

ALGEBRAIC GEOMETRY ASSOCIATED WITH MATRIX MODELS OF TWO DIMENSIONAL GRAVITY

*Edward Witten**

1. Introduction

The purpose of these notes is to describe some conjectures in algebraic geometry associated with matrix models of two dimensional quantum gravity. The original conjecture of this type gave a detailed description of intersection theory on the moduli space of Riemann surfaces, by relating it to the most basic matrix model, which was solved nonperturbatively by several groups [12, 9, 4]. (Many aspects of the solution were later clarified by Neuberger [18].) This original conjecture was described in [21], along with an introduction to the matrix models and detailed references. The original version of the conjecture has since been proved by Kontsevich [14]. (Kontsevich used methods of differential topology—in particular, a certain triangulation of the moduli space of Riemann surfaces developed in [19, 3]. I also discussed the last step of the proof in [23].)

There are more elaborate matrix models that describe two dimensional quantum gravity coupled to matter. Topological field theories associated with these models have been identified by K. Li [17], and their further study has led [22] to a generalization of the original conjecture, involving intersection theory not on moduli space of Riemann surfaces but on certain covers of it that are obtained by extracting n^{th} roots of the canonical line bundle of a surface. Our aim here is to explain this more general conjecture just as a statement in geometry, referring the reader interested in its field theoretic origin to [22].

* Research supported in part by NSF Grant 86-20266.

That field theoretic origin involves twisted $N = 2$ superconformal field theories that are close cousins of the models that are studied in the world of “mirror manifolds” [11]. Their study has yielded a remarkable family of (conjectural) algebrogeometric formulas, counting the rational curves on certain threefolds [5]. The formulas presented in the present paper should also have a mirror version, but this is not yet understood.

Formulas of the type to be presented here actually are expected to have an **A–D–E** classification; the case considered here corresponds to the **A** series. The more general formulas, which have not been determined in full detail, will involve the **D** and **E** singularities (the function $s \rightarrow s'$ implicit in §1.3 corresponds to the A_{r-1} singularity) and the Drinfel’d-Sokolov generalization of the Gelfand-Dikii equations.

1.1 Roots of the canonical line bundle. Let Σ be a smooth Riemann surface of genus g , with s marked points x_1, x_2, \dots, x_s . Fix an integer $r \geq 2$. Label each x_i by an integer m_i with $0 \leq m_i \leq r - 1$. (Eventually, the x_i will be labeled by an additional non-negative integer n_i .)

For each i , let $\mathcal{O}(x_i)$ be the line bundle, of degree 1, whose sections are functions that may have a simple pole at x_i . The canonical line bundle K of Σ has degree $2g - 2$. The line bundle $\mathcal{S} = K \otimes_i \mathcal{O}(x_i)^{-m_i}$ has degree $2g - 2 - \sum_i m_i$. If this is divisible by r , then \mathcal{S} possesses r^{th} roots. Indeed, there are r^{2g} isomorphism classes of line bundle \mathcal{T} such that

$$(1.1.1) \quad \mathcal{T}^{\otimes r} \cong \mathcal{S}.$$

Let $\mathcal{M}_{g,s}$ be the moduli space of complex Riemann surfaces of genus g with s punctures, and $\overline{\mathcal{M}}_{g,s}$ its Deligne-Mumford compactification. The choice of an isomorphism class of \mathcal{T} determines a cover $\mathcal{M}'_{g,s}$ of $\mathcal{M}_{g,s}$; this is an unramified cover of degree r^{2g} . (r and m_i entered in the definition of $\mathcal{M}'_{g,s}$, but will not be indicated explicitly in the notation.)

A line bundle \mathcal{T} endowed with an isomorphism as in equation (1.1.1) has μ_r (the r^{th} roots of unity) as an isomorphism group. As a result, a line bundle \mathcal{T} with such an isomorphism may not exist universally over $\mathcal{M}_{g,s}$. This causes some constructions below to be possible only rationally.

1.2 Behavior near a double point. We now want to explain how we wish to extend this definition to get a cover (ramified at infinity) of the compactified moduli space $\overline{\mathcal{M}}_{g,s}$.

