity. Only (1-dimensional) operators integrated over the surface or operators of zero dimension can appear as operators $\mathcal{K}$—they are independent of the location of the points on the world sheet, for example, $\int d^2z \mathcal{O}_\Delta\{b,c,\varphi\} e^{A_\Delta \phi_\mathcal{L}}$, where $\mathcal{O}_\Delta$ is an operator of dimension Δ from the matter sector, and A_Δ is chosen such that

$$\Delta_\mathcal{L}+\Delta=1, \tag{2.2}$$

where the dimension of the Liouville part $e^{A_\Delta \phi_\mathcal{L}}$ is determined by the energy–momentum tensor of the Liouville field:

$$T_\mathcal{L}=\frac{25-c}{12}\left(-\frac{1}{2}(\partial\phi_\mathcal{L})^2+\partial^2\phi_\mathcal{L}\right); \quad \Delta_\mathcal{L}=-\frac{6}{25-c}A_\Delta^2+A_\Delta \tag{2.3}$$

[we note that $\phi_\mathcal{L}(z,\bar z)\phi_\mathcal{L}(0,0)=-[12/(25-c)] \times \log z\bar z+\dots$]. Solving (2.2) for A_Δ, we find

$$A_\Delta=\frac{1}{12}[25-c-\sqrt{(25-c)(1-c+24\Delta)}], \tag{2.4}$$

where the sign in front of the square root is chosen such that $A_\Delta=0$ for $\Delta=1$.

Let us consider operators independent of the matter. The simplest local operators constructed from the ghosts bc have the form

$$s_n(z)=b\partial b\dots\partial^{n-2}b(z), \quad n>1;$$

$$s_n(z)=c\partial c\dots\partial^{|n|}c(z), \quad n<1;$$

$$\Delta(s_n)=2(n-1)+\frac{(n-2)(n-1)}{2}=\frac{n^2+n}{2}-1, \tag{2.5}$$

or, in the representation of a single bosonic field Φ,

$$|b|^2=e^{i\Phi}; \quad |c|^2=e^{-i\Phi};$$

$$T_{bc}=\frac{1}{2}(\partial\Phi)^2-i\left(j-\frac{1}{2}\right)\partial^2\Phi; \quad |s_n|^2=e^{i(n-1)\Phi}. \tag{2.6}$$

The situation becomes much simpler if the operators $\sigma_n = e^{B_n\phi_\mathcal{L}}|s_n|^2$ have *zero* dimension and $c=-2$. In fact, instead of (2.2) and (2.4) we then have

$$\Delta_\mathcal{L}+\Delta=0; \quad B_\Delta=\frac{1}{12}[25-c-\sqrt{(25-c)(25-c+24\Delta)}], \tag{2.7}$$

from which it follows that

$$B_n=\frac{1}{12}(25-c-\sqrt{(25-c)(1-c+12[n^2+n])})$$

$$=\frac{3}{2}(1-n). \tag{2.8}$$

The topological operator of zero dimension can be written as

$$\sigma_n=e^{B_n\phi_\mathcal{L}}|s_n|^2=e^{3/2(1-n)\phi_\mathcal{L}}|s_n|^2=e^{(1-n)(\Phi_\mathcal{L}-i\Phi)} \tag{2.9}$$

(where we have introduced the normalized Liouville field $\Phi_\mathcal{L}=\sqrt{(25-c)/12}\phi_\mathcal{L}=3/2\phi_\mathcal{L}$) and obviously has zero dimension. In the absence of the gravitational anomaly [the case of the theory (2.1)], the correlators of operators of zero dimension do not depend on the location on the surface and can play the role of (some of) the observables in quantum gravity.

The operators σ_n possess an important property: any set of them $\Pi_{i=1}^N \sigma_{n_i}(z_i,\bar z_i)$, satisfying the selection rule $\Sigma_{i=1}^N$ (No. of ghosts σ_{n_i}) $=\Sigma_{i=1}^N(n_i-1)=3g-3=\dim_\mathbb{C}\mathcal{M}_g$ in the ghost number, *automatically* satisfies the law of charge conservation with respect to the Liouville field:

$$\sum_{i=1}^N A_{n_i}=\sum_{i=1}^N \frac{3}{2}(1-n_i)=\frac{25-c}{6}(1-g)=\frac{9}{2}(1-g). \tag{2.10}$$

This remarkable coincidence occurs only for the particular value $c=-2$.

It is easily verified that tree ($g=0$) nonzero [i.e., satisfying the selection rule (2.10)] correlators of the operators σ_n are equal to a constant—the contribution of the ghosts exactly cancels that of the Liouville field. This is not so obvious for multiloop correlation functions: in order to ensure cancellation between the bc and Liouville contributions for higher genera $g>0$, $\phi_\mathcal{L}$ must not be treated as an "ordinary" scalar field with values in the field of real numbers. The problem is to cancel the instanton sector of the $\Phi-bc$ system, whose contribution is described by the Θ function in the expression

$$\left\langle \prod_i e^{i(n_i-1)\Phi(\xi_i)}\right\rangle \sim \prod_{i<j} E(\xi_i,\xi_j)^{(1-n_i)(1-n_j)}\Theta$$

$$\times\left(2\sqrt{2}\left(\sum_i (1-n_i)\xi_i+3\sqrt{2}\Delta\right)\Big|4T\right), \tag{2.11}$$

where $E(\xi_i,\xi_j)$ is the principal form on the surface of highest genus [the analog of $1/(\xi_i-\xi_j)$ on the sphere], and Θ denotes the Riemann theta function on the Jacobian (g-dimensional torus) of a surface of genus g (see, for example, Refs. 52, 91, and 92). It turns out to be much more natural to treat the field $\phi_\mathcal{L}$ as arising in the bosonization of the $\beta\gamma$ system, and having nontrivial global behavior for $p>0$ (see Refs. 86 and 87 for details). In contrast to the usual scalar field, the correlators of $\Phi_\mathcal{L}$ (treated as a field arising in the bosonization of the $\beta\gamma$ system) are calculated from the expression[86]

$$\left\langle \xi(z)\prod_i e^{(1-n_i)\Phi_\mathcal{L}(\xi_i)}\right\rangle \sim \prod_{i<j} E(\xi_i,\xi_j)^{-(1-n_i)(1-n_j)}$$

$$\times\left(\Theta\left(2\sqrt{2}\left(\sum_i (1-n_i)\xi_i +3\sqrt{2}\Delta\right)\Big|4T\right)\right)^{-1}, \tag{2.12}$$

which cancels the nontrivial factors in (2.11), apart from a constant.

In the foregoing discussion an ordinary ghost system can easily be replaced by an arbitrary system of anticommuting fields of spin j ($j\neq 2$), b_jc_{1-j} (following Ref. 83, we refer to the special case of $j=0$ as the $\eta\xi$ system, where $\eta=b$ is a spin-1 field), which corresponds to the very special case of matter with negative central charge $c=-2$. For the case of $c=-2$ conformal matter realized as the $\eta\xi$ system, the total action has the form

$$S_{\text{total}}=\frac{1}{4\pi}\int_{d^2z}\left(\frac{1}{2}|\partial\Phi_{\mathcal{L}}|^2+\frac{3}{2}R_0\Phi_{\mathcal{L}}+\eta\bar\partial\xi+\text{c.c.}+b\bar\partial c+\text{c.c.}\right)=\frac{1}{4\pi}\int_{d^2z}(\beta\bar\partial\gamma+b\bar\partial c+\text{c.c.}). \quad (2.13)$$

The terms in front of $b\bar\partial c$ give exactly the action arising in the bosonization of *commuting* $\beta\gamma$ systems with spin $j=2$ ($\beta=\partial\xi e^{\Phi_{\mathcal{L}}}$, $\gamma=\eta e^{-\Phi_{\mathcal{L}}}$). Thus, the full theory is transformed into a combination of bc and $\beta\gamma$ systems with the same spin $j=2$. It is quite natural to assume, and this is the main observation of the present subsection, that such a supersymmetric combination is a topological theory.

The measure in the functional integral for a similar $c=-2$ theory with 2D gravity can be rewritten as $Dg_{ab}D(c=-2\text{CFT})=(d\mu(y)DbDcD\phi)(D\xi D\eta)$, or

$$d\mu(y)(DbDc)(D\phi D\xi D\eta)=d\mu(y)(DbDc)(D\beta D\gamma). \quad (2.14)$$

Thus, the original functional integral is rewritten as an integral for the $bc-\beta\gamma$ system.

The proposed approach can be generalized directly to a theory with a richer symmetry on the world sheet—so-called W strings, where the internal geometry is formulated in terms of the W gravity associated with extended Virasoro algebras or W algebras.[93,94] W algebras are closely related to the theory of integrable systems[95] [W_N algebras with N reductions of the KP hierarchy: the KdV ($N=2$), Boussinesq ($N=3$), and other hierarchies]. The main goal is to obtain a nonperturbative formulation or a "summed" perturbation theory based on the technique of the *universal moduli space*—the infinite-dimensional Grassmannian[58,59] parametrizing different solutions of integrable systems of the KP or Toda type. The W geometry leads to the appearance of Virasoro conditions in physical space-time[25–27] (in general, literally, W conditions; henceforth, if not specially stipulated, by Virasoro conditions we shall mean the conditions for invariance under the action of the "Borel" part $W_{N,k\geq -N+1}$ of the generators of the W algebras themselves, which are literally generators of the Virasoro algebra only for $N=2$), arising from the W_∞ symmetries of the Grassmannian.[59,60]

Let us begin with the representation of the topological sector of W gravity in terms of free fields. Two-dimensional gravity is defined by the functional integral (2.1), and, as will be shown above, the simplest topological example corresponds to the special choice of $\eta\xi$ matter (2.13) in (2.1), for which the topological subsector (2.13) is formulated in terms of $j=2$ bc and $\beta\gamma$ systems:

$$S=S_{\text{graviry}}+S_{\text{matter}}=\int b\bar\partial c+\beta\bar\partial\gamma+\text{c.c.}, \quad (2.15)$$

where we have used the bosonization rules

$$\beta=e^{-\phi}\partial\xi,\quad \gamma=e^{\phi}\eta,$$

$$c_{\beta\gamma}=-c_{bc}=2(6j^2-6j+1)=26=c_{\text{matter}}+c_{\text{Liouville}}. \quad (2.16)$$

The generalization of these expressions to the case of W gravity is obvious. The topological action (2.15) has the form

$$S_{W-\text{gravity}}+S_{\text{matter}}=\int\sum_{j=2}^{N}(b_j\bar\partial c_{1-j}+\beta_j\bar\partial\gamma_{1-j}+\text{c.c.})=\int\sum_{j=2}^{N}\left(b_j\bar\partial c_{1-j}+\eta\bar\partial\xi+\text{c.c.}+\frac{1}{2}|\partial\phi_j|^2+\left(j-\frac{1}{2}\right)R_0(y)\phi_j\right). \quad (2.17)$$

First we note that the total central charge of the system of W ghosts

$$c_N=\sum_{j=2}^{N}-2(6j^2-6j+1)=2(1-N)(2N(N+1)+1) \quad (2.18)$$

(in particular, $c_2=-26$, $c_3=-100$, etc.) restricts the possible values of the central charge of the W algebra:

$$c_{W_N}=\sum_{j=2}^{N}\left\{1+12\left(j-\frac{1}{2}\right)^2\right\}=4(N-1)(N^2+N+1) \quad (2.19)$$

(for $N=2$ this equation (2.19) reproduces the "topological" Virasoro central charge $c=28$). The number of zero modes $\{b_j,c_{1-j}\}$ of the ghost fields is equal to the (complex) dimension of the W-moduli space. It follows from the Riemann–Roch theorem that

$$\sum_{j=2}^{N}(\text{No. of } b_j^{(0)})=(g-1)\sum_{j=2}^{N}(2j-1)=(g-1)(N^2-1)=(g-1)\dim SL(N) \quad (2.20)$$

[for $N=2$, $(g-1)\dim SL(2)=3g-3=\dim_{\mathbf{C}}\mathcal{M}_g$ coincides with the dimension of the moduli space of the complex structures]. Equation (2.20) was used by Hitchin for studying the moduli spaces of flat $SL(N,\mathbf{R})$ connections[96] and indicates the relation between W_G-moduli spaces and flat G connections on Riemann surfaces.

The most general expression for the functional integral in conformal W gravity has the form [cf. (2.1)]

$$\int_{\mathcal{M}_g}\{dy\}\int D\phi e^{-\int_{d^2z}(1/2|\partial\phi|^2+\beta_0R_0(y)\rho\phi+\Sigma_\alpha e^{\alpha\varphi})}\times\prod_{j=2}^{N}\int|Db_jDc_{1-j}e^{\int b_j\bar\partial c_{1-j}}|^2\times\int D\varphi e^{-S_{\text{matter}}\{\varphi,g_0\}}\mathcal{K}\{\varphi;\{b_j\},\{c_1-j\},\phi\}. \quad (2.21)$$

Let us now consider a very simple example allowing a better understanding of the features of the (classical) W symmetry. We deform the partition function of the theory with W symmetry

$$\langle\langle 1\rangle\rangle \equiv \langle e^{\int \mu_2 T+\mu_3 W_3+\dots}\rangle = \int e^{-S} e^{\int \mu_2 T+\mu_3 W_3+\dots}, \tag{2.22}$$

where $\mu_n d\bar{z}(dz)^{-n}$ are generalized Beltrami differentials, and we first consider the case in which only μ_2 and μ_3 are nonzero. Calculating the deformation in the first order, we have

$$\begin{aligned}\bar\partial u_2 \equiv \bar\partial\langle T(z)\rangle &\equiv \int d^2\xi\, \partial_{\bar z}\{\mu_2(\xi)\langle T(z)T(\xi)\rangle + \mu_3(\xi)\\ &\quad\times\langle T(z)W(\xi)\rangle\} + O(\delta_y^2)\\ &= -\frac{c}{12}\partial^3\mu_2 - 2u_2\partial\mu_2 - \mu_2\partial u_2\\ &\quad -3\partial\mu_3 u_3 - 2\mu_3\partial u_3 + \dots,\end{aligned} \tag{2.23}$$

where

$$u_2(z)=\langle T(z)\rangle,\quad u_3(z)=\langle W_3(z)\rangle,\dots,u_n(z)=\langle W_n(z)\rangle, \tag{2.24}$$

and c is the central charge. Similarly,

$$\begin{aligned}\bar\partial u_3 &= 3\partial\mu_2 u_3 + \mu_2\partial u_3 + \frac{c}{360}\partial^5\mu_3 + \frac{1}{3}\partial^3\mu_2 u_2\\ &\quad + \frac{1}{2}\partial^2\mu_3\partial u_2 + \partial\mu_3\left[2b^2\Lambda + \frac{3}{10}\partial^2 u_2\right]\\ &\quad + \mu_3\left[b^2\partial\Lambda + \frac{1}{15}\partial^3 u_2\right].\end{aligned} \tag{2.25}$$

For $\mu_3=0$, $\mu_2\equiv\mu$, and $u_2\equiv u$, (2.23) becomes

$$-\bar\partial u = 2\partial\mu u + \mu\partial u + \frac{c}{12}\partial^3\mu. \tag{2.26}$$

This equation can be treated as a compatibility condition for the auxiliary linear problem

$$\left(\frac{c}{6}\partial^2 + u\right)\Psi_{-1/2}=0;\quad \left(\bar\partial + \mu\partial - \frac{1}{2}\partial\mu\right)\Psi_{-1/2}=0, \tag{2.27}$$

where $c/6$ is a known semiclassical coefficient and $\Phi_{-1/2}$ denotes the $-\frac{1}{2}$ differential. The combatibility of these conditions implies that the complex and projective structures are consistent. Choosing $\mu=\bar\partial\epsilon$, it is easily seen that the last equation in (2.27) is a simple consequence of the transformation law for the $-\frac{1}{2}$ differential:

$$\delta\Psi_{-1/2} = \epsilon\partial\Psi_{-1/2} - \frac{1}{2}\partial\epsilon\Psi_{-1/2}. \tag{2.28}$$

For the W_3 Ward identity (2.25) the auxiliary linear problem has the form

$$\begin{gathered}\left(\frac{c}{24}\partial^3 + u_2\partial + \frac{1}{2}\partial u_2 + u_3\right)\Psi_{-1}=0,\\ \left(\bar\partial + \mu_2\partial - \partial\mu_2\partial - \frac{1}{6}\partial^2\mu_3 + \frac{1}{2}\partial\mu_3\partial\right.\\ \left. -\mu_3\left[\partial^2 - \frac{16}{c}u_2\right]\right)\Psi_{-1}=0\end{gathered} \tag{2.29}$$

[using the renormalization $\mu_3\to\sqrt{\frac{2}{5}}\mu_3$, $u_3\to\sqrt{\frac{5}{2}}u_3$; the coefficients in (2.29) are simplified for the special value $c=24$, which will be used below]. The second equation in (2.29) corresponds to the transformation law

$$\delta\Psi_{-1} = \epsilon_2\partial\Psi_{-1} - \epsilon_3\left(\partial^2 - \frac{2}{3}u_2\right)\Psi_{-1} + \epsilon\text{ - derivative terms}, \tag{2.30}$$

which *explicitly depends* on the "external field" u_2. This is the main difference between the cases of W_2 and the higher W_n. The general form of the transformations (2.30) is

$$\delta f = \sum \epsilon_n D_n(u_0,\dots,u_{n-1})f + \epsilon - \text{derivatives terms}, \tag{2.31}$$

where

$$D_n(u_0,\dots,u_{n-1}) = \partial^n + u_{n-1}\partial^{n-1} + \dots + u_0 \tag{2.32}$$

are nth-order differential operators. The appearance in this case of the operators (2.32) with nontrivial coefficients implies a relation to the KP hierarchy and the algebra of pseudodifferential operators:[97]

$$L_{KP} = \partial + \sum_{i=1}^{\infty} a_i\partial^{-i};\quad (L_{KP}^n)_+ = \partial^n + \dots + u_0. \tag{2.33}$$

In fact, the trivial differential operator ∂^n corresponds to the choice of the trivial point of the Grassmannian $\mathcal{W}_0=\{1,\lambda,\lambda^2,\lambda^3,\dots\}$, while the differential operator of general form (2.32) corresponds, in general, to any point of the Grassmannian. It is therefore natural to treat functions on the Riemann surface as objects depending on the Grassmannian point. For example, let us take some function on the curve and write it (locally) as a Fourier or Laplace integral $f(\xi)=\int e^{\lambda\xi}\hat f(\lambda)d\lambda$. Now this function can be raised to the section of some bundle over the Grassmannian specified by the flows of the KP hierarchy $f(t_1,\dots,t_n)=\int e^{\lambda t_1+\dots+\lambda^n t_n+\dots}\hat f(\lambda)d\lambda$, where $t_1\equiv\xi$. Then the symmetries corresponding to the action of differential operators of higher order in ξ are related to the action of the flows of the KP hierarchy $\partial_\xi^n f = \partial_{t_1}^n f = \partial_{t_n} f$ for the expression at the trivial point of the Grassmannian. At an arbitrary point of the Grassmannian, instead of this expression we have the transformation

$$f(t_1,\dots,t_n) = \int \Psi_{\mathcal{W}}(\lambda,\{t_n\})\hat f(\lambda)d\lambda, \tag{2.34}$$

where $\Phi_{\mathcal{W}}(\lambda,\{t_n\})$ is a Baker–Akhiezer function. This replacement gives rise to the nontrivial flows

$$\partial_{t_n} f = (\partial^n + u_{n-1}\partial^{n-1} + \dots + u_0)f, \tag{2.35}$$

as a consequence of changing over to the general point $\mathcal{W}$. In fact, a general point of the Grassmannian corresponds to W_∞ gravity.[59] The more frequently encountered case of *finite* W_N corresponds to special reductions of W_∞. Let us consider, for example, a special reduction of the Baker–Akhiezer function

$$\Psi_{\mathcal{W}}(\lambda,\{t_n\})e^{\Sigma\lambda^k t_k}\frac{\tau\left(t_1-\frac{1}{\lambda},\ldots,t_n-\frac{1}{n\lambda^n},\ldots\right)}{\tau(t_1,\ldots,t_n,\ldots)}$$

$$=e^{\Sigma t_n\lambda^n}\left[1+\sum_{i=1}^{\infty} w_i(\{t_n\})\lambda^{-i}\right] \qquad (2.36)$$

such that the sum on the right-hand side is *finite*, of order N. The corresponding τ function is expressed in terms of the solutions of the equation $(\partial^N+w_1\partial^{N-1}+\ldots+w_N)f_i=0$, $i=1,\ldots,N$ (Ref. 98), $\tau=\det_{ij}f_i^{(j-1)}$. It is easy to calculate the corresponding Ψ function:

$$\Psi(\lambda,t_1,\ldots,t_n,\ldots)=e^{\Sigma\lambda^k t_k}\frac{\tau\left(t_1-\frac{1}{\lambda},\ldots,t_n-\frac{1}{n\lambda^n},\ldots\right)}{\tau(t_1,\ldots t_n,\ldots)}$$

$$=e^{\Sigma\lambda^k t_k}\left(1+\frac{w_1}{\lambda}+\cdots+\frac{w_N}{\lambda^N}\right), \qquad (2.37)$$

where $w_{N-i}=\det_{kl}f_k^{(l)}|_{l\neq i}\cdot(\det_{kl}f_k^{(l-1)})^{-1}$. $GL(N)$ transformations of the functions f_i do not change the values of the Ψ function. Imposing the auxiliary condition $w_1=0$, we verify that $\partial\tau/\partial t_1=0$. There is a simple relation between the expressions for the w_n from (2.37) and the u_n from (2.32). Expanding the logarithms of both sides of (2.37) in series in $1/\lambda$ and using the equation $\partial_{t_1}\partial_{t_n}\log\tau=(L^n_{\mathrm{KP}})_{-1}$ (i.e., simply the coefficient of the term involving ∂^{-1}), we immediately find

$$w_2=\partial^2_{t_1}\log\tau+\partial_{t_2}\log\tau;$$

$$w_3=\partial^3_{t_1}\log\tau+\partial_{t_1}\partial_{t_2}\log\tau+\partial_{t_3}\log\tau;\ \ldots . \qquad (2.38)$$

Let us now consider the case where the only independent function is u_2 [i.e., the $\widehat{SL(2)}$ or KdV reduction of the KP hierarchy]. Then $\partial_{t_2}\log\tau=0$ and $w_2=u_2$. Similarly, for the $\widehat{SL(3)}$ reduction, $\partial_{t_3}\log\tau=0$ and $w_3=\frac{1}{2}\partial u_2+u_3$.

2.2. Algebras of the observables in 2D and *W* gravity

Let us now consider the *algebra of the observables* in $c=1$ theories of 2D gravity and theories of W gravity with integer central charge. These algebras can be treated as fundamental invariant characteristics of topological string models also *outside* the framework of perturbation theory, because they, in particular, are independent of the order of the term in the perturbative expansion. In the approach discussed here, the primary fields of the two-dimensional conformal theory—the (extended) Virasoro algebra, dressed by Liouville fields and ghosts—are representatives of BRST cohomology classes[38,100] or physical operators. From the viewpoint of the Lagrangian approach this is a remarkable simplification, because the set of primary fields, as a rule, possesses additional structures or symmetries which are not seen when studying all the fields of the two-dimensional conformal theory. In this subsection we shall show that an example of such a structure is essentially the $SU(2)$-group *model* found[101] at the $SU(2)$-invariant point of the $c=1$ conformal theory, i.e., the direct sum of all the unitary representations of $SU(2)$ of highest weight, in which each representation enters only once. The structure of the *model* is preserved upon dressing by the Liouville field and unifies a class of observables in the corresponding string model, or more precisely, in the open-string sector. Moreover, this structure also defines to some degree the algebra of the observables in the corresponding string theory. Below we shall propose a natural generalization of this structure to arbitrary groups G of the $A-D-E$ series, physically corresponding to $c=\mathrm{rank}\ G$ W_G gravity.

The construction of Ref. 101 proceeds as follows.

1. We consider a scalar (matter) field X compactified on a circle of radius $r=\sqrt{2}$ (i.e., $X\sim X+2\pi r=X+2\pi\sqrt{2}$) with energy–momentum tensor $-\frac{1}{2}(\partial X)^2$. For this radius the obvious $U(1)\times U(1)$ symmetry is increased to $SU(2)\times SU(2)$ symmetry, and we can study the chiral sector with $SU(2)$ symmetry. This corresponds to the self-dual point of the $c=1$ Gaussian model,[102] where $SU(2)\times SU(2)$ acts naturally on the Virasoro primary fields. Below, we shall consider the holomorphic sector in such a theory ($\equiv$ the open-string sector), in which the left- (or right-) handed $SU(2)$ group acts naturally. Here all the chiral primary vertex operators in the theory form the $SU(2)$ *model*: $M[SU(2)]$.

2. To isolate the structure of the primary fields, we can eliminate the descendants by introducing an interaction with 2D gravity, i.e., by going to string theory. Among the string observables there is a subsector consisting of integrals over surfaces (integrals over contours for the holomorphic sector) of gravitationally dressed primary fields with total unit dimension. The corresponding vertex operators have the form

$$q_{J,m}=\oint\psi_{J,m}(x)e^{(J-1)\sqrt{2}\phi}. \qquad (2.39)$$

3. These operators form a *Lie algebra* $\mathcal{G}$ (in contrast to the OPE in the conformal theory), whose structure constants are 3-point correlators on the sphere. Thus, we have obtained a mapping of $M[SU(2)]$ onto an algebra of (some of) the observables which is the Lie algebra $\mathcal{G}[SU(2)]$. Moreover, the mapping $M[SU(2)]\to\mathcal{G}[SU(2)]$ is a representation, i.e., it preserves the structure of the model:

(a) $q_{1,m}=Q_m$ forms the *adjoint* representation of $SU(2)$ in $\mathcal{G}[SU(2)]$;

(b) since the Q_m act trivially on the Liouville field ϕ, the $\{q_{J,m}\}$ form the $SU(2)$ *model*: the set of representations of highest weight of spin J.

4. Owing to the special selection rules determined by the properties of the Liouville sector, the commutator $[q_{J',m'},q_{J'',m''}]$ contains a single term $q_{J,m}$ with $J=J'+J''-1$ (and $m=m'+m''$):

$$[q_{J',m'},q_{J'',m''}]=C^{J'+J''-1}_{J',J''}\,q_{J'+J''-1,m'+m''}. \qquad (2.40)$$

The coefficients $C[SU(2)]$ are the $3j$ symbols (Clebsch–Gordan coefficients) of the $SU(2)$ group in some basis.