Restricted to a circle $C \subset \Sigma$, there are r isomorphism classes of \mathcal{T} . The compactification divisor in $\overline{\mathcal{M}}_{g,s}$ can be thought of as the locus on which

some C is collapsed to a point. Corresponding to the r possibilities for $\mathcal{T}|_C$, there are r possibilities for the behavior of \mathcal{T} at infinity.

The behavior near a double point can be described by a family of curves X

$$(1.2.1) \quad xy = \epsilon$$

in the x - y plane, parameterized by a complex variable ϵ . For stable curves, the marked points, say $x_i(\epsilon)$, do not coincide with the double point at $x = y = \epsilon = 0$, so they can be neglected in describing the local behavior near the double point. \mathcal{T} can therefore be thought of as just an r^{th} root of the canonical bundle K .

Algebraically, the r possibilities for \mathcal{T} differ by which sections of K have r^{th} roots. $r - 1$ of the possibilities are generalizations of the Neveu-Schwarz sector of string theory, and the remaining one is a generalization of the usual Ramond sector. To describe the first $r - 1$ possibilities, fix an integer m with $0 \leq m \leq r - 2$. Let $c = (m + 1, r)$ and $r' = r/c$. Over the r' -fold cover of the ϵ plane given by $\delta^{r'} = \epsilon$, let \mathcal{T}_m be the sheaf generated by objects u_1 and u_2 with relations

$$(1.2.2) \quad yu_1 = \delta^{(m+1)/c}u_2, \quad xu_2 = \delta^{(r-1-m)/c}u_1.$$

Intuitively, u_1 and u_2 correspond to

$$(1.2.3) \quad u_1 = (dx \cdot x^m)^{1/r}$$

and

$$(1.2.4) \quad u_2 = (-dy \cdot y^{r-2-m})^{1/r}.$$

Thus, on the complement of the double point, \mathcal{T}_m is locally free, and we can define an isomorphism $\psi: \mathcal{T}_m^{\otimes r} \cong K$ with $\psi(u_1^{\otimes r}) = dx \cdot x^m$, $\psi(u_2^{\otimes r}) = -dy \cdot y^{r-2-m}$.

For the r^{th} possibility, we consider a sheaf \mathcal{T}_{r-1} that is freely generated by a section v that we think of intuitively as

$$(1.2.5) \quad v = \left(\frac{dx}{x}\right)^{1/r} = \left(-\frac{dy}{y}\right)^{1/r}.$$

Thus, we endow \mathcal{T}_{r-1} with an isomorphism $\psi: \mathcal{T}_{r-1}^{\otimes r} \cong K$ such that $\psi(v^{\otimes r}) = dx/x$.

By an r^{th} root \mathcal{T} of $\mathcal{S} = K \otimes_i \mathcal{O}(x_i)^{-m_i}$, over the family of curves X , we mean a coherent sheaf \mathcal{T} over X such that on the complement of the double point, \mathcal{T} is locally free and has an isomorphism $\psi: \mathcal{T}^{\otimes r} \cong \mathcal{S}$, and in a neighborhood of the double point, the pair \mathcal{T}, ψ is isomorphic to one of the r possibilities just described.

Thus we can characterize the compactification that we want. It is the moduli space $\bar{\mathcal{M}}'_{g,s}$ of pairs Σ, \mathcal{T} , where Σ is a stable curve with marked points x_i labeled by m_i , and \mathcal{T} is a coherent sheaf that is an r^{th} root of $\mathcal{S} = K \otimes \mathcal{O}(x_i)^{-m_i}$ in the sense just indicated.¹

More detailed behavior at the double point. For computations in section 2, we will need to understand the behavior near the double point in a little more detail. Let Σ_0 be the singular fiber of the family X and $\pi: \bar{\Sigma} \rightarrow \Sigma_0$ its normalization. The inverse image on $\bar{\Sigma}$ of the double point consists of two points P' and P'' . If \mathcal{T} is isomorphic near the double point to one of the \mathcal{T}_m with $m < r - 1$, then $\mathcal{T} \cong \pi_* \mathcal{T}'$ where \mathcal{T}' is a locally free sheaf on $\bar{\Sigma}$ with a natural isomorphism

$$(1.2.6) \quad \psi': \mathcal{T}'^{\otimes r} \cong K \otimes_i \mathcal{O}(x_i)^{-m_i} \otimes \mathcal{O}(P')^{-m} \otimes \mathcal{O}(P'')^{-(r-2-m)}.$$

Thus, on the normalization P' and P'' behave like marked points labeled by integers m and $r - 2 - m$.