5. The structure constants specified by the $3j$ symbols $C_{J',J''}^{J'+J''-1}$ can also be interpreted as the structure constants of the algebra of area-preserving diffeomorphisms of the two-dimensional plane $\boldsymbol{R}^2 \sim \boldsymbol{C}$ (i.e., the algebra of Hamiltonian vector fields on the plane, often identified with the algebra W_∞).[5)]

Let us consider in more detail the theory of a single free scalar field X compactified on a circle, with the Lagrangian $\int \partial X \bar{\partial} X$. In this theory the chiral algebra usually contains $\hat{U}(1)\times\hat{U}(1)$, generated by the current $\mathcal{J}_0 = \partial X$, and the Virasoro algebra generated by $T = -\frac{1}{2}(\partial X)^2$ (the plus sign corresponds to the conjugates). The standard set of primary fields in this theory is given by the exponentials[6)] $e^{ipx}e^{i\bar{p}\bar{x}}$, where

$$p+\bar{p}=\frac{n}{R}, \quad p-\bar{p}=2mR, \tag{2.41}$$

$n,m \in \mathbf{Z}$, and R is the compactification radius[102] (we note that in fact this quantity is *half* the actual compactification radius: $x \sim x+2\pi R$ and $\bar{x} \sim \bar{x}+2\pi R$ implies that $x+\bar{x}=X \sim X+2\pi r$, where $r=2R$). However, in some cases the holomorphic chiral algebra increases to $S\hat{U}(2)_{k=1}$, formed by the currents $\mathcal{J}_\pm = e^{\pm i\sqrt{2}x}$ and $\mathcal{J}_0$. This occurs at the self-dual point for $R=1/\sqrt{2}$ [at which the theory is related to the $\hat{S}U(2)_{k=1}$ WZNW model]. Here the set of primary fields becomes $SU(2)\times SU(2)$-invariant, and for the holomorphic sector (or the sector corresponding to the theory of open strings) this implies that for each e^{ipx} the spectrum contains all nonzero $(Q_-)^k e^{ipx}$ [for $p>0$, or $(Q_+)^k e^{ipx}$ for $p<0$; $Q_\pm = \oint \mathcal{J}_\pm$ are the $SU(2)$ generators and commute with the energy–momentum tensor]. If $p=\text{integer}\times\sqrt{2}$, this sequence is finite, $k \leqslant |p|/\sqrt{2}$, and forms an $SU(2)$ representation of spin J $(J=|p|/2)$. Each representation appears once, and we obtain the $SU(2)$ *model.*

This conclusion is confirmed by the calculation of the one-loop partition function in the theory. In fact,[87,102,103] for $R=1/\sqrt{2}$ we have $(q=e^{2\pi i\tau})$

$$\mathcal{Z}(\tau,\bar{\tau})=\frac{\left|\theta\begin{bmatrix}0\\0\end{bmatrix}(2\tau)\right|^2+\left|\theta\begin{bmatrix}1/2\\0\end{bmatrix}(2\tau)\right|^2}{|\eta(q)|^2}, \tag{2.42}$$

where θ is the Jacobi theta function, which coincides with the partition function of the $SU(2)_{k=1}$ WZNW model. We introduce the quantity

$$Z(\tau)=\frac{\theta\begin{bmatrix}0\\0\end{bmatrix}(2\tau)}{\eta(q)}+\frac{\theta\begin{bmatrix}1/2\\0\end{bmatrix}(2\tau)}{\eta(q)}, \tag{2.43}$$

which can be interpreted as the "holomorphic square root" of the partition function or, what is practically the same thing, the partition function of the corresponding open-string model.[104] This holomorphic partition function can be represented as the sum of Virasoro characters over all the Virasoro primary fields from the spectrum of the theory. According to the foregoing arguments, (2.43) is exactly the partition function of the $SU(2)$ *model* or all the representations of $SU(2)$:

$$Z(\tau)=\sum_{J\in \mathbf{Z}_+/2}(2J+1)\chi_J(\tau), \tag{2.44}$$

where $\chi_J(\tau)$ denote the characters of the representations of the Virasoro algebra for $c=1$, $\chi_J(\tau)=(q^{J^2}-q^{(J+1)^2})/\eta(q)$ (Ref. 105), and the factor $(2J+1)=\dim R_J$ literally reflects the fact that we are dealing with the $SU(2)$ *model.* We therefore obtain

$$\begin{aligned}\eta(q)Z(\tau)&=\frac{1}{2}\left(\sum_{J\in\mathbf{Z}_+/2}(2J+1)\chi_J(\tau)+(J\to -J-1)\right)\\&=\sum_{n=-\infty}^{\infty}q^{(n/2)^2}=\theta\begin{pmatrix}0\\0\end{pmatrix}(\tau/2)\\&=\theta\begin{pmatrix}0\\0\end{pmatrix}(2\tau)+\theta\begin{pmatrix}1/2\\0\end{pmatrix}(2\tau),\end{aligned} \tag{2.45}$$

in accordance with (2.43).

Let us turn to the case of an arbitrary $A-D-E$ Lie algebra G, namely, we consider $r_G=\text{rank}(G)$ *free* scalar fields $X=\{X_1,\ldots,X_{r_G}\}$ with Lagrangian $\int \partial \boldsymbol{X}\bar{\partial}\boldsymbol{X}$. Usually the chiral algebra in an r_G-dimensional free theory is $\hat{U}(1)^{r_G}\times\hat{U}(1)^{r_G}$, and the generator of the Virasoro algebra is $T=-\frac{1}{2}\partial X\partial X$. In the case of many fields it is more natural to consider not the Virasoro subalgebra of the chiral algebra, but the algebra (of higher spins) W_G, formed by generators of the form $\Sigma(\boldsymbol{\mu}\partial \boldsymbol{X})^n$. For compactification on an r-dimensional lattice $\Gamma=\{\gamma\}: X\sim X+2\pi\gamma$, the primary Virasoro fields are $e^{ipx}e^{i\bar{p}\bar{x}}$, where

$$p=\gamma^*+\frac{1}{2}\gamma, \quad \bar{p}=\gamma^*-\frac{1}{2}\gamma; \quad \gamma\in\Gamma, \quad \gamma^*\in\Gamma^*, \tag{2.46}$$

and Γ^* is the lattice dual to Γ, i.e., $\boldsymbol{\gamma\gamma^*}$ is an integer. However, in the case of special lattices the chiral algebra increases to $\hat{G}_{k=1}$, with the generators $\mathcal{J}_{\boldsymbol{\alpha}}=e^{i\boldsymbol{\alpha}\boldsymbol{x}}$, $\mathcal{H}_{\boldsymbol{\nu}}=\boldsymbol{\nu}\partial \boldsymbol{x}$, where $\boldsymbol{\alpha}$ are all the roots of G, and $\boldsymbol{\nu}$ is some basis in the Cartan (hyper)plane. This occurs when Γ $(\gamma\in\Gamma)$ is the *root lattice* of the algebra G; then the charges $Q_{\boldsymbol{\alpha}}=\oint\mathcal{J}_{\boldsymbol{\alpha}}$ and $Q_{\boldsymbol{\nu}}=\oint\mathcal{H}_{\boldsymbol{\nu}}$, which are the generators of the algebra of the global symmetry G, commute with the energy–momentum tensor, and the Virasoro primary fields form a representation of G. As in the $SU(2)$ case, this implies that in addition to the "naive" primary fields (or tachyons) $e^{ipx}e^{i\bar{p}\bar{x}}$ there also exist others formed by the action of G. When $r_G>1$ this is not all, namely, there are many more *Virasoro* primary fields (gravitons, and so on—all the higher spins). In order to restrict the class of primary fields and reveal the structure of the *model* of the group G, it is necessary to change over to the primary fields of the W_G algebra. The generators of the W_G algebra are combinations of the type $\Sigma_{a=0}(\boldsymbol{\nu}_a\partial\mathbf{x})^n$, where $n=1,\ldots,r_G$ and $\boldsymbol{\nu}_a$ are vectors in the Cartan (hyper)plane, related to the fundamental weights. The W_G algebra itself (or its universal covering algebra) is defined as the part of the chiral algebra (in our case, the universal covering algebra $\hat{G}_1$) commuting with the charges Q_α, Q_ν. Therefore, as before, the primary fields of W_G form multiplets of G, and the complete set of them forms the *model* $M[G]$. In order to demonstrate this, we turn again to the equations for the one-loop partition functions:

$$\mathcal{Z}(\tau,\bar\tau)=|\eta(q)^{-r_G}|^2\sum_{\nu\in\Gamma^*/\Gamma}\sum_{\epsilon}\left|\Theta\begin{bmatrix}\nu+\epsilon\\0\end{bmatrix}(\tau)\right|^2, \quad (2.47)$$

where ϵ runs over the vectors $\{\frac{1}{2}e_i\}$ and $\mathbf{0}$ ($\{e_i\}$ is the basis of the lattice Γ).[87,103] The term with $\epsilon=0$,

$$\mathcal{Z}(\tau,\bar\tau)=|\eta(q)^{-r_G}|^2\sum_{\nu\in\Gamma^*/\Gamma}\left|\Theta\begin{bmatrix}\nu\\\mathbf{0}\end{bmatrix}(\tau)\right|^2, \quad (2.48)$$

is modular-invariant and is the one-loop partition function of the $k=1$ WZNW model for the $A-D-E$ group G. The corresponding one-loop partition function in the "chiral" or "open" sector is

$$Z(\tau)\equiv\eta(q)^{-r_G}\sum_{\nu\in\Gamma^*/\Gamma}\Theta\begin{bmatrix}\nu\\0\end{bmatrix}(\tau)=\sum_{\Lambda\in\Gamma_+^*}D_\Lambda\chi_\Lambda(\tau), \quad (2.49)$$

where the representation of highest weight $R_G[\Lambda]$ of the group G is in one-to-one correspondence with the highest vector Λ lying in the "positive" Weyl sector Γ^+. The dimension of the representation $R_G[\Lambda]$ is calculated as the product over all *positive* roots α (Ref. 106):

$$D_\Lambda=\prod_{\alpha\in\Delta_+}\frac{\langle\Lambda+\rho,\alpha\rangle}{\langle\rho,\alpha\rangle} \quad (2.50)$$

$[\rho=\frac{1}{2}\Sigma_{\alpha\in\Delta_+}\alpha$, and $\langle,\rangle$ is the scalar product in the Cartan (hyper)plane]. According to the foregoing arguments, the same vectors Λ correspond to the primary fields of the algebra W_G, or the irreducible representations $\mathcal{R}_{W_G}[\Lambda]$ with $c=r_G$. (However, we note that the *representations* $\mathcal{R}_{W_G}[\Lambda]$ and $R_G[\Lambda]$ themselves do not coincide: they are representations of different algebras!). We use $\chi_\Lambda(\tau)$ to denote the analogs of the Kac–Roch–Cardy Virasoro characters for the irreducible representations $\mathcal{R}_{W_G}[\Lambda]$ with conformal dimensions $\Delta_\Lambda=\frac{1}{2}\Lambda^2$:

$$\chi_\Lambda(\tau)=\eta(q)^{-r_G}\sum_{\sigma\in\mathcal{W}}\det(\sigma)q^{1/2(\Lambda+\rho-\sigma(\rho))^2}, \quad (2.51)$$

where $\mathcal{W}$ is the Weyl group of the algebra G, and $\det(\sigma)$ denotes the determinant of the transformation $\sigma\in\mathcal{W}$. These characters depend only on the dimension Δ, and so they are identical for all the primary fields D_Λ of the representation $R[\Lambda]$ and give identical contributions to (2.49), leading to the appearance of factors of D_Λ. Equation (2.49) proves that the W_G primary fields form a *model* of the group G.

To prove (2.49) we calculate the sum on the right-hand side. First we rewrite it as the sum of the weights over the entire lattice, using the fact that, owing to (2.50) and (2.51), for any $\sigma\in\mathcal{W}$ and ν we have

$$D_\nu\chi_\nu(\tau)=D_{\nu_\sigma}\chi_{\nu_\sigma}(\tau), \quad (2.52)$$

where $\nu_\sigma\equiv\sigma(\nu)+\sigma(\rho)-\rho$. For any lattice $\mathcal{T}$ this gives

$$\sum_{\nu\in\mathcal{T}_+}D_\nu\chi_\nu(\tau)=\frac{1}{\mathrm{ord}\mathcal{W}}\left(\sum_{\nu\in\hat{\mathcal{T}}}D_\nu\chi_\nu(\tau)\right), \quad (2.53)$$

where $\mathrm{ord}\mathcal{W}$ is the order (number of elements) of the Weyl group, $\mathcal{T}_+$ is the intersection of $\mathcal{T}$ with the Weyl sector, and $\hat{\mathcal{T}}$ is

$$\hat{\mathcal{T}}=\bigcup_{\sigma\in\mathcal{W}}[\sigma(\mathcal{T}_+)+\sigma(\rho)-\rho]. \quad (2.54)$$

In general, $\hat{\mathcal{T}}$ does not coincide with the initial lattice $\mathcal{T}$: this is true for the *root* lattice $\hat\Gamma=\Gamma$, but not for $\hat\Gamma^*\neq\Gamma^*$. [In the simplest $SU(2)$ example, $\Gamma^*=\{n/\sqrt{2},\ n\in\mathbf{Z}\}$, $\rho=1/\sqrt{2}$, $\Gamma_+^*=\{n/\sqrt{2},\ n\in\mathbf{Z},\ n\geqslant0\}$, and $\hat\Gamma_+^*=\{n/\sqrt{2},\ n\in\mathbf{Z},\ n\neq-1\}$. Nevertheless, the difference between Γ_+^* and $\hat\Gamma_+^*$ consists of the single point $\nu=-\rho$, and, according to (2.50), at this point $D_{-\rho}=0$, i.e., it does not contribute to the sum on the right-hand side of (2.53), so that the summation can run over the entire lattice Γ_+^*. The last statement is true for arbitrary $A-D-E$ algebras G; in general, the difference between Γ_+^* and $\hat\Gamma_+^*$ is no longer a point, but a hyperplane of codimension 1, so that for any $\nu\in\Gamma^*-\hat\Gamma^*$ the sum $\nu+\rho$ is orthogonal to at least one of the positive roots. Then, according to (2.50), the corresponding $D_\nu=0$ and the summation on the right-hand side of (2.53) can run over the entire lattice Γ^* instead of $\hat\Gamma^*$. After this we obtain

$$Z(t)=\sum_{\nu\in\Gamma_+^*}D_\nu\chi_\nu(t)=\frac{1}{\mathrm{ord}\mathcal{W}}\sum_{\nu\in\hat\Gamma^*}D_\nu\chi_\nu(\tau)$$
$$=\frac{1}{\mathrm{ord}\mathcal{W}}\sum_{\nu\in\Gamma^*}D_\nu\chi_\nu(t), \quad (2.55)$$

where we have also used the fact that $D_\nu=0$ for $\nu\in\Gamma^*-\hat\Gamma^*$. Substituting (2.55) and making the change of summation variable $\Lambda=\nu+\rho-s(\rho)$, we obtain

$$\eta(q)^{r_G}Z(\tau)=\frac{1}{\mathrm{ord}\mathcal{W}}\sum_{\nu\in\Gamma^*}\sum_{s\in\mathcal{W}}\prod_{\alpha\in\Delta_+}\frac{\langle\nu+\rho,\alpha\rangle}{\langle\rho,\alpha\rangle}$$
$$\times\det(s)q^{1/2[\nu+\rho-s(\rho)]^2}=\frac{1}{\mathrm{ord}\mathcal{W}}$$
$$\times\sum_{s\in\mathcal{W}}\det(s)\sum_{\Lambda\in\Gamma^*}\prod_{\alpha\in\Delta_+}\frac{\langle\Lambda+s(\rho),\alpha\rangle}{\langle\rho,\alpha\rangle}q^{1/2\Lambda^2}$$
$$=\sum_{\Lambda\in\Gamma^*}q^{1/2\Lambda^2}, \quad (2.56)$$

where we have used the fact that $s^2=1$ for $s\in\mathcal{W}$, and also

$$\sum_{\sigma\in\mathcal{W}}\det(\sigma)\frac{\langle\Lambda+\sigma(\rho),\alpha\rangle}{\langle\rho,\alpha\rangle}=\mathrm{ord}\mathcal{W} \quad (2.57)$$

for any Λ. Finally, for the right-hand side of (2.56) we obtain

$$\sum_{\nu\in\Gamma^*}q^{\nu^2/2}=\sum_{\nu\in\Gamma^*/\Gamma}\left(\sum_{\Lambda\in\Gamma}q^{(\Lambda+\nu)^2/2}\right)=\sum_{\nu\in\Gamma^*/\Gamma}\Theta\begin{bmatrix}\nu\\0\end{bmatrix}(\tau), \quad (2.58)$$

where a *lattice* Θ function defined as the sum over the lattice of *roots* Γ has appeared.

The characters of the irreducible representations of the algebra W_G (2.51) can be obtained by taking the limit $c \to r_G$ or $\alpha_0 \to 0$, namely, by using the characters of the "minimal" series arising for the values $c = r_G - 12\alpha_0^2\rho^2 = r_G - 6[(p-p')^2/pp']\rho^2$. According to Ref. 94,

$$\chi_{\Lambda_1,\Lambda_2}(\tau) = \eta(q)^{-r_G} \sum_{s_1,s_2 \in W} \frac{\det(s_1)\det(s_2)}{\mathrm{ord} W} \times \Theta\begin{bmatrix} ps_1\Lambda_1 - p's_2\Lambda_2 \\ 0 \end{bmatrix}(pp'\tau) = \eta(q)^{-r_G} \times \sum_{s \in W} \det(s) \sum_{\alpha \in \Gamma} \exp\left(\frac{i\pi\tau}{pp'}(ps\Lambda_1 - p'\Lambda_2 - 2pp'\alpha)^2\right), \tag{2.59}$$

of which only the term with $\alpha = 0$ remains in the limit $p \to \infty$, $p' \to \infty$ for fixed difference $p' - p$. Redefining $\Lambda_i \to \Lambda_i + \rho$ and setting $\Lambda_1 = 0$ [$n = 1$ in the $SU(2)$ case] and $\Lambda_2 \equiv \Lambda$, we arrive at (2.51).

Interaction with W_G gravity. Finally, in order to reveal the structure of the *model*, it is necessary to eliminate the descendants: to make the W_G symmetry a gauge symmetry or go to W_G strings. The main difference between W_G gravity and ordinary 2D gravity is the formulation, in which the physical operators (observables) are represented as integrated simple operators of unit dimension not containing ghost fields. Actually, ghost-free operators have the natural dimension $\Delta^{\{G\}} = 2\rho^2 = \frac{1}{6}C_V[G]\dim G$. Even the three-point correlation functions of such operators contain a nontrivial ghost part (i.e., the moduli space of W_G gravity is nontrivial even for a sphere with three labeled points).

The analog of the Liouville action is the action of the conformal G-Toda field theory (2.21):

$$\int_{d^2z}\left(|\partial\phi|^2 + \beta_0 R_0(y)\rho\phi + \sum_{\text{simple }\alpha} \eta_i e^{\alpha\phi}\right), \tag{2.60}$$

where the summation runs over the r_G *simple* roots of G and, as usual in the DDK formalism, we consider the point with all $\eta_i = 0$. In addition to the r_G-component W-Toda field ϕ, it is necessary to introduce r_G ghost pairs, bc systems $\int_{d^2z}\sum_{j\in S_G}(b_j\bar\partial c_{1-j} + \text{c.c.})$ with spins $j \in S_G$, where, in general, S_G is the set of G invariants or Casimir operators for the three A, D, and E series: for $SU(r+1)$, $j = 2, \ldots, r_G + 1$ [cf. (2.17)]; for $SO(2r)$, $j = 2, 4, \ldots, 2r-2$ and r; for E_6, $j = 2,5,6,8,9,12$; for E_7, $j = 2,6,8,10,12,14,18$; for E_8, $j = 2,8,12,14,18,20,24,30$, respectively. The central charge of the ghost system in general is $c_{\text{ghosts}} = \Sigma_{j\in S_G}[-2(6j^2 - 6j + 1)] = -48\rho^2 - 2r_G$, and the central charge of the W-Toda fields is $c_\phi = r_G + 48\beta_0^2\rho^2$. From the condition $c_{\text{matter}} + c_\phi + c_{\text{ghosts}} = 0$ we have $48(\beta_0^2 - 1)\rho^2 = c_{\text{matter}} - r_G$, and so for $c_{\text{matter}} = r_G$ we obtain $\beta_0 = \pm 1$.

For the construction of the observables in the W_G string model, the natural generalization of the DDK formalism leads to the following algorithm.

A. We choose any W_G primary matter field $\Psi_{\nu,\xi}(x) = \Pi_{i=1}^{r_G}(Q_{-\alpha_i})^{(\mu_i\xi)}\Psi_{\nu,0}$, $\Psi_{\nu,0} = e^{i\nu x}$, where ν corresponds to the representation $R_\nu[G]$ with the highest vector ν, and $\xi \equiv \xi_R$ is an element of this representation for which the conformal dimension $\Delta_{\nu,\xi} = \nu^2/2$ does not depend on ξ.

B. We dress the matter field by a W-Toda exponential $\Xi_{\nu,\xi}(x,\phi) = \Psi_{\nu,\xi}(x)e^{\beta_\nu\phi}$, such that the field $\Xi_{\nu,\xi}$ has fixed dimension $\Delta^{\{G\}}$. This gives the condition $\Delta_{\nu,\xi} - \frac{1}{2}\beta_\nu^2 - 2\beta_0\beta_{\nu\rho} = \Delta^{\{G\}}$, or, in our case, where $\Delta_{\nu,\xi} = \frac{1}{2}\nu^2$ and $\beta_0 = 1$,

$$\frac{1}{2}\nu^2 = \frac{1}{2}(\beta_\nu + 2\rho)^2 + (\Delta^{\{G\}} - 2\rho^2). \tag{2.61}$$

It is certain that the single (scalar) equation for r_G of the quantities (vector) β_ν (for fixed ν) has many solutions; nevertheless, there exists a special situation where $\Delta^{\{G\}} = 2\rho^2$, and

$$\beta_\nu = \nu - 2\rho. \tag{2.62}$$

C. Let us add the ghost factor in order to form an operator of zero dimension from $\Delta^{\{G\}}$. In ordinary gravity with $\Delta^{\{SU(2)\}} = 1$ it is sufficient to multiply $\Xi(x,\phi)$ by the ghost field corresponding to reparametrizations, $c_{-1} \equiv c$:

$$\mathcal{O}_{\nu,\xi}(x,\phi,c) = \Xi_{\nu,\xi}(x,\phi)c_{-1} = \psi_{\nu,\xi}(x)e^{\beta_\nu\phi}c_{-1}. \tag{2.63}$$

Here the correlation functions of observables are calculated with additional insertions of the form $\Pi_{\alpha=1}^{N^{(2)}}\int_{d^2z} b_2\mu_\alpha^{(2)}$, where the $\mu_\alpha^{(2)}$ are Beltrami differentials, and $N^{(2)}$ is the dimension of the moduli space. The alternative definition

$$\hat{\mathcal{O}}_{\nu,\xi}(x,\phi) = \int_{dz}\Xi_{\nu,\xi}(x,\phi) = \int_{dz}\psi_{\nu,\xi}(x)e^{\beta_\nu\phi} \tag{2.64}$$

is an integrated ghost-free operator of unit dimension.

For $G \neq SU(2)$ the situation is more complicated, since (at least at present) there is no natural definition of the form (2.64), and only a formulation of the BRST type analogous to (2.63) remains. Now instead of dressing $\Xi(x,\phi)$ by the single ghost field c_{-1}, we must use a combination of ghost fields:

$$\mathcal{O}_{\nu,\xi}(x,\phi,c) = \Xi_{\nu,\xi}(x,\phi) \times \prod_{j\in S_j}(c_{1-j}\partial c_{1-j}\partial^2 c_{1-j}\ldots\partial^{j-2}c_{1-j}) = \psi_{\nu,\xi}(x)e^{(\nu-2\rho)\phi}e^{i\Sigma(j-1)\varphi_j}. \tag{2.65}$$

In the last equation we have used the explicit expression (2.62) for β_ν and the bosonization of the ghost fields, i.e., $b_j = e^{-i\phi_j}$, $c_{1-j} = e^{i\varphi_j}$. The dimension of the combination $\{c_{1-j}\partial c_{1-j}\partial^2 c_{1-j}\ldots\partial^{j-2}c_{1-j}\} = :(c_{1-j})^{j-1}: = e^{i(j-1)\varphi_j}$ is $\Delta_j = -j(j-1)/2$, and the derivative of the ghost contributions in (2.65) acquires the dimension

$$\sum_{j\in S_G}\Delta_j = \frac{1}{24}\sum_{j\in S_G}[-2(6j^2 - 6j + 1) + 2] = \frac{1}{24}(c_{\text{ghosts}} + 2r_G) = -2\rho^2. \tag{2.66}$$

Thus, an operator corresponding to an observable has zero dimension, but acquires a large ghost charge. This ghost charge is canceled in the calculation of the correlators by insertions $\Pi_{j\in S_G}(\Pi_{\alpha=1}^{N^{(j)}}\int d^2z b_j \mu_\alpha^{(j)})$, which now include the Beltrami differentials corresponding to the moduli of the W structures. We note that the operators of observables (2.65) in W_G strings with $c=r_G$ can be chosen as representatives of BRST cohomology classes [at least for the case $G=SU(3)$; Ref. 107]. The operator algebra in the sector defined by the model of the group G reduces to the rules for the product of representations of the corresponding group, but here the calculation of the structure constants is difficult, owing to the absence or the awkwardness of the expressions for the Clebsch–Gordan coefficients in all cases except $SL(2)$.

2.3. Nonperturbative formulation of 2D quantum gravity: solution of the Virasoro conditions

By definition, the nonperturbative partition function (or the generating function for physical amplitudes) can be written as the sum of a series, each term of which is represented by a Polyakov functional integral over a Riemann surface of definite genus [see (2.1)]:

$$\mathcal{F}(\lambda)=\sum_{\text{genus}}\lambda^g F_g;\quad F_g=\int_{\Sigma_p} Dg\exp\gamma\int R\Delta^{-1}R. \tag{2.67}$$

Above, we showed that the functional integral (2.1), (2.67) can be calculated even perturbatively (i.e., each term separately) only for certain special cases in the simplest theories of 2D gravity. The calculation of the *nonperturbative* effects or summation of the series (2.67) is a complicated problem which has not yet been solved. The exact result can be found only by indirect methods, of which the first historically was the formulation in terms of matrix models,[22] where instead of the continuum theory (2.67) one deals with an effective discretization which in a sense is exact for the simplest string models, i.e., for special requirements on the space-time [for example, on the space-time dimension: effective matrix theories are known only for spaces of low dimension—in the limit for the case of pure gravity (2.67)].

As a rule, problems with the continuum formulation (2.67) are related to the fact that it contains excess information associated with the internal structure of the world sheet (for example, information about the structure of the representations of the chiral algebra of the 2D conformal field theory), which is not important for the formulation of a finite effective theory directly in physical space-time. In other words, the interaction with gravity converts the conformal descendants into gauge degrees of freedom, which do not carry any physics, and so it can be hoped that one can find an effective formulation which ignores the structure on the world sheet and which, if we are lucky, can be stated in terms of an integrable system. In a sense, integrability can be viewed as an auxiliary principle allowing the series (2.67) to be summed, namely, if a differential equation is found for which the (asymptotic) series (2.67) is a solution, then its exact solution will correspond to the nonperturbative regime.

Let us first consider an example in which the effective theory is given by the continuum limit of matrix models defined by integrals like (1.3), or more precisely, in the case of a single matrix taking the form

$$Z_N=\int DM_{N\times N}\exp-\mathrm{Tr}\sum t_k M^k;$$

$$DM_{N\times N}\equiv\frac{\Pi dM_{ij}}{\mathrm{Vol}U(N)}. \tag{2.68}$$

The continuum limit, in particular, the limit $N\to\infty$, $\log Z \underset{N\to\infty}{\to} \mathcal{F}$, gives the exact *nonperturbative* solution of (2.67) for the class of $(2,2k+1)$ models.[23]

The main difference between the continuous (2.67) and effective matrix formulations is the fact that the former assumes some auxiliary set of unitarity or factorization conditions in order to relate the normalization of different terms in the sum over topologies in (2.67) to each other, while in the matrix formulation (2.68) these relations appear *automatically*. Moreover, at least for known solutions they have the form of the Virasoro (in general, W) conditions,[7)] which in fact can be viewed as the *definition* of the nonperturbative theory. We shall refer to solutions of hierarchies of integrable equations satisfying Virasoro conditions as string solutions.

Below, we shall define nonperturbative theories as solutions of the Virasoro conditions. It turns out that these conditions lead directly to the *integrability* of the corresponding effective theories, in particular, solutions of the Virasoro conditions are the τ functions of well known hierarchies of integrable equations.[24–26]

In terms of the generating function (2.67), this means that $\mathcal{F}(T)=\log\tau(T)$, where $T\equiv\{T_k\}$ is the set of *times* of the integrable hierarchy or the set of coupling constants of the nonperturbative theory of 2D gravity. It is the appearance of an integrable system which is the new feature of the effective formulation allowing considerably more progress to be made by studying the properties of (2.68) rather than the original formulation (2.67).

The solution of the discrete Virasoro conditions[27]

$$L_n Z_N(t)=0,\quad n\geqslant -1;$$

$$L_n\equiv\sum_{k=0}^{\infty} k t_k\frac{\partial}{\partial t_{k+n}}+\sum_{a+b=n}\frac{\partial^2}{\partial t_a\partial t_b}, \tag{2.69}$$

with the auxiliary condition (making the variable t_0 meaningful) $\partial Z_N/\partial t_0=-NZ_N$, where N is identified with the dimension of the matrices in (2.68), in the special double scaling limit[23] gives a nonperturbative formulation of 2D quantum gravity as the solutions of differential equations:

$$\mathcal{L}_n\tau=0,\quad n\geqslant -1,$$

$$\mathcal{L}_n=\sum_{k=0}\left(k+\frac{1}{2}\right)T_{2k+1}\frac{\partial}{\partial T_{2(k+n)+1}}$$
$$+G\sum_{0\leq k\leq n-1}\frac{\partial^2}{\partial T_{2k+1}\partial T_{2(n-k)-1}}+\frac{\delta_{0,n}}{16}+\frac{\delta_{-1,n}T_1^2}{16G}, \tag{2.70}$$

where τ is the τ function of the KdV hierarchy, i.e., in addition to (2.70) it also satisfies an infinite system of nonlinear differential equations (bilinear Hirota relations; see, for example, Ref. 56). The final formulation of this family of models of 2D gravity in terms of integrable systems is based on the following statements:

• The generating function of the matrix model (2.69) as a function of time is the τ function of a semi-infinite Toda chain. The corresponding Riemann–Hilbert problem is the scalar product

$$\langle A(\lambda),B(\lambda)\rangle=\oint A(\lambda)B^*(\lambda)e^{-V(\lambda)};$$
$$V(\lambda)\equiv\sum_{k\geq0}t_k\lambda^k. \tag{2.71}$$

This scalar product allows the introduction of a set of orthogonal polynomials $P_n(\lambda)=\lambda^n+O(\lambda^{n-1})$, differing in normalization from the standard Baker–Akhiezer functions $\Psi_n(\lambda)=P_n(\lambda)\exp[-V(\lambda)/2-\phi_n/2]$,

$$\langle P_n(\lambda),P_m(\lambda)\rangle=e^{\phi_n}\delta_{mn}, \tag{2.72}$$

so that

$$Z_N(t)=\tau_N(t)=\tau_0\prod_{n=1}^{N-1}e^{\phi_n}, \tag{2.73}$$

and the variables ϕ_n as functions of the times $\{t_k\}$ satisfy the equations of the Toda-chain hierarchy:

$$\frac{\partial^2\phi_n}{\partial t_1^2}=e^{\phi_{n+1}-\phi_n}-e^{\phi_n-\phi_{n-1}}\equiv R_{n+1}-R_n,$$
$$\frac{\partial\phi_n}{\partial t_2}=-\left(e^{\phi_{n+1}-\phi_n}+\left(\frac{\partial\phi_n}{\partial t_1}\right)^2-e^{\phi_n-\phi_{n-1}}\right)$$
$$\equiv-(R_{n+1}+p_n^2-R_n), \tag{2.74}$$

and so on. In the reduced model the system (2.74) degenerates into a Volterra hierarchy, the first equation of which (in the variables $R_n\{t_{2k}\}\equiv e^{\phi_n-\phi_{n-1}\{t_{2k}\}}$) has the form $\partial R_n/\partial t_2=-R_n(R_{n+1}-R_{n-1})$. The compatibility between the Toda equations and the Virasoro conditions (2.69) gives the string equation.[27]

• The continuum limit is defined as the *double scaling* limit,[23] for which $N\to\infty$ simultaneously with the condition that the coupling constants attain their critical values, with the string coupling constant (the parameter describing the genus expansion) fixed. Continuous quantities are obtained from discrete ones by renormalization, which is a consequence of the nontrivial replacement of the times $\{\mathbf{t}\}\to\{\tilde{\mathbf{T}}\}\to\{\mathbf{T}\}$ and renormalization of the generating function. More precisely, below we shall introduce a parameter a such that $a\to0$ in the continuum limit, and all discrete quantities are functions of the parameter a, for example, the discrete times $t_k\equiv t_k(a,\mathbf{T})$, the matrix dimension $N\equiv N(a,\mathbf{T})\to\infty$ as $a\to0$, etc. This limit is nontrivial, both for the Toda equations[8)] and for the Virasoro conditions. The simplest case in which it can be determined and leads to a family of $(2,2k+1)$ models of 2D gravity is the one-matrix Hermitian model with zero odd-numbered times.[28] The relation between the discrete and continuous theories in the language of free scalar fields corresponding to bosonization of the Virasoro conditions (2.69) and (2.70) is obtained by replacement of the spectral parameter $u^2=1+az$. Here the continuous generators of the Virasoro algebra (2.70) are the modes of the energy–momentum tensor:

$$T(z)=\frac{1}{2}:\partial\Phi^2(z):-\frac{1}{16z^2}=\sum\frac{\mathcal{L}_n}{z^{n+2}}. \tag{2.75}$$

Details of the procedure of taking the continuum limit can be found in Refs. 28 and 29.

• The compatibility condition for the Toda equations and the Virasoro conditions, i.e., the discrete string equation

$$n+\frac{1}{2}=G_n^{(k)}\{R\}\ \text{ or }\ 1=G_{n+1}^{(k)}\{R\}-G_{n-1}^{(k)}\{R\}, \tag{2.76}$$

is equivalent to the extremum condition $\delta S/(\delta\log R_n)=0$ for the functional

$$S=\sum_n\left(\phi_n+\sum_k t_kG_n^{(k)}\{R\}\right). \tag{2.77}$$

The action (2.77), written in terms of the Lax operator L with the standard normalization of the matrix elements $L_{mn}\equiv\langle m|\lambda|n\rangle/\sqrt{\langle m|m\rangle\langle n|n\rangle}$ (Ref. 27), takes the form

$$S=\sum_n\left(\phi_n+\frac{1}{2}\sum_k t_k\mathrm{Tr}L^{2k}\right). \tag{2.78}$$

Thus, we have explicitly presented a construction in which the family of solutions of the integrable system (which is a reduction of the KP hierarchy or the two-dimensional Toda lattice) selected by the condition of invariance of the corresponding τ functions under the action of some of the generators of the Virasoro algebra (the W algebra) $L_n\tau=0$, $n\geq-1$, is formulated in Lagrangian language, i.e., as equations of motion $\delta S=0$. As usual, the action allows us to go beyond the equations of motion: a functional integral of the form $\int\mathcal{D}\phi\exp[-(1/\hbar)S\{\phi\}]$ for nonzero Planck constant $\hbar\neq0$ in principle allows us to study the dynamics in the configuration space of the string field theory. In particular, it can be hoped that by changing the values of the parameters t in (2.78) by means of the renormalization group, it will become possible to describe transitions between different multicritical points or between different points of the effective theory of two-dimensional gravity. In this formulation these deformations of the action $S\{\phi\}$ are specified by derivatives with respect to the times T, i.e., by (mutually commuting) flows of integrable hierarchies.

2.4. Topological 2D gravity as an explicit solution of the Virasoro conditions

In this subsection we shall construct a solution of the continuous Virasoro conditions without reference to their discrete analogs, i.e., we shall propose a procedure completely different from the one given above for constructing the solution of nonperturbative two-dimensional quantum gravity as the limit of an auxiliary discrete problem with simpler "unitarity conditions." In contrast to the discrete Virasoro conditions, where the solution is found directly in the form of a conformal correlator of ordinary scalar fields,[108] the continuum case turns out to be much more complicated. The main reason is that the continuum case differs from the discrete one [in explicit form; cf. (2.69) and (2.70)] in that an ordinary scalar field is replaced by a scalar field with *antiperiodic* boundary conditions. It is this singular transformation which corresponds to the meaning of the double scaling limit. It is much more complicated to construct a conformal solution for fields with antiperiodic boundary conditions,[9)] and therefore we shall use a different method to solve the continuum problem.

It turns out that the continuous Virasoro conditions (in special Miwa variables, which will be discussed in detail below) can be "summed" to give definite *matrix* differential operators. Specifically, for the $W^{(p)}$ algebra (Virasoro $=W^{(2)}$) these operators are related to the Laplacians (or Casimir operators) of the corresponding finite-dimensional algebras $[SL(n)$ for $W^{(n)}]$ and have the form

$$\frac{\partial^p}{\partial\Lambda^p}+\ldots, \tag{2.79}$$

where Λ is an $N\times N$ Hermitian matrix [for the case of $SL(N)$]. The conditions for invariance under the operators (2.79) can be interpreted as the Ward identities in several effective matrix theories.

We shall show that from the (system of) equations for the invariance of the function under the action of an operator like (2.79),

$$\mathrm{Tr}\epsilon(\Lambda)\left(W\left(\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\right)-\Lambda\right)\mathcal{Z}[\Lambda]=0 \tag{2.80}$$

[where $W(X)$ is a polynomial, and $(\Lambda_{\mathrm{tr}})_{ij}\equiv\Lambda_{ji}$], it will follow (strictly speaking, in the $N\to\infty$ limit) that this function is a solution of the continuous Virasoro (or W) conditions. It follows from the form of (2.80) that it can be rewritten as a Ward identity for a matrix integral, which (for a definite normalization) gives the exact nonperturbative solution of 2D (topological) gravity. The exact equation for the corresponding generating function has the form[34] $[W(M)\equiv V'(M)]$

$$Z^{(N)}[V|M]\equiv C^{(N)}[V|M]e^{\mathrm{Tr}V(M)-\mathrm{Tr}MW(M)}\times\int DXe^{-\mathrm{Tr}V(X)+\mathrm{Tr}W(M)X}, \tag{2.81}$$

where the integration runs over the space of $N\times N$ Hermitian matrices, and the normalization factor can be written as a Gaussian integral:

$$C^{(N)}[V|M]^{-1}\equiv\int DYe^{-\mathrm{Tr}\,U_2[M,Y]},$$

$$U_2\equiv\lim_{\epsilon\to0}\frac{1}{\epsilon^2}\mathrm{Tr}[V(M+\epsilon Y)-V(M)-\epsilon YV'(M)]. \tag{2.82}$$

First we shall discuss only special potentials of the monomial type, $V_p(X)=X^{p+1}/(p+1)$, leading to equations like (2.79) after substitution into (2.80).[10)]

In the simplest case $p=2$ we have a quadratic differential operator (Laplacian), and therefore we must prove that

$$\frac{1}{\mathcal{Z}}\mathrm{Tr}\left(\epsilon\frac{\partial^2}{\partial\Lambda_{\mathrm{tr}}^2}-\epsilon\Lambda\right)\mathcal{Z}=\frac{1}{Z}\sum_{n\geqslant-1}\mathcal{L}_nZ\mathrm{Tr}(\epsilon\Lambda^{-n-2}) \tag{2.83}$$

for

$$\mathcal{Z}^{\{2\}}\{\Lambda\}\equiv\int DX\exp\left(-\frac{1}{3}\mathrm{Tr}X^3+\mathrm{Tr}\Lambda X\right)$$

$$=C[\sqrt{\Lambda}]\exp\left(\frac{2}{3}\mathrm{Tr}\Lambda^{3/2}\right)Z^{\{2\}}(T_m),$$

$$T_m=\frac{1}{m}\mathrm{Tr}M^{-m}=\frac{1}{m}\mathrm{Tr}\Lambda^{-m/2},\quad m\ \text{odd}, \tag{2.84}$$

with

$$C[\sqrt{\Lambda}]=\det(\sqrt{\Lambda}\otimes I+I\otimes\sqrt{\Lambda})^{-1/2} \tag{2.85}$$

and

$$\mathcal{L}_n=\frac{1}{2}\sum_{\substack{k>\delta_{n+1,0}\\ k\ \text{odd}}}kT_k\frac{\partial}{\partial T_{k+2n}}+\frac{1}{4}\sum_{\substack{a+b=2n\\ a,b>0;a,b\ \text{odd}}}\frac{\partial^2}{\partial T_a\partial T_b}$$

$$+\delta_{n+1,0}\cdot\frac{T_1^2}{4}+\delta_{n,0}\cdot\frac{1}{16}-\frac{\partial}{\partial T_{2n+3}}. \tag{2.86}$$

Equation (2.83) is true for *any* dimension of the matrix Λ;[11)] moreover, in the limit of infinite dimension of the matrix Λ $(N\to\infty)$ all quantities $\mathrm{Tr}(\epsilon\Lambda^{-n-2})$, for example, $\mathrm{Tr}\Lambda^{p-n-2}$ for $\epsilon=\Lambda^p$, become algebraically independent, and so from (2.83) it follows that $\mathcal{L}_nZ\{T\}=0$, $n\geqslant-1$. We note that the function $\mathcal{Z}\{\Lambda\}$, which must be differentiated in (2.84) in order to prove that the Virasoro identities (2.83) are satisfied, depends only on the eigenvalues $\{\lambda_k\}$ of the matrix Λ. Therefore, it is natural to study (2.83) at the "diagonal point" $\Lambda_{ij}=0$, where $i\neq j$. The only nondiagonal piece in (2.83) surviving after diagonalization is proportional to

$$\left.\frac{\partial^2\lambda_k}{\partial\Lambda_{ij}\partial\Lambda_{ji}}\right|_{\Lambda_{mn}=0,\ m\neq n}=\frac{\delta_{ki}-\delta_{kj}}{\lambda_i-\lambda_j}\quad\text{for } i\neq j. \tag{2.87}$$

This equation (2.87) is just the second-order correction, well known from any quantum-mechanics course, to the eigenvalue of the Hamiltonian in traditional quantum-mechanical perturbation theory. This expression can easily be derived by variation of the determinant:

$$\delta\log(\det\Lambda)=\mathrm{Tr}\frac{1}{\Lambda}\delta\Lambda-\frac{1}{2}\mathrm{Tr}\left(\frac{1}{\Lambda}\delta\Lambda\frac{1}{\Lambda}\delta\Lambda\right)+\ldots. \tag{2.88}$$

For diagonal $\Lambda_{ij}=\lambda_i\delta_{ij}$ but, in general, nondiagonal $\delta\Lambda_{ij}$, Eq. (2.88) gives

$$\sum_k \frac{\delta\lambda_k}{\lambda_k}=-\frac{1}{2}\sum_{i\neq j}\frac{\delta\Lambda_{ij}\delta\Lambda_{ji}}{\lambda_i\lambda_j}$$
$$=\frac{1}{2}\sum_{i\neq j}\left(\frac{1}{\lambda_i}-\frac{1}{\lambda_j}\right)\frac{\delta\Lambda_{ij}\delta\Lambda_{ji}}{\lambda_i-\lambda_j}+\dots, \tag{2.89}$$

which proves (2.87).

Since the *matrix* ϵ is arbitrary (and therefore may be a function of Λ), it can be chosen to depend only on the eigenvalues λ_i. Thus, we are actually using only two conditions:

(i) the specific form of the normalization factor (2.82);

(ii) the fact that the generating function $Z[T(\lambda_i)]$ is a *complex* function, i.e., it must be differentiated as though it depended on the eigenvalues $\{\lambda_i\}$ only through the variables T_k.

After these conditions are satisfied, (2.83) can be rewritten as

$$\frac{e^{-2/3\mathrm{Tr}\Lambda^{3/2}}}{C(\sqrt{\Lambda})Z\{T\}}\left[\mathrm{Tr}\epsilon\left\{\frac{\partial^2}{\partial\Lambda^2}-\Lambda\right\}\right]C(\sqrt{\Lambda})e^{2/3\mathrm{Tr}\Lambda^{3/2}}Z\{T\}$$
$$=\frac{1}{Z}\sum_{a,b>0}\frac{\partial^2 Z}{\partial T_a\partial T_b}\sum_i \epsilon(\lambda_i)\frac{\partial T_a}{\partial\lambda_i}\cdot\frac{\partial T_b}{\partial\lambda_i}$$
$$+\frac{1}{Z}\sum_{n\geqslant 0}\frac{\partial Z}{\partial T_n}\left[\sum_{i,j}\epsilon(\lambda_i)\frac{\partial^2 T_n}{\partial\Lambda_{ij}\partial\Lambda_{ji}}\right.$$
$$+2\sum_i \epsilon(\lambda_i)\frac{\partial T_n}{\partial\lambda_i}\frac{\partial\log C}{\partial\lambda_i}$$
$$\left.+2\sum_i \epsilon(\lambda_i)\frac{\partial T_n}{\partial\lambda_i}\left(\frac{2}{3}\right)\frac{\partial}{\partial\lambda_i}\mathrm{Tr}\Lambda^{3/2}\right]+\left[\sum_i \epsilon(\lambda_i)\right.$$
$$\times\left(\frac{\partial}{\partial\lambda_i}\left(\frac{2}{3}\right)\mathrm{Tr}\Lambda^{3/2}\right)^2-\sum_i\lambda_i\epsilon(\lambda_i)+\sum_{i,j}\epsilon(\lambda_i)$$
$$\times\left(\frac{\partial^2}{\partial\Lambda_{ij}\partial\Lambda_{ji}}\left(\frac{2}{3}\right)\mathrm{Tr}\Lambda^{3/2}\right)+2\sum_i\epsilon(\lambda_i)$$
$$\left.\times\left(\frac{2}{3}\right)\frac{\partial\mathrm{Tr}\Lambda^{3/2}}{\partial\lambda_i}\frac{\partial\log C}{\partial\lambda_i}+\frac{1}{C}\sum_{i,j}\epsilon(\lambda_i)\frac{\partial^2 C}{\partial\Lambda_{ij}\partial\Lambda_{ji}}\right], \tag{2.90}$$

where $\mathrm{Tr}\Lambda^{3/2}=\Sigma_k\lambda_k^{3/2}$ and $C=\Pi_{i,j}(\sqrt{\lambda_i}+\sqrt{\lambda_j})^{-1/2}$. Now the calculation of all the quantities in (2.90) reduces to an exercise in differentiation, using (2.87). Explicit calculation shows that after differentiation, the resulting terms contain only *negative* powers of $\sqrt{\lambda_i}$ and can be "absorbed," i.e., rewritten in terms of the times T_k. As a result, we find

$$\frac{e^{-2/3\mathrm{Tr}\Lambda^{3/2}}}{C(\sqrt{\Lambda})Z\{T\}}\left[\mathrm{Tr}\epsilon\left\{\frac{\partial^2}{\partial\Lambda^2}-\Lambda\right\}\right]C(\sqrt{\Lambda})e^{2/3\mathrm{Tr}\Lambda^{3/2}}Z\{T\}$$
$$=\frac{1}{Z}\sum_{n\geqslant -1}\mathrm{Tr}(\epsilon_p\Lambda^{-n-2})\left(\frac{1}{2}\sum_{k>\delta_{m+1,0}}kT_k\frac{\partial}{\partial T_{2n+k}}\right.$$
$$+\frac{1}{4}\sum_{\substack{a+b=2n\\ a>0,b>0}}\frac{\partial^2}{\partial T_a\partial T_b}+\frac{1}{16}\delta_{n,0}+\frac{1}{4}\delta_{n+1,0}T_1^2$$
$$\left.-\frac{\partial}{\partial T_{2n+3}}\right)Z(T)=0, \tag{2.91}$$

namely, a tower of continuous Virasoro conditions for the case $p=2$.

The derivation for an arbitrary value of p is completely analogous and involves the following steps:

• We write $\mathcal{Z}[\Lambda]$ as $\mathcal{Z}^{\{p\}}[\Lambda]=g_p[\Lambda]Z^{\{p\}}(T_n)$, where

$$g_p[\Lambda]=\frac{\Delta(M)}{\Delta(\Lambda)}\prod_i[V''(\mu_i)^{-1/2}e^{(\mu_i V'(\mu_i)-V(\mu_i))}]$$
$$=\frac{\Delta(\Lambda^{1/p})}{\Delta(\Lambda)}\prod_i[\lambda_i^{-p-1/2p}e^{p/p+1\lambda_i^{i+1/p}}], \tag{2.92}$$

i.e., we explicitly separate the normalization factor from the function of time.

• We substitute the expression for $\mathcal{Z}^{\{p\}}[\Lambda]$ into (2.80), which in the case of the monomial potential $V_p(X)=X^{p+1}/(p+1)$ has the form

$$\left\{\mathrm{Tr}\epsilon(\Lambda)\left[\left(\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\right)^p-\Lambda\right]\right\}g_p[\Lambda]Z^{\{p\}}(T_n)=0, \tag{2.93}$$

The higher derivatives $\partial^i Z/\partial\Lambda_{\mathrm{tr}}^i$ are calculated by means of relations like (2.87).

• We make the variable shift

$$T_n\to\hat{T}_n=T_n-\frac{p}{n}\delta_{n,p+1} \tag{2.94}$$

(this procedure does not change the derivatives themselves).

• After all the substitutions, the left-hand side of (2.93) takes the form of an infinite series in which each term is a product of $\mathrm{Tr}[\bar{\epsilon}(M)M^{-k}]$ and a linear combination of generators of the W_p algebra acting on $Z^{\{p\}}(T_n)$. For example, if $p=3$, the resulting equation has the form

$$\frac{1}{27}\mathrm{Tr}\left[\bar{\epsilon}(M)M^{-3}\left(\sum M^{-3n}\mathcal{W}_{3n}^{(3)}+9\sum M^{-3n-1/3}\left(\sum(3k\right.\right.\right.$$
$$\left.-2)\hat{T}_{3k-2}\mathcal{W}_{3n+3k}^{(2)}+\sum\frac{\partial}{\partial T_{3a+1}}\mathcal{W}_{3b-3}^{(2)}\right)$$
$$+9\sum M^{-3n-2/3}\left(\sum(3k-2)\hat{T}_{3k-2}\mathcal{W}_{3n+3k}^{(2)}\right.$$
$$\left.\left.\left.+\sum\frac{\partial}{\partial T_{3a+1}}\mathcal{W}_{3b-3}^{(2)}\right)\right)\right]Z^{\{3\}}=0. \tag{2.95}$$

• In the limit $N\to\infty$ all expressions $\mathrm{Tr}\bar{\epsilon}(M)M^{-k}$ with fixed k and arbitrary $\bar{\epsilon}(M)$ become independent, and Eq.

(2.93) gives a tower of invariance conditions with respect to the action of the W generators. The exact proof for the $p=3$ case by the method proposed in this subsection has been given by Mikhaĭlov.[109] In Sec. 3 below, we shall propose a different existence proof for the Virasoro conditions, using the integrability of topological string theories existing for any p.

Finally, let us discuss the meaning of the shift (2.94). In the preceding subsection we studied a complicated procedure for obtaining the exact nonperturbative solutions of 2D quantum gravity as the solutions of the continuous Virasoro conditions, and we did not obtain explicit representations for these solutions. In the present subsection we have proved that the continuous Virasoro conditions have explicit solutions, which at least have an explicit integral representation. For the $p=2$ case this representation gives a solution of pure *topological* gravity and is called the Kontsevich model.[40] Thus, we have proved that the generating functions of two-dimensional quantum and topological gravity satisfy the same Virasoro conditions and in this sense are *equivalent.* Nevertheless, a more detailed treatment shows that the perturbative expansions for topological and quantum gravity occur at completely different points (see below), and the time shift (2.94) corresponds only to topological gravity.

3. EXACT SOLUTIONS OF TOPOLOGICAL STRING MODELS

3.1. Integrability of topological string models

In this subsection we shall show that the Virasoro conditions (2.70), (2.86), and (2.95) considered above and determining the nonperturbative string solutions specify a completely specific solution of the integrable KP hierarchy, namely:

• The partition function $Z_N^V[M]$ (2.81) as a function of the times[57]

$$T_k=\frac{1}{k}\operatorname{Tr}M^{-k},\quad k\geqslant 1, \tag{3.1}$$

is, for any N, the τ function of the KP hierarchy for *any* potential $V[X]$.

• If the potential $V[X]$ is a homogeneous polynomial of degree $p+1$, the partition function $Z_N^{\{V\}}[M]=Z_N^{\{p\}}[M]$ in fact is the τ function of the p-reduced KP hierarchy or, equivalently, the pth KdV hierarchy (Ref. 59).[12)]

To prove this, we first rewrite (2.81) as the determinant expression

$$Z_N^{\{V\}}[M]=\frac{\det_{(ij)}\phi_i(\mu_j)}{\Delta(\mu)},\qquad i,j=1,...,N, \tag{3.2}$$

and then show that this form is, in a sense, the definition of *any* τ function of the KP hierarchy, written in Miwa variables.[13)]

The principal feature distinguishing string solutions of hierarchies of integrable equations of the KP type is the special form of the functions $\{\phi_i(\mu)\}$ in (3.2), which is not at all arbitrary. Moreover, in the case considered here, the entire *infinite* set of functions in (3.2) is expressed in terms of a single function, the potential $V[X]$ [i.e., instead of an arbitrary *matrix* A_{ij} determining $\phi_i(\mu)=\Sigma A_{ij}\mu^j$ in the general case, the solutions corresponding to the nonperturbative regime in (topological) string theories are parametrized by the *vector* V_i or the function $V[\mu]=\Sigma V_i\mu^i$]. This is due to the presence of additional $\mathcal{L}_{-1}$ and other $\mathcal{W}$ conditions, which in the context of integrable hierarchies of the KP type can be viewed as a consequence of the $\mathcal{L}_{-1}$. All these conditions, in particular, follow from the Ward identities (2.80).