Now consider the remaining case in which \mathcal{T} is isomorphic near the double point to \mathcal{T}_{r-1} . Intuitively, \mathcal{T} is the sheaf of $1/r$ differentials on Σ_0 which may have poles of order $1/r$ with equal and opposite residues on the two branches (the residue being the coefficient of $(dx/x)^{1/r}$ or $(-dy/y)^{1/r}$). There is thus a “residue” map $\mathcal{T} \rightarrow \mathcal{O}$, which extracts the residue, and an exact sequence

$$(1.2.7) \quad 0 \rightarrow \mathcal{T}' \rightarrow \mathcal{T} \xrightarrow{\text{Res}} \mathcal{O} \rightarrow 0.$$

\mathcal{T}' is intuitively the sheaf of $1/r$ differentials on Σ_0 with zeroes of order $1 - 1/r$ at the double point. Though \mathcal{T} is not the direct image of a locally free sheaf on the normalization $\bar{\Sigma}$, \mathcal{T}' is such a direct image. Indeed, $\mathcal{T}' = \pi_* \mathcal{T}''$ where \mathcal{T}'' is a line bundle on $\bar{\Sigma}$ with a natural isomorphism

$$(1.2.8) \quad \psi'': \mathcal{T}''^{\otimes r} \cong K \otimes \mathcal{O}(x_i)^{-m_i} \otimes \mathcal{O}(P')^{-(r-1)} \otimes \mathcal{O}(P'')^{-(r-1)}.$$

Thus, in the definition of \mathcal{T}'' , the inverse images of the double point appear as new marked points labeled by $r - 1$.

¹ When $c > 1$, as explained by Deligne, some additional structure is present on the special fibers that is important to avoid unwanted automorphisms. Essentially, near the double point, $\mathcal{T}^{\otimes r'}$ should be mapped, by gluing special fibers, to a line bundle that is a local c^{th} root of K .

1.3 The top Chern class. For Σ, \mathcal{T} as above, let V' be the vector space

$$(1.3.1) \quad V' = H^1(\Sigma, \mathcal{T})$$

and let V be the dual space, which according to Serre duality is $V = H^0(\Sigma, K \otimes \mathcal{T}^{-1})$ if Σ is a smooth curve, or $H^0(\Sigma, \text{Hom}(\mathcal{T}, K))$ in general. Generically, these are vector spaces of dimension

$$(1.3.2) \quad D = (g-1)(1-2\gamma) + \gamma \sum_i m_i,$$

where $\gamma = 1/r$. This fails precisely when

$$(1.3.3) \quad W' = H^0(\Sigma, \mathcal{T})$$

is non-zero.

If W' vanishes everywhere, then V' and V vary as fibers of D dimensional vector bundles \mathcal{V}' and \mathcal{V} over moduli space. (This is actually true in genus zero, as we will see in section 2.1.) These bundles have Chern classes, and in particular top Chern classes $c_D(\mathcal{V}')$ and

$$(1.3.4) \quad c_D(\mathcal{V}) = (-1)^D c_D(\mathcal{V}').$$

We want to describe a substitute for $c_D(\mathcal{V}')$ when W' is not identically zero.¹ Then we will define $c_D(\mathcal{V})$ by (1.3.4). The construction will be analogous to the definition of the index bundle of a family of Fredholm operators [1].

Let E and F be vector bundles of dimensions T and $D+T$ over a compact topological space X . Let $\pi: E \rightarrow X$ be the projection, $i: X \rightarrow E$ the zero section, and π^*F the pullback of F to (the total space of) E . We would like to define $c_D(F, E) = \pi_*(c_{D+T}(\pi^*F))$. This is ill-defined since the fibers of $\pi: E \rightarrow X$ are not compact. But suppose we are given a continuous section $w: E \rightarrow \pi^*F$ which vanishes only on $i(X) \subset E$. As a vector bundle with a nonvanishing section has zero top Chern class, the choice of such a section permits us to define a top Chern class of π^*F in $H_{\text{cpt}}^*(E)$; we call this $c_{D+T}(\pi^*F; w)$. Then we can set

$$(1.3.5) \quad c_D(F, E; w) = \pi_*(c_{D+T}(\pi^*F; w)).$$

¹ A definition of $c_D(\mathcal{V}')$, but not the one we want, can be given by considering the Chern classes of the index bundle of the $\bar{\partial}$ operator (or of the direct image sheaves $R\pi^i(\mathcal{T})$, with $\pi: \bar{\mathcal{C}}'_{g,s} \rightarrow \bar{\mathcal{M}}'_{g,s}$ being the fibering of the universal curve).