For the proof we first reduce the representation as a matrix integral

$$\mathcal{Z}_N^{\{V\}}[\Lambda]\equiv\int DXe^{-\operatorname{Tr}[V(X)-\operatorname{Tr}\Lambda X]}, \tag{3.3}$$

where it is easy to integrate over the "angular" $U(N)$ matrices,[110] to an N-fold integral over the eigenvalues of the matrices X and Λ (denoted by $\{x_i\}$ and $\{\lambda_i\}$, respectively). The integral (3.3) takes the form

$$\frac{1}{\Delta(\Lambda)}\left(\prod_{i=1}^{N}\int dx_ie^{-V(x_i)+\lambda_ix_i}\right)\Delta(X), \tag{3.4}$$

where $\Delta(X)$ and $\Delta(\Lambda)$ are Vandermonde determinants, for example, $\Delta(X)=\Pi_{i>j}(x_i-x_j)$.

Now (3.4) can be rewritten as

$$\Delta^{-1}(\Lambda)\Delta\left(\frac{\partial}{\partial\Lambda}\right)\prod_i\int dx_ie^{-V(x_i)+\lambda_ix_i}$$
$$=\Delta^{-1}(\Lambda)\det_{(ij)}F_i(\lambda_j) \tag{3.5}$$

with the matrix elements

$$F_{i+1}(\lambda)\equiv\int dxx^ie^{-V(x)+\lambda x}=\left(\frac{\partial}{\partial\lambda}\right)^iF_1(\lambda). \tag{3.6}$$

We note that $F_1(\lambda)=\mathcal{Z}_{N=1}^{\{V\}}[\lambda]$. Recalling that $\Lambda=V'(M)$ and going over to the eigenvalues of the matrix M, $\{\mu_i\}$, we find

$$\mathcal{Z}_N^{\{V\}}[V'(M)]=\frac{\det\Phi_i(\mu_j)}{\Pi_{i>j}(V'(\mu_i)-V'(\mu_j))}, \tag{3.7}$$

where

$$\Phi_i(\mu)=F_i(V'(\mu)). \tag{3.8}$$

We now change to the normalization (2.82) given by the Gaussian integral

$$C^{(N)}[V|M]^{-1}\equiv\int DXe^{-U_2(M,X)}. \tag{3.9}$$

Using the $U(N)$ invariance of the Haar measure dX, we can easily diagonalize M. Then the Gaussian integral (3.9) is easily calculated:

$$\int DXe^{-\Sigma_{i,j}^NU_{ij}X_{ij}X_{ji}}\sim\prod_{i,j}^{N}U_{ij}^{-1/2}, \tag{3.10}$$

and it remains only to substitute the specific form of $U_{ij}(M)$. If the potential is written as a formal series $V(X)=\Sigma(v_n/n)X^n$, then

$$U_2(M,X)=\sum_{n=0}^{\infty} v_{n+1}\left(\sum_{a+b=n-1} \mathrm{Tr} M^a X M^b X\right),$$

$$U_{ij}=\sum_{n=0}^{\infty} v_{n+1}\left(\sum_{a+b=n-1} \mu_i^a \mu_j^b\right)=\sum_{n=0}^{\infty} v_{n+1}\frac{\mu_i^n-\mu_j^n}{\mu_i-\mu_j}$$
$$=\frac{V'(\mu_i)-V'(\mu_j)}{\mu_i-\mu_j}. \tag{3.11}$$

Returning to (2.81), we obtain

$$Z_N^{\{V\}}[M]=e^{\mathrm{Tr}[V(M)-MV'(M)]}C^{(N)}[V|M]\mathcal{Z}_N[V'(M)]$$
$$\sim[\det\tilde{\Phi}_i(\mu_j)]\prod_{i>j}^{N}\frac{U_{ij}}{(V'(\mu_i)-V'(\mu_j))}$$
$$\times\prod_{i=1} s(\mu_i)=\frac{[\det\tilde{\Phi}_i(\mu_j)]}{\Delta(M)}\prod_{i=1}^{N} s(\mu_i), \tag{3.12}$$

$$s(\mu)=[V''(\mu)]^{1/2}e^{V(\mu)-\mu V'(\mu)}. \tag{3.13}$$

The product of the s factors on the right-hand side of (3.12) can be absorbed in the definition of the $\tilde{\Phi}$ functions:

$$Z_N^{\{V\}}[M]=\frac{\det\Phi_i(\mu_j)}{\Delta(M)}, \tag{3.14}$$

where

$$\Phi_i(\mu)=s(\mu)\tilde{\Phi}_i(\mu)\xrightarrow[\mu\to\infty]{}\mu^{i-1}\left(1+\mathcal{O}\left(\frac{1}{\mu}\right)\right), \tag{15}$$

and the asymptotic behavior is important in order that the determinant in (3.14) give a solution of the KP hierarchy in the sense of Refs. 58 and 59.

It follows from (3.8), (3.13), and (3.15) that the $\Phi_i(\mu)$ can be obtained from the basic function $\Phi_1(\mu)$ by means of the relation

$$\Phi_i(\mu)=[V''(\mu)]^{1/2}\int x^{i-1}e^{-V(x)+xV'(\mu)}dx$$
$$=A_{\{V\}}^{i-1}(\mu)\Phi_1(\mu), \tag{3.16}$$

where $A_{\{V\}}(\mu)$ is the first-order differential operator

$$A_{\{V\}}(\mu)=s\frac{\partial}{\partial\lambda}s^{-1}=\frac{e^{V(\mu)-\mu V'(m)}}{[V''(\mu)]^{1/2}}\frac{\partial}{\partial\mu}\frac{e^{-V(\mu)+\mu V'(\mu)}}{[V''(\mu)]^{1/2}}$$
$$=\frac{1}{V''(\mu)}\frac{\partial}{\partial\mu}+\mu-\frac{V'''(\mu)}{2[V''(\mu)]^2}. \tag{3.17}$$

In the special case $V(x)=x^{p+1}/(p+1)$, the operator $A_{\{p\}}(\mu)=(1/p\mu^{p-1})(\partial/\partial\mu)+\mu-(p-1)/2p\mu^p$ coincides [apart from a scale transformation of μ and $A_{\{p\}}(\mu)$] with the operator selecting a finite-dimensional subspace in the infinite-dimensional Grassmannian.[63] It should be specially noted that it is the relation $\Phi_{i+1}(\mu)=A_{\{V\}}(\mu)\Phi_i(\mu)$ $[F_{i+1}(\lambda)=(\partial/\partial\lambda)F_i(\lambda)]$ which is responsible for isolating the partition function of topological (W) gravity—the GKM among the solutions (and τ functions) of general form written in Miwa variables:

$$\tau_N^{\{\phi_i\}}[M]=\frac{[\det\phi_i(\mu_j)]}{\Delta(M)} \tag{3.18}$$

with an arbitrary set of functions $\phi_i(\mu)$. It will be shown below that (3.18) is exactly the τ function of the KP hierarchy in the Miwa representation.

The best-known representation of the (general) τ function of the KP hierarchy is that in the form of the fermionic correlator $\tau^G\{T_n\}=\langle 0|:e^{\Sigma T_n J_n}:G|0\rangle$,[56] where

$$J(z)=\tilde{\psi}(z)\psi(z),\quad G=:\exp\mathcal{G}_{mn}\tilde{\psi}_m\psi_n: \tag{3.19}$$

in the (two-dimensional) theory of free fermionic fields $\psi(z)$, $\tilde{\psi}(z)$ with (holomorphic) action $\int\tilde{\psi}\bar{\partial}\psi$. The vacuum states are determined by the conditions $\psi_n|0\rangle=0$, $n<0$, and $\tilde{\psi}_n|0\rangle=0$, $n\geqslant 0$, where $\psi(z)=\Sigma_Z\psi_n z^n dz^{1/2}$, $\tilde{\psi}(z)=\Sigma_Z\tilde{\psi}_n z^{-n-1}dz^{1/2}$. An important constraint on the form of the correlator with the insertions (3.19) is the fact that the operator $:e^{\Sigma T_n J_n}:G$ is Gaussian, and its insertion can be viewed as a modification of the quadratic action and the fermionic propagator $\langle\tilde{\psi}\psi\rangle$, so that, as before, the Wick theorem is applicable, namely, the correlators $\langle 0|\Pi_i\tilde{\psi}(\mu_i)\psi(\lambda_i)G|0\rangle$ for *any* suitable operator G are expressed in terms of pair correlators:

$$\langle 0|\prod_i\tilde{\psi}(\mu_i)\psi(\lambda_i)G|0\rangle=\det_{(ij)}\langle 0|\tilde{\psi}(\mu_i)\psi(\lambda_j)G|0\rangle. \tag{3.20}$$

To understand what happens to the operator $e^{\Sigma T_n J_n}$ after the Miwa transformation (3.1), it is simplest to go over to the representation of the current $J(z)=\partial\varphi(z)$ by free bosonic (scalar) fields. Then

$$\sum T_n J_n=\sum_i\left(\sum_n\frac{1}{n\cdot\mu_i^n}\varphi_n\right)=\sum_i\varphi(\mu_i)$$

and

$$:e^{\Sigma_i\varphi(\mu_i)}:=\frac{1}{\Pi_{i<j}(\mu_i-\mu_j)}\prod_i :e^{\varphi(\mu_i)}:. \tag{3.21}$$

In the fermionic representation it is better to start with

$$T_n=\frac{1}{n}\sum_i\left(\frac{1}{\mu_i^n}-\frac{1}{\tilde{\mu}_i^n}\right) \tag{3.22}$$

instead of (3.1). Then

$$:e^{\Sigma T_n J_n}:=\frac{\Pi_{i,j}^N(\tilde{\mu}_i-\mu_j)}{\Pi_{i>j}(\mu_i-\mu_j)\Pi_{i>j}(\tilde{\mu}_i-\tilde{\mu}_j)}\prod_i\tilde{\psi}(\tilde{\mu}_i)\psi(\mu_i), \tag{3.23}$$

and to reconstruct the form of the replacement (3.1) it is necessary to take the limit in which all the $\tilde{\mu}_i$ go to infinity. In other words, this implies that the left-hand vacuum is replaced by

$$\langle N|\sim\langle 0|\tilde{\psi}(\infty)\tilde{\psi}'(\infty)\ldots\tilde{\iota}^{(N-1)}(\infty). \tag{3.24}$$

Now the τ function can be represented as

$$\tau_N^G[M]=\langle 0|:e^{\Sigma T_n J_n}:G|0\rangle=\Delta(M)^{-1}\langle N|\prod_i :e^{\varphi(\mu_i)}:G|0\rangle$$

$$=\lim_{\tilde{\mu}_j\to\infty}\frac{\Pi_{i,j}(\tilde{\mu}_i-\mu_j)}{\Pi_{i>j}(\mu_i-\mu_j)\Pi_{i>j}(\tilde{\mu}_i-\tilde{\mu}_j)}$$

$$\times\langle 0|\prod_i \tilde{\psi}(\tilde{\mu}_i)\psi(\mu_i)G|0\rangle, \tag{3.25}$$

where, applying the Wick theorem (3.20) and taking the limit $\tilde{\mu}_i\to\infty$, we obtain

$$\tau_N^G[M]=\frac{\det\phi_i(\mu_j)}{\Delta(M)} \tag{3.26}$$

in which the matrix elements are the functions

$$\phi_i(\mu)\sim\langle 0|\tilde{\psi}^{(i-1)}(\infty)\psi(\mu)G|0\rangle \xrightarrow[\mu\to\infty]{} \mu^{i-1}$$

$$\times\left(1+\mathcal{O}\left(\frac{1}{\mu}\right)\right). \tag{3.27}$$

Thus, we have shown that the τ function of the KP hierarchy in Miwa variables (3.1) takes the determinant form (3.2) or, equivalently, (3.2) is the τ function of the KP hierarchy. Now we shall show how to go from a general point of the Grassmannian described by the (infinite) matrix $G=\exp\Sigma A_{ij}\tilde{\psi}_i\psi_j$ with two indices (∞^2) to special solutions determined, in particular, by a single (∞) or two ($2\times\infty$) functions of the same variable.

Let us return to the question of isolating the string solutions from all the solutions of the KP hierarchy, i,e., in some sense the problem of specifying the auxiliary conditions. Using the integrability, it is sufficient to prove only one of the infinite set of auxiliary conditions, the so-called string equation or the action of the $\mathcal{L}_{-1}$th generator of the Virasoro algebra. All the rest of the tower of Virasoro conditions follows from these two properties by induction.[25,111]

The action of the $\mathcal{L}_{-1}$th Virasoro generator (differentiation with respect to the spectral parameter) is associated with the operator

$$\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}=\mathrm{Tr}\frac{1}{V''(M)}\frac{\partial}{\partial M_t r}, \tag{3.28}$$

and so it is natural to ascertain first how the operator (3.28) acts on the partition function:

$$Z^{\{V\}}[M]=\frac{\det\Phi_i(\mu_j)}{\Delta(M)}\prod_i s(\mu_i),$$

$$s(\mu)=(V''(\mu))^{1/2}e^{V(\mu)-\mu V'(\mu)},$$

$$\Phi_i(\mu)=F_i(\lambda)=\left(\frac{\partial}{\partial\lambda}\right)^{i-1}F_1(\lambda),\quad \lambda=V'(\mu). \tag{3.29}$$

First, if we treat $Z^{\{V\}}$ as a function of the times T, then

$$\frac{1}{Z^{\{V\}}}\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}Z^{\{V\}}=-\sum_{n\geq 1}\mathrm{Tr}\left[\frac{1}{V''(M)M^{n+1}}\right]\frac{\partial\log Z^{\{V\}}}{\partial T_n}. \tag{3.30}$$

On the other hand, by directly applying (3.28) to the explicit expression (3.29), we obtain

$$\frac{1}{Z^{\{V\}}}\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}Z^{\{V\}}=-\mathrm{Tr}M+\frac{1}{2}$$

$$\times\sum_{i,j}\frac{1}{V''(\mu_i)V''(\mu_j)}\frac{V''(\mu_i)-V''(\mu_j)}{\mu_i-\mu_j}$$

$$+\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\log\det F_i(\lambda_j). \tag{3.31}$$

Below, we shall show that the expression

$$\frac{1}{Z^{\{V\}}}\mathcal{L}_{-1}^{\{V\}}Z^{\{V\}}=-\frac{\partial}{\partial T_1}\log Z^{\{V\}}+\mathrm{Tr}M$$

$$-\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\log\det F_i(\lambda_j) \tag{3.32}$$

can be used to determine the universal operator $\mathcal{L}_{-1}^{\{V\}}$. This definition has the form

$$\mathcal{L}_{-1}^{\{V\}}=\sum_{n\geq 1}\mathrm{Tr}\left[\frac{1}{V''(M)M^{n+1}}\right]\frac{\partial}{\partial T_n}$$

$$+\frac{1}{2}\sum_{i,j}\frac{1}{V''(\mu_i)V''(\mu_j)}\frac{V''(\mu_i)-V''(\mu_j)}{\mu_i-\mu_j}-\frac{\partial}{\partial T_1}, \tag{3.33}$$

which becomes the expression already known for monomial potentials $V(X)=X^{p+1}/(p+1)$ [we note that terms with $i=j$ are included in the sum on the right-hand side of (3.33)].

Thus, to prove the $\mathcal{L}_{-1}^{\{V\}}$ condition, it remains to be shown that the right-hand side of (3.32) is equal to zero, i.e.,

$$\frac{\partial}{\partial T_1}\log Z_N^{\{V\}}=\mathrm{Tr}\,M-\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\log\det F_i(\lambda_j). \tag{3.34}$$

In order to prove this, it is very important that the partition function is a τ function: $Z_N^{\{V\}}=\tau_N^{\{V\}}$. Here the left-hand side of the equation can be written as the residue of a ratio of τ functions:

$$\mathrm{res}_\mu\frac{\tau_N^{\{V\}}(T_n+\mu^{-n}/n)}{\tau_N^{\{V\}}(T_n)}=\frac{\partial}{\partial T_1}\log\tau_N^{\{V\}}(T_n). \tag{3.35}$$

If we now go over to Miwa variables, the τ function in the numerator will be given by the same expression as in the denominator, but with the additional parameter μ, i.e., it will be $\tau_{N+1}^{\{V\}}$. This observation is almost sufficient for deriving (3.34). For the simplest case $N=1$ we have $[\lambda=V'(\mu)]$

$$\tau_1^{\{V\}}(T_n)=\tau_1^{\{V\}}[\mu_1]=e^{V(\mu_1)-\mu_1V'(\mu_1)}[V''(\mu_1)]^{1/2}F(\lambda_1), \tag{3.36}$$

$$\tau_1^{\{V\}}(T_n+\mu^{-n}/n)=\tau_2^{\{V\}}[\mu_1,\mu]$$

$$=e^{V(\mu_1)-\mu_1V'(\mu_1)}e^{V(\mu)-\mu V'(\mu)}$$

$$\times\frac{[V''(\mu_1)V''(\mu)]^{1/2}}{\mu-\mu_1}$$

$$\times\left[F(\lambda_1)\frac{\partial F(\lambda)}{\partial\lambda}-F(\lambda)\frac{\partial F(\lambda_1)}{\partial\lambda_1}\right]$$
$$=\frac{e^{V(\mu)-\mu V'(\mu)}[V''(\mu)]^{1/2}F(\lambda)}{\mu-\mu_1}$$
$$\times\tau_1^{\{V\}}[\mu_1]\cdot\left[-\frac{\partial\log F(\lambda_1)}{\partial\lambda_1}+\frac{\partial\log F(\lambda)}{\partial\lambda}\right]. \quad (3.37)$$

The function F has the asymptotic behavior

$$F(\lambda)=\int dx\, e^{-V(x)+\lambda x}\sim e^{V(\mu)-\mu V'(\mu)}[V''(\mu)]^{-1/2}\times\left\{1+O\left(\frac{V''''}{V''V''}\right)\right\}. \quad (3.38)$$

If $V(\mu)$ increases at infinity $\mu\to\infty$ as μ^n, then $V''''/(V'')^2\sim\mu^{-n}$, and for what follows it is sufficient that $n=p+1>1$, so that in the brackets on the right-hand side of (3.38) the asymptotic form is $\{1+o(1/\mu)\}$, where $\mu\cdot o(\mu)\to 0$ for $\mu\to\infty$. Thus, the numerator of the right-hand side of (3.37) is constructed as $\sim 1+o(1/\mu)$, while the second term in the square brackets behaves as $[\partial\log F(\lambda)]/\partial\lambda\sim\mu[1+o(1/\mu)]$. Collecting all the terms, we find

$$\frac{\partial}{\partial T_1}\log\tau_1^{\{V\}}=\mathrm{res}_\mu\left(\frac{1+o\left(\frac{1}{\mu}\right)}{\mu-\mu_1}\left(-\frac{\partial\log F(\lambda_1)}{\partial\lambda_1}+\mu\left(1+o\left(\frac{1}{\mu}\right)\right)\right)\right)$$
$$=\mu_1-\frac{\partial\log F(\lambda_1)}{\partial\lambda_1}, \quad (3.39)$$

i.e., we have proved (3.34) for the special case $N=1$.

The proof for arbitrary N is practically the same. After a simple but lengthy calculation we obtain

$$\frac{\partial}{\partial T_1}\log\tau_N^{\{V\}}=\mathrm{res}_\mu\left(\frac{1+o(1/\mu)}{\Pi_{j=1}(\mu-\mu_j)}\mu^N\left([1+o(1/\mu)]-\frac{1}{\mu}\left[\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\log\det F_i(\lambda_j)\right]\cdot[1+\mathcal{O}(1/\mu)]\right)\right)$$
$$=\sum_{j=1}^N\mu_j-\mathrm{Tr}\frac{\partial}{\partial\Lambda_{\mathrm{tr}}}\log\det F_i(\lambda_j), \quad (3.40)$$

which completes the proof of (3.34) and the derivation of the form of the universal auxiliary condition of the $\mathcal{L}_{-1}^{\{V\}}$ operator.

In the special case of a monomial potential $V\equiv V_p=X^{p+1}/(p+1)$, the general expression (3.33) takes the more usual form[25,26]

$$\mathcal{L}_{-1}^{\{p\}}=\frac{1}{p}\sum_{n\geq 1}(n+p)T_{n+p}\frac{\partial}{\partial T_n}+\frac{1}{2p}\times\sum_{\substack{a+b=p\\ a,b\geq 0}}aT_abT_b-\frac{\partial}{\partial T_1}. \quad (3.41)$$

3.2. Exact solutions of topological (p,1) models and their deformation into theories of the Ginzburg–Landau type

For various potentials $V(X)$ the model (2.81) formally reproduces various theories of the (p,q) series as follows. The potential $V(X)=X^{p+1}/(p+1)$ gives the entire series of (p,q) string models with fixed p and all possible q. To fix q it is necessary to fix the times T in a special way, i.e., to set all $T_k=0$ except for T_1 and T_{p+q} (we note that this procedure spoils the symmetry between p and q present in the conformal theory).

As an illustration, let us consider two simple examples. First we set $p=2$, i.e., we start with the case of the KdV reduction of the KP hierarchy. In this case the string equation takes the form

$$\frac{1}{\tau_{\mathrm{KdV}}}\mathcal{L}_{-1}\tau_{\mathrm{KdV}}=\frac{1}{2}\sum_{\substack{k>1\\ k\ \mathrm{odd}}}kT_k\frac{\partial}{\partial T_{k-2}}\log\tau_{\mathrm{KdV}}+\frac{T_1^2}{4}=0, \quad (3.42)$$

or, differentiating again with respect to T_1, we find

$$\sum_{\substack{k>1\\ k\ \mathrm{odd}}}kT_k\frac{\partial^2}{\partial T_{k-2}\partial T_1}\log\tau_{\mathrm{KdV}}+T_1=0. \quad (3.43)$$

Using the definition of the Gel'fand–Dikiĭ polynomials

$$\frac{\partial^2}{\partial T_{k-2}\partial T_1}\log\tau_{\mathrm{KdV}}=[L^{2m-1}]_{-1}\equiv\mathcal{R}_m[u], \quad (3.44)$$

we have

$$\sum_{m\geq 0}(2m+1)T_{2m+1}\mathcal{R}_m[u]=0. \quad (3.45)$$

Now we can use the "rule" for isolating (p,q) critical points[25] for $p=2$ and $q=2m-1$, i.e., $(2,2m-1)$ solutions from (3.43) and (3.45). The simplest case is $m=1$, when $3T_3(\partial^2/\partial T_1^2)\log\tau_{\mathrm{KdV}}+T_1=0$. Using $u\sim(\partial^2/\partial T_1^2)\log\tau_{\mathrm{KdV}}$, we find the solution of the KdV equation $u\sim T_1/T_3$, or, fixing T_3, we have $F=\log\tau\sim T_1^3$, which coincides with the earlier results (2.10) for the $c=-2$ theory interacting with 2D gravity, in which $\langle P^3\rangle=1$, where P is the unit operator (2.9), $P=c\bar{c}e^{\phi}$. This example is the best known case of pure *topological* gravity.

A less trivial example is the (2,3) theory with $m=2$, in which (3.45) and the explicit expression $\mathcal{R}_2\sim u^2+u''$ lead to the appearance of the first Panlevé equation $u^2+u''=T_1$. This example corresponds to the theory of pure (physical) gravity, where the solution, as we shall see, is much less trivial than in the preceding case.

From this point of view the presence of all the (p,q) critical points in the model (2.81) is a purely formal statement. For the potential $V(X)=X^{p+1}/(p+1)$ the partition function $Z[V|T_k]=\tau_V[T_k]\equiv\tau_p[T_k]$ satisfies the string equation

$$\sum_{k=1}^{p-1} k(p-k)T_kT_{p-k}+\sum_{k=1}^{\infty}(p+k)\left(T_{p+k} - \frac{p}{p+1}\delta_{k,1}\right)\frac{\partial}{\partial T_k}\log\tau_p[T]=0, \tag{3.46}$$

i.e., the τ function is defined as an expansion in (small) Miwa times (3.1) near zero values of all the times except T_{p+1}, which is shifted by a finite factor $p/(p+1)$ corresponding to the $(p,1)$ model according to the foregoing arguments. Thus, we see that a matrix integral representation of the Kontsevich type corresponds to the solution of $(p,1)$ string models, which describe the interaction of (A_n) topological matter with topological gravity or topological W gravity.

Let us now turn to deformations of the pure $(p,1)$ theory[64,112] related to deformation of the potential and the so-called p or Whitham times, which are directly related to deformation of the *moduli* of the solutions. In fact, there exists *a priori* another integrable structure in the model (2.81) where the flows are described by times related to nontrivial coefficients of the potential V. As a result, theories with the monomial potential $V_p(X)=X^{p+1}/(p+1)$ and an arbitrary polynomial of degree $p+1$ turn out to be closely related to each other.

To show this, we return to the question of calculating the derivatives of $Z_{\rm GKM}$ with respect to the times T_k. These derivatives determine the nonperturbative correlation functions in string theory and are of particular interest from the "internal" point of view in an integrable system. The derivatives with respect to the times T_k with $k\geq p+1$ (corresponding to correlators of irrelevant operators) are rather complicated to calculate. However, for times T_k with $1\leq k\leq p$ the situation is much simpler. Using the obvious definition of the average such that the partition function (2.81) is $Z_{\rm GKM}=\langle 1\rangle$, we have

$$\left.\frac{\partial Z_{\rm GKM}}{\partial T_k}\right|_V=\langle {\rm Tr}\,M^k-{\rm Tr}\,X^k\rangle,\quad 1\leq k\leq p. \tag{3.47}$$

Here it is understood that the derivative on the left-hand side is calcuated for *fixed* potential $V(x)=\Sigma^{p+1}(v_k/k)X^k$.

The right-hand side of (3.47) can also be written as

$$\left.\frac{\partial Z_{\rm GKM}}{\partial T_k}\right|_V=\left\langle {\rm Tr}\frac{\partial V(M)}{\partial v_k}-{\rm Tr}\frac{\partial V(X)}{\partial v_k}\right\rangle,\quad 1\leq k\leq p, \tag{3.48}$$

which is already very similar to, but actually not the same as, the expression $-(\partial/\partial v_k)Z_{\rm GKM}$, which differs from (3.48) by correction factors. The problem is that the expression $(\partial/\partial v_k)Z_{\rm GKM}$ contains contributions not only from differentiating $V(X)-V(M)$ from the exponential in (2.81), but also a contribution from the term $V'(M)(X-M)\equiv W(M)(X-M)$, as well as the derivative of the normalization factor in (2.82). These corrections can be split into two parts:

$$\mathcal{O}\left(\frac{\partial}{\partial v_k}W\right)+\text{``quantum corrections''}. \tag{3.49}$$

The first part can be eliminated by introducing a new "spectral parameter" $W(M)=\tilde{M}^p$, as a result of which the new times $\tilde{T}_k=(1/k){\rm Tr}\tilde{M}^{-k}$ arise naturally. It is clear from (3.49) that the $\{v_k\}$ themselves are not the "correction" variables for an arbitrary potential. It is easy to see that it is much more convenient to work with linear combinations of them $\{t_k\}$, defined as[36]

$$t_k=-\frac{p}{k(p-k)}\,{\rm res}\,W^{1-k/p}(\mu)d\mu. \tag{3.50}$$

Using (3.50), we easily find

$$\mu=\frac{1}{p}\sum_{-\infty}^{p+1}kt_k\tilde{\mu}^{k-p},\quad V(\mu)-\mu V'(\mu)=-\sum_{-\infty}^{p+1}t_k\tilde{\mu}^k. \tag{3.51}$$

This equation (3.51) shows that the exponential factor in (2.81) is none other than the standard essential singularity of the Baker–Akhiezer function for the hierarchy where the integrable flows are parametrized by the p times.

Finally, direct calculation gives

$$Z[V|T_k]=\tau_V[T_k]=\exp\left(-\frac{1}{2}\sum A_{ij}(t)(t_i+\tilde{T}_i)(t_j+\tilde{T}_j)\right)\tau_p[t_k+\tilde{T}_k], \tag{3.52}$$

where $A_{ij}={\rm res}_\mu W^{i/p}dW_+^{j/p}$ and $f(\mu)_+$ denotes the nonnegative part of the Laurent series $f(\mu)=\Sigma f_i\mu^i$: $f(\mu)_+=\Sigma_{i\geq 0}f_i\mu^i$. It is now easy to show that $\tau_p[T]\equiv\tau_{V_p}[T]$ is the τ function of the p reduction of the KP hierarchy (the hierarchy of the pth KdV equation).

The meaning of (3.52) is that "shifting" by the flows of the p times (3.50) allows the τ function to be expressed fairly simply in terms of the τ function of the p reduction, now depending only on the sum of the times $\tilde{T}_k$ and t_k. The replacement of the spectral parameter $M\to\tilde{M}=f(M)=W^{1/p}(M)$ (and the corresponding time replacement $T_k\to\tilde{T}_k$) is a natural operation in the construction of equivalent hierarchies.[113]

Actually, the relation between the τ functions of equivalent hierarchies can be obtained from the following identity transformation:

$$\tau(T)=\frac{\Delta(\tilde{\mu})}{\Delta(\mu)}\prod_i[f'(\mu_i)]^{1/2}\tilde{\tau}(\tilde{T}), \tag{3.53}$$

where the determinant representation of $\tilde{\tau}(\tilde{T})$ as a function of the times $\tilde{T}$ (3.2) is constructed from the basis vectors $\tilde{\phi}(\tilde{\mu})=[f'(\mu(\tilde{\mu}))]^{1/2}\phi_i(\mu(\tilde{\mu}))$. It can be shown by direct calculation that the factor in front of the τ function on the right-hand side of (3.53) can be rewritten as

$$\frac{\Delta(\tilde{\mu})}{\Delta(\mu)}\prod_i[f'(\mu_i)]^{1/2}=\exp\left(-\frac{1}{2}\sum_{i,j}A_{ij}\tilde{T}_i\tilde{T}_j\right), \tag{3.54}$$

where $A_{ij}={\rm res}\,f^i(\lambda)d_\lambda f_+^i(\lambda)$. It follows from (3.53) that

$$\tau(T(\tilde{T}))=\tilde{\tau}(\tilde{T})\exp\left(-\frac{1}{2}\sum_{i,j}A_{ij}\tilde{T}_i\tilde{T}_j\right). \tag{3.55}$$

We introduce the τ function $\hat{\tau}(\tilde{T})$ of the p-reduced KP hierarchy, defined as

$$\tilde{\tau}(\tilde{T})\equiv\frac{\hat{\tau}(\tilde{T})}{\tau_0(t)}\exp\left(\sum_j jt_{-j}\tilde{T}_j\right),\quad \tau_0(t)=e^{-1/2\Sigma A_{ij}t_it_j}, \tag{3.56}$$

for which instead of

$$\tau_V[T]=\frac{\det\phi_i(\mu_j)}{\Delta(\mu)} \tag{3.57}$$

we have

$$\frac{\tau_p[t+\tilde{T}]}{\tau_p[t]}=\frac{\det\hat{\phi}_i(\tilde{\mu}_j)}{\Delta(\tilde{\mu})}, \tag{3.58}$$

and the Grassmannian points corresponding to (3.57) and (3.58) are determined by the basis vectors

$$\phi_i(\mu)=[W'(\mu)]^{1/2}\exp(V(\mu)-\mu W(\mu))\int x^{i-1}e^{-V(x)+xW(\mu)}dx \tag{3.59}$$

and

$$\hat{\phi}_i(\tilde{\mu})=[p\tilde{\mu}^{p-1}]^{1/2}\exp\left(-\sum_{j=1}^{p+1}t_j\tilde{\mu}_j\right)\times\int x^{i-1}e^{-V(x)+x\tilde{\mu}^p}dx. \tag{3.60}$$

Here it is easy to show that $\hat{\tau}_p(T)$ satisfies the string equation (the L_{-1} condition) with the KP times shifted in the following manner:

$$\sum_{k=1}^{p-1}k(p-k)(t_k+\tilde{T}_k)(t_{p-k}+\tilde{T}_{p-k})+\sum_{k=1}^{\infty}(p+k)(t_{p+k}+\tilde{T}_{p+k})\frac{\partial}{\partial\tilde{T}_k}\log\hat{\tau}_p[t+\tilde{T}]=0, \tag{3.61}$$

where the t_i defined in (3.50) are *identically* equal to zero for $i\geqslant p+2$.

We can extract at least two different consequences from (3.52) and (3.61). First, the generating function in the case of deformation of a monomial potential (equivalent to a polynomial of the same degree) is expressed in terms of the τ function of the equivalent (in the sense of Ref. 113) p-reduced KP hierarchy. Second, in the deformed case, not only t_{p+1} but all the t_k with $k\leqslant p+1$ are nonzero. We shall refer to such theories as *topologically deformed* $(p,1)$ models [in order not to confuse them with the *intrinsically* $(p,1)$ models specified by monomial potentials $V_p(X)$], since the deformation is topological in the sense that it preserves all the properties of topological models. From the viewpoint of field theories, these models correspond to 2D $N=2$ twisted Ginzburg–Landau theories interacting with topological gravity. The relation obtained above has been reproduced in the spherical limit by another method.[35]

3.3. Nontopological solutions and *pq* duality

The scheme presented above for constructing topological solutions has a clear interpretation in the language of canonical quantization. In fact, the exact nonperturbative solutions of topological $(p,1)$ theories are described by the generating function (2.81), which can be viewed in a certain sense as a representation of the functional integral for the equations[24] $[\hat{P},\hat{Q}]=1$, i.e., simply the Heisenberg algebra in a realization in which $\hat{P}$ and $\hat{Q}$ are differential operators of finite order (p and q, respectively), the pth order of the operator $\hat{P}$ specifying the p reduction, while q corresponds to the qth critical point. The semiclassical commutator is transformed into a Poisson bracket[36,114] $\{P,Q\}=1$, where now $P(x)$ and $Q(x)$ are definite functions (polynomials). It is easily seen that the case considered above corresponds to the situation in which $Q(x)\equiv x$ is a first-degree polynomial, and the total degree $p-P(x)$ must be identified with $W(x)\equiv V'(x)$.[14)] Here the expression in the exponent in (2.81), (3.59), and (3.60) acquires the natural meaning of the action functional:

$$S_{p,1}(x,\mu)=-V(x)+xW(\mu)=-\int_0^x dyW(y)Q'(y)+Q(x)W(\mu),$$

$$W(x)=V'(x)=x^p+\sum_{k=1}^{p}v_kx^{k-1},\quad Q(x)=x^q, \tag{3.62}$$

and, naturally, its generalization to the case of arbitrary (p,q) models becomes

$$S_{W,Q}=-\int_0^x dyW(y)Q'(y)+Q(x)W(\mu),$$

$$W(x)=V'(x)=x^p+\sum_{k=1}^{p}v_kx^{k-1},$$

$$Q(x)=x^q+\sum_{k=1}^{q}\bar{v}_kx^{k-1}. \tag{3.63}$$

As before, variation of the action (3.63) gives $W(x)=W(\mu)$ with one of the solutions $x\equiv\mu$, and the value of the action at the extremum $x=\mu$ has the form

$$S_{W,Q}|_{x=\mu}=\int_0^\mu dyW'(y)Q(y)=\sum_{k=-\infty}^{p+q}t_k\tilde{\mu}^k, \tag{3.64}$$

where $\tilde{\mu}^p=W(\mu)$ and

$$t_k\equiv t_k^{(W,Q)}=-\frac{p}{k(p-k)}\operatorname{res}W^{1-k/p}dQ. \tag{3.65}$$

It should be noted that the value of the action (3.63) at the extremum, written in the form (3.64), determines the semiclassical (or dispersionless) limit of the p-reduced KP hierarchy[36,114] with $p+q-1$ independent flows. It was shown above that in the case of the topologicaly deformed

$(p,1)$ model, the semiclassical hierarchy is *exact* in the following sense: solutions of the hierarchy in which the flows are determined by the p times are also solutions of the *full* equations of the KP hierarchy [the $t+T$ formula (3.52)], and the first of the set of basis vectors is exactly the Baker–Akhiezer function restricted to the "small" phase space. This is certainly not so for the general case of (p,q) models: here the semiclassical limit is no longer exact, and in order to find the explicit form of the basis vectors it is necessary to solve the "full" problem, i.e., to find the exact solutions of the complete (reduced) KP hierarchy along the first $p+q-1$ flows. Nevertheless, the presence of a "semiclassical component" in the full integrable structure of these models can in principle give some useful information. For example, it can be assumed that the coefficients of the asymptotic expansions of the basis vectors are expressed only in terms of the derivatives of the semiclassical τ function.

Returning to (3.65), we immediately note that now the p times are identically zero only for $k \geqslant p+q+1$, while for the $(k=p+q)$th time we have

$$t_{p+q} \equiv t_{p+q}^{(W,Q)} = \frac{p}{p+q}, \tag{3.66}$$

and the correct critical behavior is obtained by "twisting" all the times $\{t_k\}$ with $k<p+q$ so that they become equal to zero. The exact expression for the basis vectors in the Grassmannian in the general case of (p,q) models takes the form

$$\phi_1(\mu) = [W'(\mu)]^{1/2} \exp(-S_{W,Q}|_{x=\mu}) \times \int d\mathcal{M}_Q(x) f_i(x) \exp S_{W,Q}(x,\mu), \tag{3.67}$$

where $d\mathcal{M}_Q(x)$ is the integration measure. As will be seen in what follows, for the general case of (p,q) models the measure is determined by *two* polynomials W and Q and has the form

$$d\mathcal{M}_Q(z) = [Q'(z)]^{1/2} dz, \tag{3.68}$$

which follows from the string equation. In choosing the measure in the form (3.68), to ensure the correct asymptotic behavior of the basis vectors $\phi_i(\mu)$ it is necessary to choose the functions $f_i(x)$ (not necessarily monomials or polynomials!) to satisfy the same asymptotic condition $f_i(x) \sim x^{i-1}[1+O(1/x)]$. Finally, for the basis vectors to satisfy the string equation, it is necessary that they satisfy two conditions: the reduction condition

$$W(\mu)\phi_i(\mu) = \sum_j C_{ij}\phi_j(\mu) \tag{3.69}$$

and invariance under the action of the Kac–Schwarz operator (3.17):

$$A^{(W,Q)}\phi_i(\mu) = \sum A_{ij}\phi_j(\mu), \tag{3.70}$$

where

$$A^{(W,Q)} \equiv N^{(W,Q)}(\mu) \frac{1}{W'(\mu)} \frac{\partial}{\partial \mu} [N^{(W,Q)}(\mu)]^{-1} = \frac{1}{W'(\mu)} \frac{\partial}{\partial \mu} - \frac{1}{2} \frac{W''(\mu)}{W'(\mu)^2} + Q(\mu),$$
$$N^{(W,Q)}(\mu) \equiv [W'(\mu)]^{1/2} \exp(-S_{W,Q}|_{x=\mu}). \tag{3.71}$$

The string equation is a consequence of (3.69) and (3.70). The structure of the action leads directly to

$$A^{(W,Q)}\phi_i(\mu) = N^{(W,Q)}(\mu) \int d\mathcal{M}_Q(z) Q(z) f_i(z) \times \exp S_{W,Q}(z,\mu), \tag{3.72}$$

and the condition (3.70) can be reformulated as the property of Q reduction of the (dual) basis $\{f_i(z)\}$:

$$Q(z) f_i(z) = \sum A_{ij} f_i(z). \tag{3.73}$$

Let us now return to the W-reduction condition. Multiplying $\phi_i(\mu)$ by $W(\mu)$ and integrating by parts, we obtain

$$W(\mu)\phi_i(\mu) = N^{(W,Q)}(\mu) \int d\mathcal{M}_Q(z) f_i(z) \frac{1}{Q'(z)} \times \frac{\partial}{\partial z}[\exp Q(z) W(\mu)] \times \exp\left[-\int_0^z dy\, W(y) Q'(y)\right]$$
$$= -N^{(W,Q)}(\mu) \int d\mathcal{M}_Q(z) \exp[S_{W,Q}(z,\mu)] \times \left(\frac{1}{Q'(z)} \frac{\partial}{\partial z} - \frac{1}{2} \frac{Q''(z)}{Q'(z)^2} - W(z)\right) f_i(z)$$
$$\equiv -N^{(W,Q)}(\mu) \int d\mathcal{M}_Q(z) \times \exp[S_{W,Q}(z,\mu)] A^{(Q,W)} f_i(z). \tag{3.74}$$

Thus, for the dual basis $\{f_i(z)\}$ the condition (3.69) becomes

$$A^{(Q,W)} f_i(z) = -\sum C_{ij} f_j(z), \tag{3.75}$$

where we have introduced the notation $A^{(Q,W)}$ $(\neq A^{(W,Q)})$ for the dual Kac–Schwarz operator:

$$A^{(Q,W)} = \frac{1}{Q'(z)} \frac{\partial}{\partial z} - \frac{1}{2} \frac{Q''(z)}{Q'(z)^2} - W(z). \tag{3.76}$$

Equations (3.67) and (3.68) are exact integral expressions for the basis vectors which are solutions of (p,q) string models.[112] The meaning of these expressions is contained in the explicit form of the integral transformation relating dual (p,q) and (q,p) nonperturbative exact string solutions. We shall call this the pq duality transformation (in general, W–Q duality) for exact nonperturbative generating functions (see also Ref. 115). The principal consequence of these expressions is that the *general solution* of $c \leqslant 1$ 2D nonperturbative string theory is formulated using *two* (polynomial) functions $W(x)$ and $Q(x)$.

3.4. String field theory and the $c \to 1$ limit

Finally, let us discuss the question of why the scheme proposed above for a unified description of the nonperturbative regime in a class of string models can be viewed as an attempt to construct a *string field theory* or an effective formulation of string theory in which the world sheet of the string no longer appears explicitly. From the very start it should be noted that by string field theory we mean something more than the traditional definition as a field theory of functionals defined on string loops. By string field theory we shall mean an effective theory allowing a unified treatment of the string vacua (the solutions of the classical equations of motion of string field theory—2D conformal field theories interacting with 2D gravity), and assuming that the effective action is a function having (perturbative) expansions about the given string vacua. Such a theory, in particular, must be able to describe transitions between different string vacua and other nonperturbative effects.

The scheme studied above can be referred to as a *string* field theory, as it is based on the two-dimensional geometry of the world sheet, which allows the correlators (in the topological sector) to be expressed in terms of integrals of differential forms on the moduli space.[40] This is the main difference from ordinary field theory, where there are no restrictions imposed by the two-dimensional geometry. The general form of the construction proposed above is related to the cell decomposition of the moduli space.[40,116]

Thus, at present we are dealing with a theory describing various (p,q) models interacting with 2D gravity, outside the framework of perturbation theory and technically based on (differential) equations in coupling-constant space imposed on the generating function of the physical correlators. The main statement is the identification of the generating function with the τ function of an integrable hierarchy of the KP/Toda type, which is not defined as a *global* function in coupling-constant space, but has some fixed expansion about each critical point, reproducing the perturbation series of the original first-quantized theory (2.67). However, for naive transitions in coupling-constant space from one solution to another, impediments appear which are related to the poor convergence of the perturbative expansions, i.e., a nontrivial analytic continuation becomes necessary, which leads to intrinsically nonperturbative contributions. For the simplest topological $(p,1)$ theories this scheme can be described in the language of the effective integral representation, which reduces to an integral over Hermitian matrices (2.81), and the exact integral expressions for the general case are constructed in a much more complicated fashion.

This scheme is applicable, in principle, also to the limiting case $c=1$. However, a $c=1$ theory of general form is much more "saturated" than $c<1$ topological models, and so the naive limit from $c<1$ effective matrix models leads only to the formulation of strongly degenerate $c=1$ theories.

All these cases are more or less based on generalized Penner models for calculating the Euler characteristics of moduli spaces of complex curves. In fact, the determinant representation of the partition function of the Penner model[116] already by itself presupposes that (for fixed times) this partition function is the τ function of the Toda lattice. The existence of an integral representation of the Penner model indicates that this theory is in some sense an analog of the generalized Kontsevich model. In fact, the solution of the Penner model,

$$\mathcal{Z} \sim \det \mathcal{H}_{ij}^{(\alpha)}, \tag{3.77}$$

with $\mathcal{H}_{ij}^{(\alpha)} = \Gamma(\alpha+i+j-1)$, is a special case of the topological theories studied above.

In order to demonstrate this, we first note that for any solution of the KP hierarchy there is an explicit relation between the determinant representation in the τ function of the Miwa variables,

$$\tau_{\mathrm{KP}}[T_k] = \frac{\det_{ij} \phi_i(\mu_j)}{\Delta(\mu)}, \tag{3.78}$$

and the determinant representation of the τ function characteristic for solutions of the Toda lattice,

$$\tau_N[T_{-k}, T_k] = \det_{ij} H_{i+N,j+N}[T_{-k}, T_k], \tag{3.79}$$

where

$$\Delta(\mu) = \prod_{i>j} (\mu_i - \mu_j), \quad \phi_i(\mu) = \mu^{i-1}\left(1 + \mathcal{O}\left(\frac{1}{\mu}\right)\right),$$

$$T_k = \frac{1}{k} \sum_i \mu_i^{-k}, \quad k>0, \tag{3.80}$$

and

$$\partial H_{ij} / \partial T_k = H_{i,j-k}, \quad j>k>0,$$
$$\partial H_{ij} / \partial T_{-k} = H_{i-k,j}, \quad i>k>0. \tag{3.81}$$

The relation between the representations (3.78) and (3.79) is most simply stated in terms of Schur polynomials, defined by

$$\mathcal{P}[z|T_k] \equiv \exp\left\{ \sum_{k>0} T_k z^k \right\} = \sum z^k P_k[T], \tag{3.82}$$

in particular, $P_{-n}=0$ for any $n>0$; $P_0[T]=1$; $P_1[T] = T_1$; $P_2[T] = T_2 + \frac{1}{2}T_1^2$; $P_3[T] = T_3 + T_2 T_1 + \frac{1}{6} T_1^3$, etc. An important property of the Schur polynomials is the equation $\partial P_k / \partial T_n = P_{k-n}$ following from $\partial \mathcal{P} / \partial T_k = z^k \mathcal{P}$. This property allows the time dependence of the matrix $H_{ij}[T]$ satisfying (3.81) to be expressed in terms of Schur polynomials:

$$H_{ij}[T_{-p}, T_p] = \sum_{\substack{k \leq i \\ l \geq -j}} P_{i-k}[T_{-p}] H_{kl} P_{l+j}[T_p], \tag{3.83}$$

where $H_{kl} \equiv H_{kl}[0,0]$ is now a matrix independent of the time T.

Let us first consider the case of zero negative times and zero time $N = T_{-k} = 0$. Then we shall allow the zero-valued time to take any (positive integer) value[15)] $N>0$ and introduce negative times T_{-k}.

For a given system of basis vectors $\phi_i(\mu)$ for $i>0$ we introduce, by definition,

$$H_{ij}[T_{-k}=0, T_k] = \oint_{z \to 0} \phi_i(z) z^{-j} \mathcal{P}[z|T_k] dz, \quad i>0. \tag{3.84}$$

The integration contour, which encircles the origin, can be deformed to an infinitely distant point. The integral then sits on the singularities of the function $\mathcal{P}[z]$ if they exist. Substituting the definition (3.82) of the function $\mathcal{P}[z]$ into (3.84), we obtain (3.83), in which $P_{k-i}[T_{-m}=0]=\delta_{ki}$, and also

$$H_{kl}=\oint_{z\to 0}\phi_k(z)z^i dz. \tag{3.85}$$

In order to prove the equality of (3.78) and (3.79), in (3.82) we transform to Miwa variables (3.80):

$$\mathcal{P}[z|T_k]=\frac{\det M}{\det(M-Iz)}=\prod_i \frac{\mu_i}{(\mu_i-z)}$$
$$=\left[\prod_i \mu_i\right]\sum_k \frac{(-)^k}{(z-\mu_k)}\frac{\Delta_k(\mu)}{\Delta(\mu)}, \tag{3.86}$$

where $\Delta_k(\mu)\equiv\Pi_{i>j;i,j\neq k}(\mu_i-\mu_j)$. Now the contribution to (3.84) comes only from the poles of the function $\mathcal{P}[z|T_k]$ at the points μ_k:

$$H_{ij}[T_{-k}=0,T_k]=\oint_{z\to 0}\phi_i(z)z^{-j}\mathcal{P}[z|T_k]dz$$
$$=\frac{\Pi_i\mu_i}{\Delta(\mu)}\sum_k (-)^k\phi_i(\mu_k)\frac{\Delta_k(\mu)}{\mu_k^j}. \tag{3.87}$$

The sum on the right-hand side of (3.87) has the form of a matrix product, and so

$$\det H_{ij}=\det \phi_i(\mu_k)\prod_k \left[\frac{\Pi_i\mu_i}{\Delta(\mu)}(-)^k\Delta_k(\mu)\right]\cdot\det\frac{1}{\mu_k^j}. \tag{3.88}$$

The last determinant on the right-hand side of (3.88) is equal to $\Delta(1/\mu)\sim\Delta(\mu)\cdot[\Pi_k\mu_k^N]^{-1}$. In addition, $\Pi_k[\Delta_k(\mu)/\Delta(\mu)]=\Delta(\mu)^{-2}$, and, collecting all the terms, we verify the existence of the equality $\det H_{ij}=\det \phi_i(\mu_j)/\Delta(\mu)$, as required.

Let us now consider the introduction of zero and negative times. Zero time n arises simply as a single-time shift of the labels i and j of the matrix H_{ij}: $H_{ij}\to H_{i+N,j+N}$. We can therefore write

$$H_{i+N,j+N}[0,T_k]=\oint_{z\to 0}\phi_i^{\{N\}}(z)z^{-j}\mathcal{P}[z]dz, \tag{3.89}$$

with the vectors $\phi_i^{\{N\}}(z)=z^{-N}\phi_{i+N}(z)$. This completely solves the problem of zero time N for positive integer values of N.

As far as negative times are concerned, as soon as the matrix H_{kl} is defined they can be introduced by means of (3.83), so that

$$H_{i+N,j+N}[T_{-k},T_k]\equiv\sum_{k\leq i}P_{i-k}[T_{-p}]H_{k+N,j+N}[0,T_l]$$
$$=\oint_{z\to 0}\phi_i^{\{T_{-k},N\}}(z)z^{-j}$$
$$\times\mathcal{P}\left[\frac{1}{z}\middle|T_{-k}\right]\mathcal{P}[z|T_k]dz, \tag{3.90}$$

with the basis vectors

$$\phi_i^{\{T_{-k},N\}}(z)\equiv\left(\mathcal{P}\left[\frac{1}{z}\middle|T_{-k}\right]\right)^{-1}\sum P_{i-k}[T_{-l}]\phi_k^{[N]}$$
$$=z^{-N}\exp\left(-\sum_{k>0}T_{-k}z^{-k}\right)$$
$$\times\sum P_k[T_{-l}]\phi_{i+N-k}(z). \tag{3.91}$$

We have introduced an auxiliary exponential factor into the definition of the basis vectors (3.91) in order to ensure the correct asymptotic behavior $\phi_i^{\{T_{-k},N\}}(z)=z^{i-1}\{1+\mathcal{O}(1/z)\}$.

An important special case of the hierarchy of the Toda lattice is its reduction to a Toda chain (see, for example, Ref. 98). The reduction to a Toda chain can easily be described in terms of a fermionic operator G specifying the Grassmannian point and satisfying the condition $[G,J_k+J_{-k}]=0$, and also in determinant language. In the last case the symmetry condition becomes

$$[H,\Lambda+\Lambda^{-1}]=0, \tag{3.92}$$

where Λ is a shifting matrix, $\Lambda_{ij}\equiv\delta_{i,j-1}$. Satisfaction of this condition leads to the appearance of the τ function of the hierarchy of the Toda chain, which now depends only on the sum of positive and negative times $t_k=\frac{1}{2}(T_k+T_{-k})$, and is independent of their difference (this can be viewed as a defining property of the Toda chain). We note that the solution of the condition (3.92) is $H_{i,j}=\mathcal{H}_{i-j}$.