Suppose A and B are two vector bundles over X and $\epsilon: A \rightarrow B$ is an isomorphism. If $\tilde{\pi}: A \rightarrow X$ is the projection, ϵ determines a section $\epsilon: A \rightarrow \tilde{\pi}^*B$ which vanishes only on $i(X) \subset A$. On each fiber of $\tilde{\pi}: A \rightarrow X$, the zero of ϵ is a simple zero, of winding number 1. Hence

$$(1.3.6) \quad c_0(A, B; \epsilon) = 1.$$

If in an obvious way we think of $w \oplus \epsilon$ as a section of the pullback of $F \oplus B$ over $E \oplus A$, then

$$(1.3.7) \quad c_D(F \oplus A, E \oplus B; w \oplus \epsilon) = c_D(F, E; w)$$

essentially by the multiplicativity of Chern classes in direct sums.

Returning to our problem, pick a metric on the universal curve $\overline{\mathcal{C}}'_{g,s} \rightarrow \overline{\mathcal{M}}'_{g,s}$, so that it can be regarded as a family of Riemannian manifolds (with some singularities at infinity). In particular, the relative canonical bundle K (whose sections are $(1, 0)$ forms along the fibers) gets a metric. Regarding the marked points x_i , $i = 1 \dots s$, as divisors on $\overline{\mathcal{C}}'_{g,s}$, pick a metric on each $\mathcal{O}(x_i)$. These metrics determine a metric on the line bundle \mathcal{T} which has a defining isomorphism $\mathcal{T}^{\otimes r} \cong K \otimes_i \mathcal{O}(x_i)^{-m_i}$.

At least over the open moduli space $\mathcal{M}'_{g,s}$, one has bundles of Hilbert spaces $\mathcal{E} = \Omega^{0,0}(\mathcal{T})$ and $\mathcal{F} = \Omega^{0,1}(\mathcal{T})$ (consisting respectively of \mathcal{T} -valued $(0, 0)$ and $(0, 1)$ forms along the fibers of the universal curve), with a family of $\bar{\partial}$ operators $\bar{\partial}: \mathcal{E} \rightarrow \mathcal{F}$, and their adjoints $\partial: \mathcal{F} \rightarrow \mathcal{E}$. Seeley and Singer have extended these definitions [20] to get continuous families of Hilbert spaces \mathcal{E} , \mathcal{F} and Fredholm operators $\bar{\partial}$, ∂ over the *compactified* moduli spaces $\overline{\mathcal{M}}'_{g,s}$.¹ Let $\pi: \mathcal{E} \rightarrow \overline{\mathcal{M}}'_{g,s}$ be the projection. We wish, as in the discussion above, to define a “top Chern class” $c_D(\mathcal{F}, \mathcal{E}; w)$, with an appropriate section w . First we will construct w , and then we will deal with the fact that \mathcal{E} and \mathcal{F} are infinite dimensional.

For \mathcal{L} a line bundle, let $\overline{\mathcal{L}}$ be the complex conjugate line bundle. ($\overline{\mathcal{L}}$ has transition functions that are the complex conjugates of those of \mathcal{L} ; more intrinsically a section of $\overline{\mathcal{L}}$ is an anti-linear map $\mathcal{L}^{-1} \rightarrow \mathcal{O}$.) A metric on \mathcal{L} determines an invertible element of $\overline{\mathcal{L}} \otimes \mathcal{L}$. Using the metric on \mathcal{T} and the defining isomorphism of \mathcal{T} , we get

$$(1.3.8) \quad \overline{\mathcal{T}}^{\otimes(r-1)} \cong \overline{\mathcal{T}}^{\otimes r} \otimes \mathcal{T} \cong \overline{K} \otimes \mathcal{T} \otimes_i \overline{\mathcal{O}(x_i)}^{\otimes(-m_i)}.$$

¹ They considered the case $r = 2$ which apparently contains the essential features.