It is now easy to reconstruct the dependence on negative and zero times in the integral expressions for the string solutions of the KP hierarchy, so that the partition function of the effective theory is transformed into the τ function of the hierarchy of the Toda lattice. The corresponding set of functions—basis vectors at the Grassmannian point corresponding to the generalized Kontsevich model—is given by the integral expressions

$$\phi_i^{\{V\}}(\mu)=e^{V(\mu)-\mu V'(\mu)}\sqrt{V''(\mu)}\int dx\, x^{i-1}e^{-V(x)+xV'(\mu)}$$
$$\equiv s(\mu)\int dx\, x^{i-1}e^{-V(x)+xV'(\mu)}\equiv\langle x^{i-1}\rangle_\mu. \tag{3.93}$$

The dependence on N and T_{-k} is introduced according to the rule[16)]

$$\phi_i^{\{V,N,T_{-k}\}}(\mu)\equiv\left\langle x^{i-1}\left[\frac{x}{\mu}\right]^N \exp\left(\sum_{l>0} T_{-l}(x^{-l}-\mu^{-l})\right)\right\rangle_\mu$$
$$=\frac{\sqrt{V''(\mu)}e^{V(\mu)-\mu V'(\mu)}}{\mu^N}$$
$$\times\int dx\, x^{N+i-1}e^{-V(x)+xV'(\mu)}$$
$$\times\exp\left(\sum_{l>0} T_{-l}(x^{-l}-\mu^{-l})\right)$$
$$=e^{\hat V(\mu)-\mu V'(\mu)}\sqrt{V''(\mu)}\int dx\, x^{i-1}e^{-\hat V(x)+xV'(\mu)}, \tag{3.94}$$

where $\hat V(X)\equiv V(X)-N\log X-\Sigma_{k>0}t_{-k}X^{-k}$. The initial potential V is identified with $\hat V_+$. It follows immediately from (3.94) that the partition function of the generalized Kontsevich model taking into account the dependence on zero and negative times (which is automatically the τ function of the hierarchy of the Toda lattice) is

$$\hat Z_{\{\hat V\}}[M]=e^{\mathrm{Tr}\hat V(M)-\mathrm{Tr}M\hat V'_+(M)}\frac{\int DXe^{-\mathrm{Tr}\hat V(X)+\mathrm{Tr}\hat V'_+(M)X}}{\int dXe^{-\mathrm{Tr}\hat U_{+,2}(X,M)}}. \tag{3.95}$$

It is now easy to introduce the dependence on positive and negative times in the Penner model (3.77) and to reconstruct $\Phi_k^{\{V\}}(z)$ from (3.91). In fact,

$$h_{ij}^{(\alpha)}=\mathcal{H}_{ij}^{(\alpha)}=\Gamma(\alpha-1+i+j)$$
$$=\int_0^\infty\frac{dy}{y}e^{-y}y^{\alpha-1+i+j}=\oint\phi_i^{(\alpha)}(z)z^j \tag{3.96}$$

immediately gives

$$\phi_i^{(\alpha)}(z)=\int_0^\infty\frac{dy}{y}e^{zy-y}y^{\alpha-1+i}, \tag{3.97}$$

which is a representation in the spirit of the topological models studied above. The essential difference from the $c<1$ examples considered above is in the definition of the integration contour in (3.97), and also in the fact that the dependence on the parameter z is trivial. The integral is easily calculated, giving

$$\phi_i^{(\alpha)}(z)=\frac{\Gamma(\alpha+i)}{(z-1)^{\alpha+i}}\equiv\phi_{\alpha+i}(z),$$
$$\left(\frac{\partial}{\partial z}\right)^j\phi_i^{(\alpha)}(z)=(-)^j\phi_{i+j}^{(\alpha)}(z)=(-)^j\frac{\Gamma(\alpha+i+j)}{(z-1)^{\alpha+i+j}}. \tag{3.98}$$

Introducing negative times, we find[118]

$$\phi_i^{(\alpha)}(z|T_{-p})=z^{-\alpha}\exp\left(-\sum_{p>0}T_{-p}z^{-p}\right)$$
$$\times\sum_k P_k[T_{-p}]\phi_{i-k}^{(\alpha)}(z) \tag{3.99}$$

or

$$Z_{c=1}\sim\int DY\exp\mathrm{Tr}\,ZY+\alpha\,\mathrm{Tr}\log Y+\sum_{k>0}T_{-k}\mathrm{Tr}\,Y^{-k} \tag{3.100}$$

with positive times $T_{+k}=(1/k)\mathrm{Tr}Z^k$. It should be noted that (3.100) has been obtained independently by comparison with the results of calculating tachyon amplitudes in a $c=1$ theory.[119]

Finally, let us make a few remarks about $c=1$ theories. In general, here we would expect to obtain the most general (nonreduced) τ function of the KP or Toda-lattice hierarchy satisfying some (again, nonreduced) string equation. This situation would more likely correspond to taking the "direct sum" of the various (p,q) theories rather than the limit from the region $c<1$. However, in some degenerate cases the direct limit $c\to1$ studied above is also meaningful. These degenerate cases essentially turn out to be $c=1$ analogs of (p,q) string models and correspond to the topological sector of the $c=1$ theory.

In fact, it is easy to see that in the special case in which $p=\pm q$, Eqs. (3.69) and (3.70) are greatly simplified and the system degenerates into a single equation. This case certainly does not correspond to the minimal series, where the numbers in the pair (p,q) must be relatively prime. Nevertheless, as before, it is possible to satisfy both conditions, reduction and invariance under the action of the Kac–Schwarz operator. The resulting solutions, as can be seen from the expression for the central charge of the matter, formally correspond to $c=1$ for $p=q$ and $c=25$ for $p=-q$.

The simplest example again arises for $q=1$. In this case the "$c=1$" theory turns out to be equivalent to an auxiliary discrete matrix model,[117] while "$c=25$" literally corresponds to the result expected from the development of the Penner approach.[118,119] In fact, if we take in general the (nonpolynomial) functions $W(x)=x^{-\beta}$ and $Q(x)=x^{\beta}$, the action acquires a logarithmic term $S_{-\beta,\beta}=-\beta\log x+x^\beta/\mu^\beta$, and Eqs. (3.69) and (3.70) lead to simple rational solutions. It is easy to see that $\beta=1$ leads directly to the Penner model in an external field, which, as we have seen, corresponds more closely to the theory "dual" to the $c=1$ theory with central charge of the matter $c_{\mathrm{matter}}=25$ and strongly nonunitary realization of the conformal matter.[17)]

4. NONPERTURBATIVE RESULTS IN 4D $N=2$ SUPERSYMMETRIC GAUGE THEORIES: COMPLEX CURVES AND INTEGRABLE SYSTEMS

In the preceding sections we studied nonperturbative solutions of topological string models for which there are *explicit* exact expressions describing the generating functions of *all* the correlators. This is primarily related to the fact that these solutions can be regarded as a deformation of trivial finite-gap solutions for which the spectral curve Σ is a complex sphere and, in particular, then only the parameters related to the residues at the points in (3.65) are important, and the eigenvalues of the operators determining the generating differential are polynomial functions (3.63) on the sphere CP^1. Here the prepotential—the logarithm of the semiclassical τ function—is a simple polynomial function of the times, $\mathcal{F}=t_1^3/6+\ldots$, and the topological correlation functions are the expansion coefficients of the prepotential; as we have seen, these are numbers corresponding to the intersection indices on the moduli space of the complex structures. The corresponding theory is called *topological* gravity.[31,32,34,40]

In the case of *physical* ($c<1$ or pq) gravity, the number of explicit expressions known is much smaller. These theories correspond to the nontrivial spectral curves $\Sigma_{g=(p-1)(q-1)/2}$ (Ref. 121).[18)] 2D topological theories with nontrivial spectral curves corresponding to space-time have been constructed formally in Refs. 36 and 37.

From now on (see also Ref. 90), we shall make a detailed study of the more interesting example of the nonperturbative solutions associated with the appearance of nontrivial complex curves of higher genus—in the effective Witten–Seiberg solutions[41,42] of 4D $\mathcal{N}\geqslant 2$ supersymmetric gauge field theories with the bare Lagrangian

$$\mathcal{L}=\int d^4\vartheta F(\Phi_i)=\ldots\frac{1}{g^2}\operatorname{Tr}\mathbf{F}_{\mu\nu}^2+i\theta\operatorname{Tr}\mathbf{F}_{\mu\nu}\tilde{\mathbf{F}}_{\mu\nu}+\ldots \tag{4.1}$$

(where the superfield is $\Phi_i=\varphi^i+\vartheta\sigma_{\mu\nu}\bar{\vartheta}G^i_{\mu\nu}+\ldots$). The nonperturbative exact solution is formally defined as the mapping

$$G,\tau,h_k\rightarrow a_i,\ a_i^D,\quad a_i^D=\frac{\partial\mathcal{F}}{\partial a_i} \tag{4.2}$$

(G is the gauge group, τ is the ultraviolet coupling constant, and $h_k=(1/k)\langle\operatorname{Tr}\Phi^k\rangle$ are the vacuum values of the Higgs field), and an elegant description exists in terms of the curve $\Sigma_{g=\operatorname{rank}G}$, for which h_k parametrizes certain (usually hyperelliptic) moduli of the complex structure. The periods

$$a_i=\oint_{A_i}dS,\quad a_i^D=\oint_{B_i}dS \tag{4.3}$$

of the generating meromorphic 1-differential

$$dS=\lambda d\log w=\operatorname{Tr}\mathcal{L}d\log T, \tag{4.4}$$

satisfying the conditions

$$\frac{\partial dS}{\partial h_k}\cong d\omega_k, \tag{4.5}$$

where $d\omega_k$ are holomorphic differentials on Σ, determine the spectrum of massive BPS states

$$M\sim|\mathbf{n}\mathbf{a}+\mathbf{m}\mathbf{a}^D|, \tag{4.6}$$

and the expression $a_i^D=\partial\mathcal{F}/\partial a^i$ determines the prepotential $\mathcal{F}$ (defining the low-energy action; see the Introduction, where a more detailed explanation of the basic concepts was given) and, thereby, the set of effective coupling constants:

$$T_{ij}=\frac{\partial^2\mathcal{F}}{\partial a_i\partial a_j}=\frac{\partial a_i^D}{\partial a_j}. \tag{4.7}$$

In the present section we shall show that the curves arising in the Witten–Seiberg solutions are the spectral curves of finite-gap solutions of the periodic Toda lattice and its natural deformations into the elliptic Calogero–Moser model and (classical) spin chains.

4.1. $N=2$ supersymmetric gluodynamics and the periodic Toda chain

We begin our analysis with the very simple case of the periodic Toda chain corresponding to the nonperturbative solution of 4D $\mathcal{N}=2$ supersymmetric gluodynamics. The periodic problem in the Toda chain can be formulated in *two* different ways, which are naturally deformed in two different directions. From the physical point of view, the different deformations of the Toda chain correspond to the inclusion of two types of interaction with 4D $\mathcal{N}=2$ supersymmetric gauge theory: with matter in the adjoint and the fundamental representations of the gauge group.

The Toda chain is specified by the equations of motion

$$\frac{\partial\phi_i}{\partial t}=p_i,\quad \frac{\partial p_i}{\partial t}=e^{\phi_{i+1}-\phi_i}-e^{\phi_i-\phi_{i-1}}, \tag{4.8}$$

where for the periodic problem (with period N_c in the "number" of particles) the conditions $\phi_{i+N_c}=\phi_i$ and $p_{i+N_c}=p_i$ are imposed. The Toda chain is a completely integrable system with N_c mutually commuting (as Poisson brackets) Hamiltonians, $h_1^{\mathrm{TC}}=\Sigma p_i$, $h_2^{\mathrm{PC}}=\Sigma(\frac{1}{2}p_i^2+e^{\phi_i-\phi_{i-1}})$, etc. Like any finite-gap solution, the periodic problem in the Toda chain can be described in terms of the eigenvalues and eigenfunctions of two operators: the Lax operator $\mathcal{L}$ [or the auxiliary linear problem for (4.8)],

$$\lambda\psi_n^{\pm}=\sum_k\mathcal{L}_{nk}\psi_k^{\pm}=e^{1/2(\phi_{n+1}-\phi_n)}\psi_{n+1}^{\pm}+p_n\psi_n^{\pm}+e^{1/2(\phi_n-\phi_{n-1})}\psi_{n-1}^{\pm}\left(=\pm\frac{\partial}{\partial t}\psi_n^{\pm}\right), \tag{4.9}$$

and the *monodromy* operator (or boundary conditions), which in this case is simply a shift in the discrete variable corresponding to the number of particles: $T\phi_n=\phi_{n+N_c}$, $Tp_n=p_{n+N_c}$, $T\psi_n=\psi_{n+N_c}$. The condition for the spectra of these two operators to be compatible,

$$\mathcal{L}\psi=\lambda\psi,\quad T\psi=w\psi,\quad [\mathcal{L},T]=0, \tag{4.10}$$

implies that the relation $\mathcal{P}(\mathcal{L},T)=0$ exists between them. It can be stated rigorously in terms of the spectral curve Σ: $\mathcal{P}(\lambda,w)=0$ (Refs. 50 and 53). The generating function of the integrals of the motion can be written in terms of the operators $\mathcal{L}$ and T, and for the Toda chain there are two different formulations of this type.

In the first version, which can be viewed as a limiting case of Hitchin systems,[96] the Lax operator (4.9) is written in the basis of eigenfunctions of the T operator. For a chain of length N_c it becomes an $N_c\times N_c$ matrix

$$\mathcal{L}^{\rm TC}(w)=\begin{pmatrix} p_1 & e^{(1/2)(\phi_2-\phi_1)} & 0 & & we^{(1/2)(\phi_1-\phi_{N_c})} \\ e^{(1/2)(\phi_2-\phi_1)} & p_2 & e^{(1/2)(\phi_3-\phi_2)} & \dots & 0 \\ 0 & e^{(1/2)(\phi_3-\phi_2)} & p_3 & & 0 \\ & & \dots & & \\ \frac{1}{w}e^{(1/2)(\phi_1-\phi_{N_c})} & 0 & 0 & & p_{N_c} \end{pmatrix} \tag{4.11}$$

defined on the cylinder. The Poisson brackets $\{p_i,\phi_j\} = \delta_{ij}$ are equivalent to the Poisson relation on the Lax operator $\{\mathcal{L}^{\rm TC}(w) \otimes \mathcal{L}^{\rm TC}(w')\} = [\mathcal{R}(w,w'), \mathcal{L}^{\rm TC}(w)\otimes \mathbf{1}+\mathbf{1} \otimes \mathcal{L}^{\rm TC}\times(w')]$ with the numerical trigonometric $\mathcal{R}$ matrix

$$\mathcal{R}(w,w')=\frac{w\Sigma(\delta_{i,i+1}\otimes\delta_{i+1,i})+(w'\Sigma\delta_{i+1,i}\otimes\delta_{i,i+1})}{w-w'}, \tag{4.12}$$

and the eigenvalues of the Lax operator determined by the spectral equation

$$\mathcal{P}(\lambda,w)=\det_{N_c\times N_c}(\mathcal{L}^{\rm TC}(w)-\lambda)=0 \tag{4.13}$$

commute with each other as Poisson brackets. Substituting the explicit expression (4.11) into (4.13), we find[122]

$$w+\frac{1}{w}=2P_{N_c}(\lambda) \tag{4.14}$$

or

$$y^2=P^2_{N_c}(\lambda)-1,\quad 2y=w-\frac{1}{w}, \tag{4.15}$$

where $P_{N_c}(\lambda)$ is a polynomial of degree N_c, whose coefficients are the Schur polynomials $S_j(h)$ from the Hamiltonians $h_k=\Sigma_{i=1}^{N_c}p_i^k+\dots$:

$$P_{N_c}(\lambda)=\sum_{k=0}^{N_c}S_{N_c-k}(h)\lambda^k=\left(\lambda^{N_c}+h_1\lambda^{N_c-1}+\frac{1}{2}(h_2-h_1^2)\lambda^{N_c-2}+\dots\right). \tag{4.16}$$

The spectral equation depends only on mutually commuting combinations of dynamical variables—Hamiltonians or action variables parametrizing the subspace in the moduli space of the complex structures of hyperelliptic curves $\Sigma^{\rm TC}$ of genus $N_c-1=\operatorname{rank}SU(N_c)$.

An alternative description of the same system arises if we explicitly *solve* the auxiliary linear problem (4.9), which is a second-order difference equation. It can be solved by simply rewriting it in the form $\tilde{\psi}_{i+1}=L_i^{\rm TC}(\lambda)\tilde{\psi}_i$, i.e., by means of a chain of 2×2 "Lax matrices,"[54] having (after a simple "gauge transformation") the form

$$L_i^{\rm TC}(\lambda)=\begin{pmatrix} p_i+\lambda & e^{\phi_i} \\ e^{-\phi_i} & 0\end{pmatrix},\quad i=1,\dots,N_c. \tag{4.17}$$

These matrices satisfy a *quadratic* Poisson relation of the r-matrix type:[55]

$$\{L_i^{\rm TC}(\lambda)\otimes L_j^{\rm TC}(\lambda')\}=\delta_{ij}[r(\lambda-\lambda'),L_i^{\rm TC}(\lambda)\otimes L_j^{\rm TC}(\lambda')]. \tag{4.18}$$

with a numerical rational r matrix (independent of the label i!) satisfying the classical Yang–Baxter equation $r(\lambda) = (1/\lambda)\Sigma_{a=1}^3\sigma_a\otimes\sigma^a$. Consequently, the monodromy matrix (usually defined for an inhomogeneous lattice with inhomogeneities λ_i)

$$T_{N_c}(\lambda)=\prod_{N_c\geqslant i\geqslant 1}L_i(\lambda-\lambda_i) \tag{4.19}$$

satisfies the same relation

$$\{T_{N_c}(\lambda)\otimes T_{N_c}(\lambda')\}=[r(\lambda-\lambda'),T_{N_c}(\lambda)\otimes T_{N_c}(\lambda')], \tag{4.20}$$

and the integrals of the motion of the Toda chain are generated by a spectral equation in a different form:

$$\det_{2\times 2}(T_{N_c}^{\rm TC}(\lambda)-w)=w^2-w\operatorname{Tr}T_{N_c}^{\rm TC}(\lambda)+\det T_{N_c}^{\rm TC}(\lambda) = w^2-w\operatorname{Tr}T_{N_c}^{\rm TC}(\lambda)+1=0 \tag{4.21}$$

or

$$\mathcal{P}(\lambda,w)=w+\frac{1}{w}-\operatorname{Tr}T_{N_c}^{\rm TC}(\lambda)=w+\frac{1}{w}-2P_{N_c}(\lambda)=0. \tag{4.22}$$

Here we have used the fact that $\det_{2\times 2}L^{\rm TC}(\lambda)=1$ leads to $\det_{2\times 2}T_{N_c}^{\rm TC}(\lambda)=1$. The right-hand side of (4.22) is a polynomial in λ of degree N_c, whose coefficients are integrals of the motion, since

$$\{\operatorname{Tr}T_{N_c}(\lambda),\operatorname{Tr}T_{N_c}(\lambda')\}=\operatorname{Tr}\{T_{N_c}(\lambda)\otimes T_{N_c}(\lambda')\} =\operatorname{Tr}[r(\lambda-\lambda'),T_{N_c}(\lambda)\otimes T_{N_c}(\lambda')]=0. \tag{4.23}$$

For the special case of L matrices (4.17), the inhomogeneities of the chain λ_i reduce to a trivial shift of the momenta: $p_i\to p_i-\lambda_i$.

Below, we shall consider possible elliptic deformations of the two different Lax representations of the Toda chain. Deformation of the $N_c\times N_c$ representation leads to the Calogero–Moser model, while deformation of the 2×2 representation leads to the XYZ model and the Sklyanin algebra.

In addition to the equation of the spectral curve (4.13), (4.14), (4.15), and (4.22), the minimal set of data for determining the Toda chain is given by the generating

1-differential dS^{TC}. The properties of this differential will be studied in detail in Ref. 90; here we only note that the equation $dS^{TC}=\lambda(dw/w)$ in the case of the Toda chain has the literal form (4.4), where λ is the hyperelliptic coordinate in the representations (4.14) and (4.15). The periods of this differential,

$$\mathbf{a}=\oint_{\mathbf{A}} dS^{TC}=\oint_{\mathbf{A}}\lambda\frac{dw}{w},\quad \mathbf{a}_D=\oint_{\mathbf{B}} dS^{TC}=\oint_{\mathbf{B}}\lambda\frac{dw}{w},\tag{4.24}$$

specify the massive BPS spectrum (4.6) and determine the prepotential and the effective coupling constants (4.7) in the low-energy limit of $\mathcal{N}=2$ supersymmetric gluodynamics—the effective $\mathcal{N}=2$ supersymmetric theory with the (Abelian) gauge group $U(1)^{\mathrm{rank}}G$.

4.2. Elliptic deformation of the $N_c\times N_c$ representation: the Calogero–Moser model and interaction with adjoint matter

The $N_c\times N_c$ Lax matrix operator for the $GL(N_c)$ Calogero system explicitly depending on the spectral parameter has the form[123]

$$\mathcal{L}^{\mathrm{cal}}(\xi)=\left(\mathbf{pH}+\sum_\alpha F(\mathbf{q}\alpha|\xi)E_\alpha\right)$$

$$=\begin{pmatrix} p_1 & F(q_1-q_2|\xi) & \cdots & F(q_1-q_{N_c}|\xi)\\ F(q_2-q_1|\xi) & p_2 & \cdots & F(q_2-q_{N_c}|\xi)\\ & & \cdots & \\ F(q_{N_c}-q_1|\xi) & F(q_{N_c}-q_2|\xi) & \cdots & p_{N_c}\end{pmatrix}.\tag{4.25}$$

Its matrix elements $F(q|\xi)=(g/\omega)[\sigma(q+\xi)/\sigma(q)\sigma(\xi)]e^{\zeta(q)\xi}$ can be expressed in terms of Weierstrass elliptic functions, i.e., the Lax operator $\mathcal{L}(\xi)$ is defined on the elliptic curve $E(\tau)$ (a complex torus with periods ω, ω' and modulus $\tau=\omega'/\omega$). From the viewpoint of the 4D interpretation, the interaction constant in the Calogero system, $g^2/\omega^2\sim m^2$, is expressed in terms of the mass m of the adjoint $\mathcal{N}=2$ supermultiplet, breaking the $\mathcal{N}=4$ supersymmetry down to $\mathcal{N}=2$ (Ref. 69).

It follows directly from (4.25) that the spectral curve Σ^{Cal} for the $GL(N_c)$ Calogero system has the form

$$\det_{N_c\times N_c}(\mathcal{L}^{\mathrm{Cal}}(\xi)-\lambda)=0,\tag{4.26}$$

and the masses of the BPS states (4.6) ($\mathbf{a}$ and $\mathbf{a}_D$) are specified by the periods of the generating 1-differential

$$dS^{\mathrm{cal}}\cong\lambda d\xi\tag{4.27}$$

associated with incontractible contours on Σ^{Cal}. The integrability of the Calogero–Moser model can be described in the language of the Poisson structure

$$\{\mathcal{L}(\xi)\otimes\mathcal{L}(\xi')\}=[\mathcal{R}^{\mathrm{Cal}}_{12}(\xi,\xi'),\ \mathcal{L}(\xi)\otimes 1]-[\mathcal{R}^{\mathrm{Cal}}_{21}(\xi,\xi'),\ 1\otimes\mathcal{L}(\xi')]\tag{4.28}$$

determined by the *dynamical* elliptic $\mathcal{R}$ matrix[124] ensuring the involution of the eigenvalues of the matrix $\mathcal{L}$.

The periodic Toda chain is obtained from the Calogero elliptic model in the special double scaling limit,[125] where $g\sim m\to\infty$, $-i\tau\to\infty$, and $q_i-q_j=\frac{1}{2}[(i-j)\log g+(\phi_i-\phi_j)]$, such that the dimensionless coupling constant τ becomes the dimensional parameter $\Lambda^{N_c}\sim m^{N_c}e^{i\pi\tau}$. In this limit the elliptic curve $E(\tau)$ degenerates into a cylinder with coordinate $w=e^{\xi}e^{i\pi\tau}$, and the generating 1-differential $dS^{\mathrm{Cal}}\to dS^{TC}\cong\lambda(dw/w)$ goes into the generating differential of the Toda chain. The Lax operator of the Calogero system becomes the Lax operator (4.11): $\mathcal{L}^{\mathrm{Cal}}(\xi)d\xi\to\mathcal{L}^{TC}(w)(dw/w)$, and the spectral curve acquires the form (4.13). In contrast to the Toda chain, Eq. (4.26) *cannot* be rewritten in the form (4.14), i.e., the specific w dependence of the spectral equation (4.13) is not preserved when the Toda chain is embedded in the system of Calogero–Moser particles. However, the form (4.14) is preserved in a natural way when the Toda chain is interpreted as a special case of spin models.

To describe the other elliptic deformation, we shall use a nonstandard normalization of the Weierstrass $\wp$ function:

$$\wp(\xi|\tau)=\sum_{m,n=-\infty}^{+\infty}\frac{1}{(\xi+m+n\tau)^2}-\sum_{m,n=-\infty}^{+\infty}\frac{1}{(m+n\tau)^2},\tag{4.29}$$

which is doubly periodic in ξ with periods 1 and $\tau=\omega'/\omega$ (this differs from the standard definition by a factor of ω^{-2} and the redefinition $\xi\to\omega\xi$). According to (4.29), the values of $\wp(\xi|\tau)$ at the half-periods, $e_a=e_a(\tau)$, $a=1,2,3$, are functions of only the modular parameter τ, also differing by a factor of ω^{-2} from the standard definition. The complex torus $E(\tau)$ can be defined as the factor $\mathbf{C}/\mathbf{Z}\oplus\tau\mathbf{Z}$ with the "flat" coordinate ξ defined modulo $(1,\tau)$. Moreover, the torus (with a point labeled) can be specified by the elliptic curve

$$y^2=(x-e_1)(x-e_2)(x-e_3),\quad x=\wp(\xi),$$
$$y=\frac{1}{2}\wp'(\xi),\quad d\xi=\frac{dx}{2y}.\tag{4.30}$$

There are three interesting degeneracies of the elliptic picture:

• *The rational limit*: The two periods are $\omega,\omega'\to\infty$, and ξ is redefined as $\xi=\omega^{-1}\zeta$ for finite $\tau=\omega'/\omega$ and ζ. Then

$$x=\wp(\xi)=\frac{\omega^2}{\zeta^2}(1+o(\omega^{-1})),$$
$$y=\frac{1}{2}\wp'(\xi)=-\frac{\omega^3}{\zeta^3}(1+o(\omega^{-1})).\tag{4.31}$$

In the other two limits, $\tau\to+i\infty$, i.e., $q=e^{i\pi\tau}\to0$.

• *The trigonometric limit*: ξ is finite for $q\to0$ and

$$x=\wp(\xi)=-\frac{1}{3}+\frac{1}{\sin^2\pi\xi}+o(q),$$
$$y=\frac{1}{2}\wp'(\xi)=-\pi\frac{\cos\pi\xi}{\sin^3\pi\xi}+o(q).\tag{4.32}$$

• *The double scaling limit*: $\xi=\log(qw)$, the branch points are

$$e_{1,2}\to -\frac{1}{3}\pm 8q+o(q^2),\quad e_3\to +\frac{2}{3}+o(q^2), \tag{4.33}$$

and, in addition,

$$x=\wp(\xi)=-\frac{1}{3}+4q(w-w^{-1})+o(q^2),$$

$$y=\frac{1}{2}\wp'(\xi)=4q(w-w^{-1})+o(q^2), \tag{4.34}$$

so that $d\xi=(dw/w)[1+\mathcal{O}(q)]$. In the simplest example, $N_c=2$ for a curve $\Sigma^{\rm Cal}$ of genus 2. In fact, in this special case (4.26) becomes

$$\mathcal{P}(\lambda,x)=\lambda^2-h_2+\frac{g^2}{\omega^2}x=\lambda^2-h_2+\frac{g^2}{\omega^2}\wp(\xi)=0, \tag{4.35}$$

showing that each value of x corresponds to two points on the curve $\Sigma^{\rm Cal}$, $\lambda=\pm\sqrt{h_2-(g^2/\omega^2)x}$, i.e., $\Sigma^{\rm Cal}$ is specified as a double covering by the elliptic curve $E(\tau)$ with the branch points $x=(\omega/g)^2h_2$ and $x=\infty$. In fact, since the coordinate x itself is elliptic on $E(\tau)$ (when the elliptic curve is treated as a double covering of the sphere CP^1), $x=(\omega/g)^2h_2$ corresponds to a *pair* of points on $E(\tau)$ differing in the sign of y. This would also be true for $x=\infty$, but $x=\infty$ is one of the branch points in the parametrization (4.30) of the curve $E(\tau)$. Thus, the *two* cuts between $x=(\omega/g)^2h_2$ and $x=\infty$ on each of the sheets of $E(\tau)$ effectively coalesce to form a *single* one between the points $((\omega/g)^2h_2,+)$ and $((\omega/g)^2h_2,-)$. The curve $\Sigma^{\rm Cal}$ can therefore be treated as two tori $E(\tau)$ glued together along one cut, i.e., $\Sigma^{\rm Cal}_{N_c=2}$ is a curve of genus 2.