A section f of $\overline{\mathcal{O}(x_i)}^{-1}$ is a complex valued function f which near $x = x_i$ can be written in the form $f = ug$, where u is an antiholomorphic function that vanishes at $x = x_i$ and g is a complex valued smooth function. In particular, such an f is a smooth function. Thus a section of $\overline{K} \otimes \mathcal{T} \otimes_i \overline{\mathcal{O}(x_i)}^{\otimes(-m_i)}$ is a section of $\overline{K} \otimes \mathcal{T}$ that happens to have certain zeroes. Hence for $s \in \mathcal{E} = \Gamma(\mathcal{T})$, we can interpret \bar{s}^{r-1} as a vector in $\mathcal{F} = \Gamma(\overline{K} \otimes \mathcal{T})$.

This enables us to define $w: \mathcal{E} \rightarrow \pi^*\mathcal{F}$ by

$$(1.3.9) \quad w(s) = \bar{\partial}s + \bar{s}^{r-1}.$$

To put this in the general framework discussed earlier, we must show that $w(s) = 0$ only if $s = 0$. Indeed we have

$$(1.3.10) \quad (\bar{\partial}s, \bar{s}^{r-1}) = \int_{\Sigma} \partial\bar{s} \cdot \bar{s}^{r-1} = \frac{1}{r} \int_{\Sigma} \partial(\bar{s}^r) = 0,$$

so

$$(1.3.11) \quad (w, w) = (\bar{\partial}s, \bar{\partial}s) + (\bar{s}^{r-1}, \bar{s}^{r-1}).$$

Hence $w = 0$ only if $0 = \bar{\partial}s = \bar{s}^{r-1}$, and so only if $s = 0$.

Now, it is possible to find a finite dimensional subbundle F of \mathcal{F} which maps surjectively to $\text{coker}(\bar{\partial}) = \mathcal{F}/\bar{\partial}(\mathcal{E})$. This is essentially a step in defining the index bundle of the $\bar{\partial}$ operator [1].¹ Define a finite dimensional subbundle E of \mathcal{E} by $E = \bar{\partial}^{-1}(F)$. Let E^\perp be the orthocomplement of E , and $F^\perp = \bar{\partial}(E^\perp)$. Replacing F by the orthocomplement of F^\perp , which still maps surjectively to $\text{coker}(\bar{\partial})$ (and redefining E to preserve $E = \bar{\partial}^{-1}(F)$), one can assume that F contains $\ker(\bar{\partial})$.

For such a pair (E, F) , let ρ_F be the orthogonal projection $\mathcal{F} \rightarrow F$, let $\pi: E \rightarrow \overline{\mathcal{M}}'_{g,s}$ be the projection, and define $w \in \Gamma(E, \pi^*F)$ by

$$(1.3.12) \quad w(s) = \bar{\partial}s + \rho_F(\bar{s}^{r-1})$$

¹ Over a compact manifold X of real dimension $2d$, there is no obstruction to finding an everywhere non-zero section s of a complex vector bundle \mathcal{F} of rank $> d$. Hence for \mathcal{F} a Hilbert space bundle, one can find a sequence of everywhere orthonormal sections s_1, s_2, \dots . To do this, let F_n be the (trivial) subbundle generated by s_1, \dots, s_n , and at each stage let s_{n+1} be an everywhere nonzero section of F_n^\perp . If $\bar{\partial}: \mathcal{E} \rightarrow \mathcal{F}$ is a continuous family of operators, parameterized by X , of finite dimensional kernel and cokernel, then for each $x \in X$, there is n such that $F_{n,x} \rightarrow \text{coker } \bar{\partial}_x$ is surjective. As this surjectivity is an open condition and X is compact, there is some n that works for all $x \in X$.