The curve $\Sigma^{\rm Cal}$ is specified analytically for $N_c=2$ by the system of equations (4.30), (4.35), and, accidentally, this curve again turns out to be hyperelliptic (only for $N_c=2$!) after the substitution of x from (4.35) into (4.30).

As the two holomorphic 1-differentials on $\Sigma^{\rm Cal}$ we can take

$$v=\frac{dx}{y}\sim\frac{\lambda d\lambda}{y},\quad V=\frac{dx}{y\lambda}\sim\frac{d\lambda}{y}, \tag{4.36}$$

so that

$$dS\cong\lambda d\xi=\sqrt{h_2-\frac{g^2}{\omega^2}\wp(\xi)}d\xi=\frac{dx}{y}\sqrt{h_2-\frac{g^2}{\omega^2}x}. \tag{4.37}$$

It is easily checked that $\partial dS/\partial h_2\cong\frac{1}{2}(dx/y\lambda)$. The presence of only one of the two holomorphic differentials (4.36) on the right-hand side is related to their different parity under the $\mathbf{Z}_2\otimes\mathbf{Z}_2$ symmetry of $\Sigma^{\rm Cal}$: $y\to -y$ and $\lambda\to-\lambda$. Since dS has definite parity, its periods along two of the four elementary cycles on $\Sigma^{\rm Cal}$ are automatically zero, leaving only the two nontrivial periods a and a_D; this corresponds exactly to the two independent variables in the four-dimensional interpretation. Moreover, the two nonzero periods can be defined in terms of a "reduced" curve of genus 1: $Y^2=(y\lambda)^2=[h_2-(g^2/\omega^2)x]\Pi_{a=1}^3(x-e_a)$, on which $dS\cong[h_2-(g^2/\omega^2)x](dx/Y)$. Since $x=\infty$ is no longer a branch point for it, dS has simple poles at $x=\infty$ (on the two different sheets of $\Sigma^{\rm Cal}_{\rm reduced}$) with residues $\pm g/\omega\sim\pm m$.

In the "opposite" limit of the Calogero–Moser system we have $g^2\sim m^2\to 0$, which corresponds to the $\mathcal{N}=4$ supersymmetric Yang–Mills theory with β function identically equal to zero. The corresponding integrable system is a system of *free* particles, and the generating 1-differential $dS\cong\sqrt{h_2}\cdot d\xi$ is simply a *holomorphic* differential on (the N_c copies of) $E(\tau)$.

4.3. From Toda to spin chains and supersymmetric QCD: the case $N_f<2N_c$ and *XXX* spin models

Let us now turn to the other deformation of the Toda chain corresponding to the inclusion of the interaction with $\mathcal{N}=2$ matter hypermultiplets in the fundamental representation of the gauge group: $\mathcal{N}=2$ supersymmetric QCD. According to Refs. 42 and 76, the spectral curves for $\mathcal{N}=2$ supersymmetric QCD when $N_f<2N_c$ have the same structure as (4.10) with a less trivial monodromy matrix satisfying the conditions

$$\mathrm{Tr}T_{N_c}(\lambda)=2P_{N_c}(\lambda)+R_{N_c-1}(\lambda),$$

$$\det T_{N_c}(\lambda)=Q_{N_f}(\lambda), \tag{4.38}$$

where $Q_{N_f}(\lambda)$ and $R_{N_c-1}(\lambda)$ are polynomials in λ *independent* of h [we recall that for the Toda chain with the Lax matrix (4.17), $\det_{2\times2}T^{\rm TC}_{N_c}(\lambda)=\Pi_{i=1}^{N_c}\det_{2\times2}L_i^{\rm TC}(\lambda-\lambda_i)=1$ and $\mathrm{Tr}T^{\rm TC}_{N_c}(\lambda)=P_{N_c}(\lambda)$]. The two formulations (4.13) and (4.21) are equivalent for the Toda chain, but their *deformations* are different: the representation in the form of a "chain" of 2×2 matrices (4.21) and (4.22) naturally generalizes to the case of the *XYZ* family of spin models.[54,55]

The main idea of deforming the (2×2)-matrix representation is to modify Eqs. (4.18)–(4.23), while preserving the Poisson brackets

$$\{L(\lambda)\otimes L(\lambda')\}=[r(\lambda-\lambda'),\ L(\lambda)\otimes L(\lambda')],$$

$$\{T_{N_c}(\lambda)\otimes T_{N_c}(\lambda')\}=[r(\lambda-\lambda'),\ T_{N_c}(\lambda)\otimes T_{N_c}(\lambda')], \tag{4.39}$$

and thereby the possibility of constructing the monodromy matrix $T(\lambda)$ by taking the product of the matrices $L_i(\lambda)$ over all sites. The equation of the spectral curve for a periodic inhomogeneous spin chain takes the form

$$\det(T_{N_c}(\lambda)-w)=0, \tag{4.40}$$

with the T matrix $T_{N_c}(\lambda)=\Pi^1_{i=N_c}L_i(\lambda-\lambda_i)$, as before satisfying (4.39). The spectral equations can be written more explicitly for $sl(2)$ chains:

$$w+\frac{\det_{2\times2}T_{N_c}(\lambda)}{w}=\mathrm{Tr}_{2\times2}T_{N_c}(\lambda) \tag{4.41}$$

or

$$W+\frac{1}{W}=\frac{\mathrm{Tr}_{2\times2}T_{N_c}(\lambda)}{\sqrt{\det_{2\times2}T_{N_c}(\lambda)}}, \tag{4.42}$$

and the generating 1-differential is now $dS=\lambda(dW/W)$, $W=w/\sqrt{\det T_{N_c}(\lambda)}$. As before, the equations involve the dynamical variables of the spin system only in the form of special combinations—as integrals of the motion and invariants. It is the special form of these equations[74] (the quadratic dependence on the parameters w and W) which allows periodic spin chains to be identified with solutions of the Witten–Seiberg problem with matter in the fundamental representation.

The 2×2 Lax matrix for the $sl(2)$ XXX chain has the form

$$L(\lambda)=\lambda\cdot 1+\sum_{a=1}^{3} S_a\cdot\sigma^a. \tag{4.43}$$

The Poisson brackets of the dynamical variables S_a with $a=1,2,3$ (taking values in the algebra of the functions) follow from (4.39) with rational r matrix:

$$r(\lambda)=\frac{1}{\lambda}\sum_{a=1}^{3}\sigma^a\otimes\sigma^a, \tag{4.44}$$

and, in the case of $sl(2)$, become

$$\{S_a,S_b\}=i\epsilon_{abc}S_c, \tag{4.45}$$

i.e., the $\{S_a\}$ have the meaning of the angular momentum (or the classical spin). The algebra (4.45) contains the Casimir operators (i.e., invariants which have zero Poisson bracket with all the generators S_a) $K^2=\mathbf{S}^2=\Sigma_{a=1}^3 S_aS_a$, so that

$$\det_{2\times 2} L(\lambda)=\lambda^2-K^2,$$

$$\det_{2\times 2} T_{N_c}(\lambda)=\prod_{i=N_c}^{1}\det_{2\times 2} L_i(\lambda-\lambda_i)=\prod_{i=N_c}^{1}((\lambda-\lambda_i)^2-K_i^2)$$

$$=\prod_{i=N_c}^{1}(\lambda+m_i^+)(\lambda+m_i^-)=Q_{2N_c}(\lambda), \tag{4.46}$$

where it is understood that the values of the spin K can differ at different sites along the chain, and[19)]

$$m_i^{\pm}=-\lambda_i\mp K_i. \tag{4.47}$$

At the same time, the determinant (4.46) depends only on the Casimir invariants K_i of the Poisson algebra. The trace of the monodromy matrix $\mathcal{T}_{N_c}(\lambda)=\frac{1}{2}\mathrm{Tr}_{2\times 2}T_{N_c}(\lambda)$ is not an invariant, and, as usual in integrable systems, depends on the variables $S_a^{(i)}$ only through the integrals of the motion, which, in contrast to the Casimir operators, commute only with each other.

In order to obtain a clear representation of the Hamiltonians, in what follows we shall analyze some explicit examples of monodromy matrices for $N_c=2$ and $N_c=3$. The integrals of the motion depend nontrivially on the inhomogeneities of the chain λ_i, and the coefficients of the spectral equation (4.40) depend only on the integrals of the motion and on symmetric functions of the mass parameters m (4.47). This property is important for identifying the parameters m with the masses of the matter hypermultiplets in $\mathcal{N}=2$ supersymmetric QCD. Explicit examples with $N_c=2,3$ have been analyzed in Ref. 77.

4.4. $N_f=2N_c$: Spin chains of general form and the Sklyanin algebra

The construction given above cannot be completed without studying the least well known "elliptic" case $N_f=2N_c$, where the 4D theory is ultraviolet-finite (at least for certain values of the moduli) and possesses an additional *dimensionless* parameter, the UV non-Abelian coupling constant $\tau=8\pi i/e^2+\theta/\pi$.

The most general theory of this type is well known: the XYZ spin chain, in which the elementary L matrix is defined on the elliptic curve $E(\tau)$ and has the form[54,55,126]

$$L^{\mathrm{Skl}}(\xi)=S^0\mathbf{1}+i\sum_{a=i}^{3}W_a(\xi)S^a\sigma_a, \tag{4.48}$$

where

$$W_a(\xi)=\sqrt{e_a-\wp(\xi|\tau)}=i\frac{\theta_{11}'(0)\theta_{a+1}(\xi)}{\theta_{a+1}(0)\theta_{11}(\xi)}. \tag{4.49}$$

The Lax matrix (4.48) satisfies the Poisson bracket (4.18) with the numerical *elliptic* r matrix

$$r(\xi)=i\frac{g}{\omega}\sum_{a=1}^{3}W_a(\xi)\sigma_a\otimes\sigma_a,$$

from which it follows that S^0, S^a form a (classical) Sklyanin algebra:[55,126]

$$\{S^a,S^0\}=2i(e_b-e_c)S^bS^c,\quad \{S^a,S^b\}=2iS^0S^c, \tag{4.50}$$

with the natural notation that abc is the triplet 123 or a cyclic permutation of it.

Accordingly, along with the above limiting (degenerate) cases of an elliptic curve (4.31)–(4.34) we can consider three interesting degeneracies of the Sklyanin algebra.

• *The rational limit.* The two periods are $\omega,\omega'\to\infty$, and (4.50) becomes

$$\{S^a,S^0\}=0,\quad \{S^a,S^b\}=2i\epsilon^{abc}S^0S^c, \tag{4.51}$$

i.e., the generator S^0 becomes a Casimir operator (a constant, for example, $\frac{1}{2}$), and the other S^a form a classical spin algebra (4.45). The corresponding Lax matrix (4.43), $L\equiv\lambda L_{XXX}=\lambda\mathbf{1}+\mathbf{S}\cdot\boldsymbol{\sigma}$, describes the XXX spin model with the rational r matrix (4.44).

• *The trigonometric limit.* For $\tau\to+i\infty$ and $q\to 0$, the Sklyanin algebra (4.50) becomes

$$\{\hat S^3,\hat S^0\}=32iq\hat S^1\hat S^2+\mathcal{O}(q)\to 0,$$

$$\{\hat S^1,\hat S^0\}=-2i\hat S^2\hat S^3+\mathcal{O}(q),$$

$$\{\hat S^2,\hat S^0\}=2i\hat S^3\hat S^1+\mathcal{O}(q),\quad \{\hat S^1,\hat S^2\}=2i\hat S^0\hat S^3+\mathcal{O}(q),$$

$$\{\hat S^1,\hat S^3\}=-2i\hat S^0\hat S^2+\mathcal{O}(q),\quad \{\hat S^2,\hat S^3\}=2i\hat S^0\hat S^1+\mathcal{O}(q). \tag{4.52}$$

The corresponding Lax matrix is

$$L_{XXZ}=\hat S^0\mathbf{1}-\frac{1}{\sin\pi\xi}(\hat S^1\sigma_1+\hat S^2\sigma_2+\cos\pi\xi\hat S^3\sigma_3), \tag{4.53}$$

and the r matrix is

$$r(\xi)=\frac{i}{\sin\pi\xi}(\sigma_1\otimes\sigma_1+\sigma_2\otimes\sigma_2+\cos\pi\xi\sigma_3\otimes\sigma_3). \tag{4.54}$$

• *The double scaling limit.* Using (4.33) and (4.34), we find

$$\sqrt{e_{1,2}-\wp(\xi)}=2\sqrt{q}\left(\sqrt{w}\pm\frac{1}{\sqrt{w}}\right)+\mathcal{O}(q),$$

$$\sqrt{e_3-\wp(\xi)}=1+\mathcal{O}(q), \tag{4.55}$$

and so the Sklyanin algebra (4.50) takes the following form after the redefinition $\hat{S}^{1,2}=(1/4\sqrt{q})\bar{S}^{1,2}$:

$$\{\bar{S}^2,\bar{S}^3\}=2i\bar{S}^0\bar{S}^1+\mathcal{O}(q),\quad \{\bar{S}^1,\bar{S}^0\}=-2i\bar{S}^2\bar{S}^3+\mathcal{O}(q),$$

$$\{\bar{S}^2,\bar{S}^0\}=2i\bar{S}^3\bar{S}^1+\mathcal{O}(q),\quad \{\bar{S}^1,\bar{S}^2\}=32iq\bar{S}^0\bar{S}^3+\mathcal{O}(q)\to 0,$$

$$\{\bar{S}^1,\bar{S}^3\}=-2i\bar{S}^0\bar{S}^2+\mathcal{O}(q),\quad \{\bar{S}^3,\bar{S}^0\}=2i\bar{S}^1\bar{S}^2, \tag{4.56}$$

with the Lax matrix

$$L_{ds}=\bar{S}^0\mathbf{1}+i\bar{S}^3\sigma_3+\frac{i}{2}\left(\sqrt{w}+\frac{1}{\sqrt{w}}\right)\bar{S}^1\sigma_1$$

$$+\frac{i}{2}\left(\sqrt{w}-\frac{1}{\sqrt{w}}\right)\bar{S}^2\sigma_2. \tag{4.57}$$

It is easily seen that (4.53) and (4.57) practically coincide. In particular, the Lax matrix (4.57) satisfies the quadratic Poisson relations (4.18) with the trigonometric r matrix (4.54). In fact, these two Lax matrices are related by the simple transformation $L_{ds}=-\sin(\pi\xi\sigma_2)L_{XXZ}$, where w is identified with $e^{2i\xi}$ and $\bar{S}^0$, $\bar{S}^1$, $\bar{S}^2$, and $\bar{S}^3$ with $\hat{S}^2$, $\hat{S}^3$, $\hat{S}^0$, and $\hat{S}^1$, respectively. We also note that the matrix (4.57) is the L matrix of the sine–Gordon lattice model.

The determinant $\det_{2\times 2}\hat{L}(\xi)$ is given by

$$\det_{2\times 2}\hat{L}(\xi)=\hat{S}_0^2+\sum_{a=1}^{3}e_a\hat{S}_a^2-\wp(\xi)\sum_{a=1}^{3}\hat{S}_a^2$$

$$=K-M^2\wp(\xi)=K-M^2x, \tag{4.58}$$

where

$$K=\hat{S}_0^2+\sum_{a=1}^{3}e_a(\tau)\hat{S}_a^2,\quad M^2=\sum_{a=1}^{3}\hat{S}_a^2 \tag{4.59}$$

are the Casimir operators of the Sklyanin algebra (which commute as Poisson brackets with all the generators $\hat{S}^0$, $\hat{S}^1$, $\hat{S}^2$, and $\hat{S}^3$). The determinant of the monodromy matrix (4.19) in turn is

$$Q(\xi)=\det_{2\times 2}T_{N_c}(\xi)=\prod_{i=1}^{N_c}\det_{2\times 2}\hat{L}(\xi-\xi_i)$$

$$=\prod_{i=1}^{N_c}(K_i-M_i^2\wp(\xi-\xi_i)), \tag{4.60}$$

and its trace $P(\xi)=\frac{1}{2}\mathrm{Tr}T_{N_c}(\xi)$ generates the integrals of the motion, since, as before,

$$\{\mathrm{Tr}T_{N_c}(\xi),\mathrm{Tr}\,T_{N_c}(\xi')\}=0. \tag{4.61}$$

For example, in the case of a *homogeneous* chain [all $\xi_i=0$ in (4.60)], $\mathrm{Tr}T_{N_c}(\xi)$ is a combination of polynomials $P(\xi)=\mathrm{Pol}^{(1)}_{[N_c/2]}(x)+y\mathrm{Pol}^{(2)}_{[(N_c-3)/2]}(x)$, where $[N_c/2]$ is the integer part of $N_c/2$, and the coefficients $\mathrm{Pol}^{(1)}$ and $\mathrm{Pol}^{(2)}$ are integrals of the motion of the XYZ model.[20] As a result, the spectral equation (4.42) for the XYZ chain takes the form

$$w+\frac{Q(\xi)}{w}=2P(\xi), \tag{4.62}$$

where, for a homogeneous chain, P and Q are polynomials in $x=\wp(\xi)$ and $y=\frac{1}{2}\wp'(\xi)$. Equation (4.62) describes double covering by the elliptic curve $E(\tau)$: each point of arbitrary location $\xi\in E(\tau)$ corresponds to two points on Σ^{XYZ}, corresponding to the two roots $w_\pm$ of Eq. (4.62). The branch points are $w_+=w_-=\pm\sqrt{Q}$, or $Y=\frac{1}{2}(w-Q/w)=\sqrt{P^2-Q}=0$.

The curve (4.62) is similar to the spectral curve of the $N_c=2$ Calogero–Moser model (4.35), but the essential difference is that now $x=\infty$ is *no longer* a branch point, and so the total number of cuts on both copies of $E(\tau)$ is N_c, and the genus of the spectral curve is N_c+1.

The curve Σ^{XYZ} can be described analytically by the system of equations

$$y^2=\prod_{a=1}^{3}(x-e_a),\quad Y^2=P^2-Q, \tag{4.63}$$

and the set of holomorphic 1-differentials on Σ^{XYZ} can be chosen as

$$v=\frac{dx}{y},\quad V_\alpha=\frac{x^\alpha dx}{yY},\quad \alpha=0,\ldots,\left[\frac{N_c}{2}\right];$$

$$\tilde{V}_\beta=\frac{x^\beta dx}{Y},\quad \beta=0,\ldots,\left[\frac{N_c-3}{2}\right]. \tag{4.64}$$

The total number of these differentials, $1+([N_c/2]+1)+([(N_c-3)/2]+1)=N_c+1$, is equal to the genus of Σ^{XYZ}.

Finally, once we have the spectral curve, we can try to write down a generating 1-differential dS possessing the defining property (4.5). There are two possible choices for the Toda chain:

$$d\Sigma^{\mathrm{TC}}\cong d\lambda\log w,\quad dS^{\mathrm{TC}}\cong\lambda\frac{dw}{w},$$

$$d\Sigma^{\mathrm{TC}}=-dS^{\mathrm{TC}}+df^{\mathrm{TC}}. \tag{4.65}$$

The two differentials $d\Sigma^{\mathrm{TC}}$ and dS^{TC} satisfy the condition (4.5), and the function f^{TC} is defined so that its variation $\delta f^{\mathrm{TC}}=\lambda(\delta w/w)$ is a (meromorphic) single-valued function on Σ^{TC}.

In the case of the XXX model, the properties of the generating differential are practically the same as those of (4.65):

$$d\Sigma^{XXX}\cong d\lambda\log W,\quad dS^{XXX}\cong\lambda\frac{dW}{W},$$

$$d\Sigma^{XXX}=-dS^{XXX}+df^{XXX}. \tag{4.66}$$

For the XYZ model (4.62) the generating differential(s) dS^{XYZ} can be defined as

$$d\Sigma^{XYZ} \cong d\xi \cdot \log W,$$

$$dS^{XYZ} \cong \xi \frac{dW}{W} = -d\Sigma^{XYZ} + d(\xi \log W). \tag{4.67}$$

Now, varying with respect to the moduli (which are all contained in P), we have

$$\delta(d\Sigma^{XYZ}) \cong \frac{\delta W}{W} d\xi = \frac{\delta P(\xi)}{\sqrt{P(\xi)^2 - Q(\xi)}} d\xi = \frac{dx}{yY} \delta P, \tag{4.68}$$

and according to (4.4) the right-hand side is a *holomorphic* 1-differential on the spectral curve (4.62).

The singularities of $d\Sigma^{XYZ}$ are located at the points $W=0$ or $W=\infty$, i.e., at the zeros of $Q(\xi)$ or the poles of $P(\xi)$. Near the singular points, $d\Sigma^{XYZ}$ is not single-valued, i.e., in making a circuit around a singular point it acquires the additional term $2\pi i d\xi$. The difference between $d\Sigma$ and dS is again a total derivative, but its variation $\delta f^{XYZ} = \xi(\delta W/W)$ is no longer single-valued on the curve. In contrast to $d\Sigma^{XYZ}$, dS^{XYZ} has simple poles at $W=0,\infty$ with residues $\xi|_{W=0,\infty}$ defined modulo $(1,\tau)$. Moreover, the differential dS^{XYZ} itself is not single-valued: it changes by $(1,\tau)\times(dW/W)$ in going around incontractible contours on $E(\tau)$.

Thus, neither $d\Sigma^{XYZ}$ nor dS^{XYZ} is a Seiberg–Witten 1-form in the literal sense; the latter must have well defined residues corresponding to the hypermultiplet masses.[42]

In the simplest example $N_c=2$, the second equation in (4.63) has the form

$$Y^2 = P^2 - Q = (H_0 - H_2 x)^2 - (K_1 - M_1^2 x)(K_2 - M_2^2 x)$$
$$\equiv A(x-x_1)(x-x_2) \tag{4.69}$$

and is a curve of genus $N_c+1=3$, obtained by gluing the two copies of $E(\tau)$ together along the two cuts between $x=x_1$ and $x=x_2$ on each sheet of $E(\tau)$. In Eq. (4.69),

$$H_0 = \hat{S}_1^0 \hat{S}_2^0 + \sum_{a=1}^{3} e_a \hat{S}_1^a \hat{S}_2^a, \quad H_2 = \sum_{a=1}^{3} \hat{S}_1^a \hat{S}_2^a, \tag{4.70}$$

and, comparing with (4.59), it is natural to assume that $H_2 = M_1 M_2 \cos h$. This splitting of the dependence on the Casimirs (M) and the moduli (h) is based on the study of different limits: the conformal limit, where all $M_i \to 0$, and the "dimensional transmutation" limit, where $M_i \to \infty$ and $\tau \to +i\infty$.

When $\tau \to +i\infty$ or $q = e^{i\pi\tau} \to 0$, the branch points e_1 and e_2 move toward each other: $e_1 - e_2 = 16q + \mathcal{O}(q^3)$, and the correct coordinates on Σ^{XYZ} become $x = -\frac{1}{3} + q\check{x}$, $y = q\check{y}$. Then Eq. (4.30) for $E(\tau)$ becomes $\check{y}^2 = \check{x}^2 - 1$ and describes the double covering of CP^1, which is again CP^1. The canonical holomorphic 1-form $d\xi = 2(dx/y)$ becomes a meromorphic differential on CP^1: $2(d\check{x}/\check{y}) = 2(d\check{x}/\sqrt{\check{x}^2-1}) = 2(dz/z)$, where $\check{x} = z + z^{-1}$.

The double scaling limit assumes that the branch points x_1 and x_2 also behave in a special way for $q \to 0$. Let $x_i = -\frac{1}{3} + q\check{x}_i$. Then, redefining $Y = q\check{Y}$, for Σ^{XYZ} in the double scaling limit we obtain $\check{y}^2 = \check{x}^2 - 1$, $\check{Y}^2 = A(\check{x} - \check{x}_1)(\check{x} - \check{x}_2)$. These equations describe two copies of CP^1 glued together along the two cuts (between $\check{x} = \check{A}_1$ and $\check{x} = \check{A}_2$ on each sheet), i.e., an elliptic curve of genus 1. The generating 1-differential is

$$d\Sigma^{XYZ} \cong d\xi \cdot \log W \to d\Sigma^{\mathrm{TC}} \cong \frac{dz}{z} \log W. \tag{4.71}$$

For higher N_c the multi-scaling limit can be taken in a similar manner, i.e., assuming that a spectral curve of genus N_c+1, the double covering $E(\tau)$, degenerates into a double covering of CP^1 of genus $N_c - 1$, associated with the Toda chain. The generating differentials $d\Sigma^{XYZ}$ and dS^{XYZ} also become the corresponding 1-forms (4.65).

Thus, it is natural to assume that the XYZ chain, which is an elliptic generalization of the XXX chain describing $\mathcal{N}=2$ supersymmetric QCD with $N_f < 2N_c$, can be related to the case $N_f = 2N_c$ (see, for example, Ref. 127, where this theory is studied in more detail in the context of multidimensional generalizations of the Seiberg–Witten solutions).

5. CONCLUSION

In this review we have attempted to explain how integrable systems like the KP hierarchy or the Toda chain arise in the description of exact nonperturbative effects in the simplest models of string theory and supersymmetric gauge field theory. We have shown that, as in models of nonperturbative two-dimensional quantum and topological gravity in four-dimensional extended-supersymmetric non-Abelian gauge theory, the effective action of light fields can be specified by the logarithm of the τ function of well known integrable systems of the Whitham type associated with Toda chains.

More specifically, the BRS spectrum and the low-energy effective action are defined in terms of an auxiliary Riemann surface—the spectral surface of an integrable system (in the case of the two-dimensional models studied in detail in the first two sections of this review, the corresponding spectral surface is simply the Riemann complex sphere with labeled points) and the generating 1-form. We have studied in detail the Riemann surfaces arising in the description of the nonperturbative behavior in $\mathcal{N}=2$ supersymmetric gluodynamics, and also in $\mathcal{N}=2$ supersymmetric QCD and the theory with matter in the adjoint representation—the $\mathcal{N}=4$ supersymmetric Yang–Mills field theory broken down to $\mathcal{N}=2$, and the integrable systems associated with these surfaces.

In the second part of this review[90] we shall discuss the finer points of the formulation of supersymmetric gauge theories in terms of integrable systems, and we refer the interested reader to that study. The questions discussed include the properties of the generating differential, the explicit associativity equations which the effective action satisfies in the case of "higher-order" gauge groups, and also the relation of the exact Seiberg–Witten solutions and the integrable systems corresponding to them to the general ideas of modern nonperturbative string theory (M theory). These are the problems of greatest interest from the viewpoint of modern elementary-particle theory.

The author is grateful to his teacher V.Ya. Faĭnberg and his coauthors A.A. Gerasimov, A.S. Gorskiĭ, A.V. Zabrodin, I.M. Krichever, A.M. Levin, Yu.M. Makeenko, A.D. Mironov, A.Yu. Morozov, M.A. Ol'shanetskiĭ, A.Yu. Orlov, V.N. Rubtsov, and S.M. Kharchev, the results of studies with whom formed the basis of this review. I would also like to thank J. Ambjorn, I.A. Batalin, D.V. Bulatov, A.I. Vaĭnshteĭn, B.L. Voronov, C. Vafa, A.V. Gurevich, B.A. Dubrovin, D.R. Lebedev, A.S. Losev, V.V. Losyakov, N.A. Nekrasov, S.P. Novikov, S.Z. Pakulyak, I.V. Polyubin, A.A. Roslyĭ, A. Sagnotti, N.A. Slavnov, A.V. Smilga, M.A. Solov'ev, I.V. Tyutin, L.D. Fadeev, V.V. Fok, D. Fong, S.M. Khoroshkin, J. Schwarz, J. Schnittger, and A.I. Yung for useful discussions. I would particularly like to thank A.P. Isaev for interesting discussions and the suggestion to write this review. This work was performed with the support of Grants RFFI-98-01-00344 and INTAS-96-482.

1)In addition to a parametric dependence on the moduli, physical quantities may depend on the topological (discrete) characteristics of the moduli spaces. Moreover, in the simplest topological string models it is this (and only this) dependence which is important, i.e., the correlation functions are numbers.

2)BPS (Bogomol'nyĭ–Prasad–Sommerfeld) states are states whose masses are proportional to the central charges of the extended $\mathcal{N}\geqslant 2$ supersymmetry algebra.

3)Here we give only a crude picture of finite-gap integration, just to explain the statements made in the text. The details of the mathematical statements can be found in Refs. 50–53.

4)$\mathcal{K}$ differs from $\tilde{\mathcal{K}}$ by the factor $\Pi_{i=1}^{3g-3}|\int_{d^2z}\mu_i b|^2$, where μ_i are $(-1,1)$ Beltrami differentials, related to the specific choice of coordinates y_i on the moduli space $\mathcal{M}_g$.

5)Moreover, $\mathcal{G}[SU(2)]$ is the algebra of the derivatives of the vacuum ring formed by another part of the algebra of the observables: the physical vertex operators of zero dimension and zero ghost number,[100,101] which in fact is isomorphic to the ring of Hamiltonians (polynomials on $\mathbf{R}^2\sim\mathbf{C}$).

6)Here x and $\bar{x}$ denote the holomorphic and antiholomorphic parts of X, respectively.

7)This effect clearly indicates the presence of a definite duality between the world sheet and space-time, for which reparametrizations of the world sheet and transformations in coupling-constant space are interchanged.

8)This is a well known limit in the theory of integrable systems corresponding to the coalescence of two singular points (singularities of the Baker–Akhiezer function, and so on) in the Toda theory into a single one which corresponds to the KdV hierarchy with the corresponding redefinition of the times.

9)One reason is the absence of any simplifying selection rules for the zero mode of the scalar field.

10)The proof of invariance under the Virasoro conditions for potentials of a general form is based on the use of integrability and will be given below.

11)The only requirement is that the matrix differential operator act on functions of the variables T_k (apart from the normalization).

12)Moreover, in this case the equation $\partial Z^{\{p\}}/\partial T_{np}=0$ is satisfied literally.

13)As a check, it is fairly easy to verify (see Ref. 34) that the determinant expression (3.2) with *any* set of functions $\{\phi_i(\mu)\}$ satisfies the bilinear Hirota relations.

14)For example, $W(\mu)=\mu^2+t_1$, $Q(\mu)=\mu$, and then

$$\{W,Q\}=\frac{\partial W}{\partial t_1}\frac{\partial Q}{\partial \mu}-\frac{\partial Q}{\partial t_1}\frac{\partial W}{\partial_\mu}=1.$$

15)See Ref. 117 for a discussion of the case $N<0$.

16)We note that the exponential of negative quantities in the normalization does not significantly affect the τ function of the KP hierarchy, since this factor reduces to an exponential of a trivial bilinear form in the times and corresponds to freedom in defining it. Actually, $\tau\sim\det(\exp[\Sigma_k a_k z_j^{-k}]\phi_i(z_j))\sim\Pi_l\exp[\Sigma a_k z_l^{-k}]\det\phi_i(z_j)\sim\exp[\Sigma k a_k T_k]\det\phi_i(z_j)$.

17)This duality between the models with $c=1$ and $c=25$ is most likely related to the well known fact that solutions of the $c=1$ matrix model of Gross and Klebanov are related[120] to the solution of the Penner model[116] by a Legendre transformation.

18)In this case only the *duality* transformation relating the generating functions of dual theories is known. It has the form of a Fourier transformation with exponent $S=\int^\lambda dS$ (3.63).

19)Equation (4.47) shows that in the limit of zero masses $m_i^\pm=0$ the chain becomes *homogeneous* (all $\lambda_i=0$) with zero spins at each site (all $K_i=0$).

20)For an inhomogeneous chain, it is more complicated to construct the explicit expression for the trace. It can be simplified by means of equations of the type

$$\wp(\xi-\xi_i)=\left(\frac{\wp'(\xi)+\wp'(\xi_i)}{\wp(\xi)-\wp(\xi_i)}\right)^2-\wp(\xi)-\wp(\xi_i)=4\left(\frac{y+y_i}{x-x_i}\right)^2-x-x_i.$$

1 A. M. Polyakov, *Gauge Theories and Strings* [in Russian] (Landau Institute of Theoretical Physics, 1995).

2 A. M. Polyakov, Phys. Lett. B **59**, 79 (1975); A. A. Belavin and A. M. Polyakov, Pis'ma Zh. Éksp. Teor. Fiz. **22**, 503 (1975) [JETP Lett. **22**, 245 (1975)].

3 A. A. Belavin, A. M. Polyakov, A. S. Schwarz, and Yu. S. Tyupkin, Phys. Lett. B **59**, 85 (1975); G. 't Hooft, Phys. Rev. D **14**, 3432 (1976); M. Atiyah, V. Drinfeld, N. Hitchin, and Yu. Manin, Phys. Lett. A **65**, 185 (1978).

4 R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976); C. Callan, R. Dashen, and D. Gross, Phys. Rev. D **19**, 1826 (1979).

5 A. I. Vaĭnshteĭn, V. I. Zakharov, V. A. Novikov, and M. A. Shifman, Usp. Fiz. Nauk **136**, 553 (1982) [Sov. Phys. Usp. **25**, 195 (1982)]; Fiz. Élem. Chastits At. Yadra **17**, 472 (1986) [Sov. J. Part. Nucl. **17**, 204 (1986)].

6 M. Atiyah and R. Ward, Commun. Math. Phys. **55**, 117 (1977); M. Atiyah, N. Hitchin, and I. Singer, Proc. R. Soc. London, Ser. A **362**, 425 (1978).

7 L. D. Faddeev and A. A. Slavnov, *Gauge Fields: Introduction to Quantum Theory* (Addison-Wesley, Redwood City, Calif., 1990) [Russ. original, 2nd ed., Nauka, Moscow, 1988].

8 Yu. A. Gol'fand and E. P. Likhtman, Pis'ma Zh. Éksp. Teor. Fiz. **13**, 452 (1971) [JETP Lett. **13**, 323 (1971)]; D. V. Volkov and V. P. Akulov, Phys. Lett. B **46**, 109 (1973); J. Wess and B. Zumino, Nucl. Phys. B **70**, 39 (1974); V. Ogievetskiĭ and L. Mesincescu, Usp. Fiz. Nauk **117**, 637 (1975) [Sov. Phys. Usp. **18**, 960 (1975)]; S. Ferrara, D. Freedman, and P. van Nieuwenhuizen, Phys. Rev. D **13**, 3214 (1976); S. Deser and B. Zumino, Phys. Lett. B **62**, 335 (1976).

9 J. Scherk, in *Geometrical Ideas in Physics* [in Russian] (Mir, Moscow, 1983).

10 G. Veneziano, Nuovo Cimento A **57**, 190 (1968); Y. Nambu, in *Proceedings of the Intern. Conf. on Symmetries and Quark Models*, Wayne State Univ. (Gordon and Breach, New York, 1970), p. 269; T. Goto, Prog. Theor. Phys. **46**, 1560 (1971).

11 J. Scherk and J. Schwarz, Nucl. Phys. B **81**, 118 (1974).

12 A. M. Polyakov, Phys. Lett. B **103**, 207, 211 (1981).

13 A. A. Belavin and V. Knizhnik, Zh. Éksp. Teor. Fiz. **91**, 364 (1986) [Sov. Phys. JETP **64**, 214 (1986)]; V. Knizhnik, Usp. Fiz. Nauk **159**, 401 (1989) [Sov. Phys. Usp. **32**, 945 (1989)].

14 M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory* (Cambridge University Press, Cambridge, 1987) [Russ. transl., Mir, Moscow, 1990].

15 E. Cremmer, S. Ferrara, and J. Scherk, Phys. Lett. B **74**, 48 (1978).

16 A. Font, L. Ibañez, D. Lust, and F. Quevedo, Phys. Lett. B **249**, 35 (1990).

17 J. Schwarz, hep-th/9607201 and references therein.

18 L. Dixon, in *Proceedings of the 1987 ICTP Summer Workshop in High Energy Physics and Cosmology* (World Scientific, Singapore, 1988); W. Lerche, C. Vafa, and N. Warner, Nucl. Phys. B **324**, 427 (1989).

19 A. M. Polyakov, Phys. Lett. A **2**, 893 (1987); V. Knizhnik, A. M. Polyakov, and A. Zamolodchikov, Mod. Phys. Lett. A **3**, 819 (1988).

20 F. David, Mod. Phys. Lett. A **3**, 1651 (1988); J. Distler and H. Kawai, Nucl. Phys. B **231**, 509 (1989).

21 A. Alekseev and S. Shatashvili, Nucl. Phys. B **323**, 719 (1989).

22 J. Ambjörn, B. Durhuus, and J. Frölich, Nucl. Phys. B **257[FS14]**, 433 (1985); F. David, Nucl. Phys. B **257[FS14]**, 45 (1985); V. Kazakov,

Phys. Lett. B **150**, 282 (1985); V. Kazakov, I. Kostov, and A. Migdal, Phys. Lett. B **157**, 295 (1985).

[23] V. Kazakov, Mod. Phys. Lett. A **4**, 2125 (1989); E. Brezin and V. Kazakov, Phys. Lett. B **236**, 144 (1990); M. Douglas and S. Shenker, Nucl. Phys. B **335**, 635 (1990); D. Gross and A. Migdal, Phys. Rev. Lett. **64**, 127 (1990).

[24] M. Douglas, Phys. Lett. B **238**, 176 (1990).

[25] M. Fukuma, H. Kawai, and R. Nakayama, Int. J. Mod. Phys. A **6**, 1385 (1991).

[26] R. Dijkgraaf, H. Verlinde, and E. Verlinde, Nucl. Phys. B **348**, 435 (1991).

[27] A. Gerasimov, A. Marshakov, A. Mironov *et al.*, Nucl. Phys. B **357**, 565 (1991).

[28] Yu. Makeenko, A. Marshakov, A. Mironov, and A. Morozov, Nucl. Phys. B **356**, 574 (1991).

[29] A. Mironov, Private Communication [in press].

[30] J. Ward, Phys. Rev. **77**, 2931 (1950); E. Fradkin, Zh. Éksp. Teor. Fiz. **29**, 288 (1955) [Sov. Phys. JETP **2**, 111 (1956)]; Y. Takahashi, Nuovo Cimento **6**, 370 (1957); A. Slavnov, Teor. Mat. Fiz. **10**, 99 (1972) [Theor. Math. Phys. (USSR)]; **13**, 175 (1972) [Theor. Math. Phys. (USSR)]; J. Taylor, Nucl. Phys. B **33**, 436 (1971).

[31] E. Witten, Surv. Diff. Geom. **1**, 243 (1991); R. Dijkgraaf, E. Verlinde, and H. Verlinde, Nucl. Phys. B **352**, 59 (1991).

[32] E. Witten, Nucl. Phys. B **340**, 281 (1990); R. Dijkgraaf and E. Witten, Nucl. Phys. B **342**, 486 (1990).

[33] A. Marshakov, A. Mironov, and A. Morozov, Phys. Lett. B **274**, 280 (1992).

[34] S. Kharchev, A. Marshakov, A. Mironov *et al.*, Nucl. Phys. B **380**, 181 (1992).

[35] A. Losev, Teor. Mat. Fiz. **95**, 307 (1993) [Theor. Math. Phys. (USSR)]; A. Losev and I. Polyubin, Int. J. Mod. Phys. A **10**, 4161 (1995).

[36] I. Krichever, Funkts. Anal. Prilozh. **22**, 37 (1988) [in Russian]; Commun. Math. Phys. **143**, 415 (1991); Commun. Pure Appl. Math. **47**, 437 (1994); hep-th/9205110.

[37] B. Dubrovin, hep-th/9407018.

[38] C. Becchi, A. Rouet, and R. Stora, Commun. Math. Phys. **42**, 127 (1975); A. Tyutin, Preprint No. 39, Lebedev Institute, Moscow (1975) [in Russian].

[39] I. Batalin and G. Vilkovysky, Phys. Rev. D **28**, 2567 (1983).

[40] M. Kontsevich, Funkts. Anal. Prilozh. **25**, 50 (1991) [in Russian]; Commun. Math. Phys. **147**, 1 (1992).

[41] N. Seiberg and E. Witten, Nucl. Phys. B **426**, 19 (1994).

[42] N. Seiberg and E. Witten, Nucl. Phys. B **431**, 484 (1994).

[43] A. Gorsky, I. Krichever, A. Marshakov *et al.*, Phys. Lett. B **355**, 466 (1995).

[44] A. Marshakov, A. Mironov, and A. Morozov, Phys. Lett. B **389**, 43 (1996).

[45] A. Marshakov, hep-th/9608161; Int. J. Mod. Phys. A **12**, 1607 (1997).

[46] A. Marshakov, A. Mironov, and A. Morozov, hep-th/9701123.

[47] S. Kachru and C. Vafa, hep-th/9505105; S. Ferrara, J. Harvey, A. Strominger, and C. Vafa, hep-th/9505162; A. Klemm, W. Lerche, P. Mayr *et al.*, hep-th/9604034.

[48] L. Lipatov, in *Perturbative QCD* (World Scientific, Singapore, 1989), p. 411; Phys. Lett. B **251**, 284 (1990); Phys. Lett. B **309**, 394 (1993); Pis'ma Zh. Éksp. Teor. Fiz. **59**, 571 (1994) [JETP Lett. **59**, 596 (1994)]; L. Faddeev and G. Korchemsky, Phys. Lett. B **342**, 311 (1994); G. Korchemsky, Nucl. Phys. B **443**, 255 (1995).

[49] V. Turaev and O. Viro, Topology **31**, 865 (1992).

[50] I. Krichever, Funkts. Anal. Prilozh. **11**, 15 (1977) [in Russian]; Usp. Mat. Nauk **32**, 180 (1977) [Russ. Math. Surveys].

[51] S. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov, *Theory of Solitons: the Inverse Scattering Method* (Consultants Bureau, New York, 1984) [Russ. original, Nauka, Moscow, 1980].

[52] B. Dubrovin, Usp. Mat. Nauk **36**, No. 2, 12 (1981) [Russ. Math. Surveys].

[53] B. Dubrovin, I. Krichever, and S. Novikov, *Current Mathematics Problems (All-Union Institute of Scientific and Technical Information), Dynamical Systems* No. 4, 179 (1985) [in Russian].

[54] L. D. Faddeev and L. A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Springer-Verlag, Berlin, 1987) [Russ. original, Nauka, Moscow, 1986].

[55] E. Sklyanin, J. Sov. Math. **47**, 2473 (1989).

[56] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa, Preprint RIMS-358 (1981).

[57] T. Miwa, Proc. Jpn. Acad., Ser. A: Math. Sci. **58**, 9 (1982).

[58] M. Sato, RIMS Kokyuroku **439**, 30 (1981); M. Sato and Y. Sato, Lect. Notes Numer. Appl. Anal. **5**, 259 (1982).

[59] G. Segal and G. Wilson, Publ. I.H.E.S. **61**, 1 (1985); A. Pressley and G. Segal, *Loop Groups* (Clarendon Press, Oxford, 1986) [Russ. transl., Mir, Moscow, 1990].

[60] A. Orlov and E. Schulman, Lett. Math. Phys. **12**, 171 (1986); P. Grinevich and A. Orlov, Preprint, Cornell University, Ithaca, NY (1988); A. Orlov, in *Proceedigns of the Kiev Intern. Workshop*, 1987 (World Scientific, Singapore, 1988), p. 116.

[61] A. Gurevich and L. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **65**, 590 (1973) [Sov. Phys. JETP **38**, 291 (1974)]; Pis'ma Zh. Éksp. Teor. Fiz. **17**, 268 (1973) [JETP Lett. **17**, 193 (1973)]; see also Ref. 51.

[62] B. Dubrovin and S. Novikov, Usp. Mat. Nauk **44**, No. 6, 29 (1989) [Russ. Math. Surveys].

[63] V. Kac and A. Schwarz, Phys. Lett. B **257**, 329 (1991); A. Schwarz, Mod. Phys. Lett. A **6**, 611, 2713 (1991).

[64] A. Marshakov, A. Mironov, A. Morozov, and S. Kharchev, Teor. Mat. Fiz. **95**, 280 (1993) [Theor. Math. Phys. (USSR)].

[65] C. Vafa and E. Witten, Nucl. Phys. B **431**, 3 (1994).

[66] F. Smirnov, *Form Factors in Completely Integrable Models of Quantum Field Theory* (World Scientific, Singapore, 1992); A. Its, A. Izergin, V. Korepin, and N. Slavnov, Phys. Rev. Lett. **70**, 1704 (1993) and references therein; M. Jimbo and T. Miwa, Regional Conf. Series in Math., Vol. 85 (1995); S. Khoroshkin, D. Lebedev, and S. Pakuliak, q-alg/9602030 and references therein.

[67] E. Fradkin and A. Tseytlin, Nucl. Phys. B **261**, 1 (1985).

[68] A. Klemm, W. Lerche, S. Theisen, and S. Yankielowicz, Phys. Lett. B **344**, 169 (1995); P. Argyres and A. Faraggi, Phys. Rev. Lett. **73**, 3931 (1995).

[69] R. Donagi and E. Witten, hep-th/9510101.

[70] E. Martinec, hep-th/9510204.

[71] A. Gorsky and A. Marshakov, Phys. Lett. B **375**, 127 (1996).

[72] E. Martinec and N. Warner, hep-th/9511052.

[73] H. Itoyama and A. Morozov, hep-th/9511126/9512161/9601168.

[74] A. Marshakov, Mod. Phys. Lett. A **11**, 1169 (1996).

[75] S. Ann and S. Nam, hep-th/9603028.

[76] A. Hanany and Y. Oz, hep-th/9505075.

[77] A. Gorsky, A. Marshakov, A. Mironov, and A. Morozov, Phys. Lett. B **380**, 75 (1996).

[78] I. Krichever and D. Phong, hep-th/9604199.

[79] M. Gell-Mann and F. Low, Phys. Rev. **95**, 1300 (1954); N. N. Bogolyubov and D. V. Shirkov, Dokl. Akad. Nauk SSSR **103**, 203, 391 (1955) [in Russian]; Nuovo Cimento **3**, 854 (1956); L. Ovsyannikov, Dokl. Akad. Nauk SSSR **109**, 112 (1956) [in Russian]; C. Callan, Phys. Rev. D **2**, 1542 (1970); H. Symanzik, Comm. Math. Phys. **18**, 227 (1970); N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, 3rd ed. (Wiley, New York, 1980) [Russ. original, 4th ed., Nauka, Moscow, 1984]; A. Vladimirov and D. V. Shirkov, Usp. Fiz. Nauk **129**, 407 (1979) [Sov. Phys. Usp. **22**, 860 (1979)].

[80] A. Belavin, A. Polyakov, and A. Zamolodchikov, Nucl. Phys. B **241**, 333 (1984).

[81] V. Kats, *Infinite-Dimensional Lie Algebras* [in Russian] (Mir, Moscow, 1993).

[82] V. Knizhnik and A. Zamolodchikov, Nucl. Phys. B **247**, 83 (1984).

[83] D. Friedan, E. Martinec, and S. Shenker, Nucl. Phys. B **271**, 93 (1986).

[84] B. Feĭgin and D. Fuks, Funkts. Anal. Prilozh. **16**, 47 (1982) [in Russian].

[85] Vl. Dotsenko and V. Fateev, Nucl. Phys. B **240**, 312 (1984); **251**, 691 (1985).

[86] E. Verlinde and H. Verlinde, Phys. Lett. B **192**, 95 (1987).

[87] A. Gerasimov, A. Marshakov, A. Morozov *et al.*, Int. J. Mod. Phys. A **5**, 2495 (1990).

[88] J. Distler, Nucl. Phys. B **342**, 523 (1990).

[89] A. Gerasimov, A. Marshakov, and A. Morozov, Phys. Lett. B **242**, 345 (1990).

[90] A. Marshakov, Fiz. Élem. Chastits At. Yadra [Phys. Part. Nuclei] (in press).

[91] J. Fay, *Theta Functions on Riemann Surfaces* (Springer-Verlag, New York, 1973).

[92] D. Mumford, *Tata Lectures on Theta* (Birkhäuser, Boston, 1983) [Russ. transl., Mir, Moscow, 1988].

[93] A. Zamolodchikov, Teor. Mat. Fiz. **63**, 347 (1985) [Theor. Math. Phys.

(USSR)]; V. Fateev and A. Zamolodchikov, Nucl. Phys. B **280[FS18]**, 644 (1987).
[94] V. Fateev and S. Lukyanov, Int. J. Mod. Phys. A **3**, 507 (1988); Preprints ITF-88-74R, ITF-88-76R, Institute of Theoretical Physics, Kiev (1988) [in Russian].
[95] V. Drinfel'd and V. Sokolov, *Current Mathematics Problems (All-Union Institute of Scientific and Technical Information)*, Vol. 24 (1984) [in Russian].
[96] N. Hitchin, Duke Math. J. **54**, 91 (1987); Commun. Math. Phys. **131**, 347 (1990).
[97] D. Lebedev and Yu. Manin, Preprint, ITEP, Moscow (1976).
[98] K. Ueno and K. Takasaki, Adv. Studies Pure Math. **4**, 1 (1984).
[99] A. Marshakov and A. Morozov, Zh. Éksp. Teor. Fiz. **97**, 721 (1990) [Sov. Phys. JETP **70**, 403 (1990)].
[100] B. Lian and G. Zuckerman, Phys. Lett. B **254**, 417 (1991); **266**, 21 (1991); P. Bouwknegt, J. McCarthy, and K. Pilch, Preprint CERN-TH-6196/91, CERN, Geneva (1991).
[101] I. Klebanov and A. Polyakov, Mod. Phys. Lett. A **6**, 3273 (1991); E. Witten, Nucl. Phys. B **373**, 187 (1992).
[102] R. Dijkgraaf, E. Verlinde, and H. Verlinde, Commun. Math. Phys. **115**, 649 (1988); P. Ginsparg, Nucl. Phys. B **295[FS21]**, 153 (1988).
[103] A. Morozov and M. Olshanetsky, Nucl. Phys. B **299**, 389 (1988).
[104] A. Morozov and A. Roslyĭ, Yad. Fiz. **49**, 256 (1989) [Sov. J. Nucl. Phys. **49**, 161 (1989)]; J. Cardy, Nucl. Phys. B **324**, 581 (1989); A. Sagnotti, Preprint ROM2F-91/11, University of Rome, Rome (1991) and references therein.
[105] V. Kac, Lect. Notes Phys. **94**, 441 (1979); see also Ref. 81.
[106] D. P. Zhelobenko and A. I. Shtern, *Representations of Lie Groups* [in Russian] (Nauka, Moscow, 1983).
[107] J. Thierry-Mieg, Phys. Lett. B **197**, 368 (1987).
[108] A. Marshakov, A. Mironov, and A. Morozov, Phys. Lett. B **265**, 99 (1991); S. Kharchev, A. Marshakov, A. Mironov *et al.*, Nucl. Phys. B **404**, 717 (1993).
[109] A. Mikhailov, Int. J. Mod. Phys. A **9**, 873 (1994).
[110] C. Itzykson and J.-B. Zuber, J. Math. Phys. **21**, 411 (1980); M. Mehta, Commun. Math. Phys. **79**, 327 (1981).
[111] M. Fukuma, H. Kawai, and R. Nakayama, Commun. Math. Phys. **143**, 371 (1992).
[112] S. Kharchev and A. Marshakov, in *String Theory, Quantum Gravity and the Unification of Fundamental Interactions* (World Scientific, Singapore, 1993), p. 331; Int. J. Mod. Phys. A **10**, 1219 (1995).
[113] T. Shiota, Invent. Math. **83**, 333 (1986); see also Ref. 92; T. Takebe, Preprint RIMS-779 (1991).
[114] K. Takasaki and T. Takebe, Preprint RIMS-814 (1991); Preprint KUCP-0050/92.
[115] M. Fukuma, H. Kawai, and R. Nakayama, Commun. Math. Phys. **148**, 101 (1992).
[116] J. Harer and D. Zagier, Invent. Math. **85**, 457 (1986); R. Penner, Commun. Math. Phys. **113**, 299 (1987); J. Diff. Geom. **27**, 35 (1988); J. Distler and C. Vafa, Mod. Phys. Lett. A **6**, 259 (1991).
[117] S. Kharchev, A. Marshakov, A. Mironov, and A. Morozov, Nucl. Phys. B **397**, 339 (1993).
[118] A. Marshakov, in *Pathways to Fundamental Theories* (World Scientific, Singapore, 1993), p. 251.
[119] R. Dijkgraaf, G. Moore, and R. Plessner, Preprint IASSNS-HEP-92/48 (1992).
[120] D. Gross and I. Klebanov, Nucl. Phys. B **344**, 475 (1990).
[121] I. Krichever, Preprint ETH, Zurich (1990); S. Novikov, Funkts. Anal. Prilozh. **24**, 43 (1990) [in Russian]; G. Moore, Commun. Math. Phys. **133**, 261 (1990); R. Schimmrigk, Phys. Rev. Lett. **65**, 2483 (1990).
[122] I. Krichever, Appendix to Ref. 52.
[123] I. Krichever, Funkts. Anal. Prilozh. **14**, 282 (1980) [in Russian].
[124] E. Sklyanin, Alg. Anal. **6**, 227 (1994); H. Braden and T. Suzuki, Lett. Math. Phys. **30**, 147 (1994); B. Enriquez and V. Rubtsov, alg-geom/9503010; N. Nekrasov, hep-th/9503157; G. Arutyunov and P. Medvedev, hep-th/9511070.
[125] V. Inozemtsev, Commun. Math. Phys. **121**, 629 (1989).
[126] E. Sklyanin, Funkts. Anal. Prilozh. **16**, 27 (1982); **17**, 34 (1983) [in Russian].
[127] A. Marshakov and A. Mironov, Nucl. Phys. B **518**, 59 (1998).

Translated by Patricia A. Millard