As ρ_F is self-adjoint and $\rho_F \bar{\partial}s = \bar{\partial}s$, we have

$$(1.3.13) \quad (\bar{\partial}s, \rho_F \bar{s}^{r-1}) = (\bar{\partial}s, \bar{s}^{r-1}) = 0$$

as before. Hence

$$(1.3.14) \quad (w, w) = (\bar{\partial}s, \bar{\partial}s) + (\rho_F \bar{s}^{r-1}, \rho_F \bar{s}^{r-1}),$$

so $w(s)$ vanishes only if $\bar{\partial}s = \rho_F \bar{s}^{r-1} = 0$. However, if $\bar{\partial}s = 0$, then $\bar{\partial}(\bar{s}^{r-1}) = 0$, so $\bar{s}^{r-1} \in F$ (which contains $\ker(\bar{\partial})$ by construction) and $\rho_F \bar{s}^{r-1} = \bar{s}^{r-1}$. So in fact $w(s) = 0$ only if $s = 0$. Thus we can define

$$(1.3.15) \quad c_D(F, E; w) = \pi_*(c_{D+T}(\pi^* F; w))$$

where $T = \dim F$. Henceforth we frequently abbreviate $c_D(F, E; w)$ as $c_D(F)$, it being understood that $E = \bar{\partial}^{-1}(F)$ and that w is defined as in equation (1.3.9).

It remains to show that $c_D(F') = c_D(F)$ if F' is any other finite dimensional subbundle of \mathcal{F} containing $\ker(\bar{\partial})$. It suffices to consider the case that $F \subset F'$ since up to homotopy any F and F' chosen as above are subbundles of some F'' .¹

For $F \subset F'$, we set $F' = F \oplus \tilde{F}$ with \tilde{F} the orthocomplement of F in F' . Let $E' = \bar{\partial}^{-1}(F')$. Then $E' = E \oplus \tilde{E}$ (a direct sum decomposition, not necessarily orthogonal) where $\tilde{E} = \bar{\partial}^{-1}(\tilde{F}) \cap \ker(\bar{\partial})^\perp$. So $\bar{\partial}(\tilde{E}) = \tilde{F}$ and $\bar{\partial}: \tilde{E} \rightarrow \tilde{F}$ is an isomorphism.

Let $\pi': F' \rightarrow \bar{\mathcal{M}}'_{g,s}$ be the projection. We want to compare two natural elements w', w'' of $\Gamma(E', \pi'^*(F'))$. The first is

$$(1.3.16) \quad w'(s) = \bar{\partial}s + \rho_{F'} \bar{s}^{r-1}$$

with $\rho_{F'}$ the orthogonal projection onto F' . This is our “standard” choice, so $c_D(F') = c_D(F', E'; w')$. But with $F' = F \oplus \tilde{F}$, $E' = E \oplus \tilde{E}$, we can also take the direct sum of the “standard” map $w: E \rightarrow F$ (given by $w(s) = \bar{\partial}s + \rho_F \bar{s}^{r-1}$, as above) and the $\bar{\partial}$ isomorphism $\bar{\partial}: \tilde{E} \rightarrow \tilde{F}$. This is thus

$$(1.3.17) \quad w''(e \oplus \tilde{e}) = w(e) \oplus \bar{\partial}\tilde{e}.$$

¹ The argument is similar to the one in the last footnote. One can construct an ascending family of subbundles $F \subset F_1 \subset F_2 \subset \dots \subset \mathcal{F}$ by picking at each stage a non-zero section of F_n^\perp . Let θ_n be the orthogonal projection on F_n . For $x \in \bar{\mathcal{M}}'_{g,s}$, $\theta_n: F'_x \rightarrow F_{n,x}$ is injective for large enough n ; as this injectivity is an open condition and $\bar{\mathcal{M}}'_{g,s}$ is compact, there is some n such that $\theta_n: F' \rightarrow F_n$ is an embedding. For $0 \leq t \leq 1$, let $F'_t = (1-t + t\theta_n)F'$. By continuity $c_D(F'_t)$ is independent of t . F and F'_1 are both subbundles of F_n .

As $\bar{\partial}: \tilde{E} \rightarrow \tilde{F}$ is an isomorphism, we can use equation (1.3.7) to get

$$(1.3.18) \quad c_D(F', E'; w'') = c_D(F, E; w).$$

Thus we will get the desired result $c_D(F, E; w) = c_D(F', E'; w')$ if we can show $c_D(F', E'; w') = c_D(F', E'; w'')$. To this aim, it is enough to find a homotopy from w' to w'' , that is a continuous family of sections w_t , $0 \leq t \leq 1$, with $w_1 = w'$, $w_0 = w''$, and with the w_t all “injective” ($w_t(s) = 0$ only if $s = 0$).

This can be done as follows. Define $a_t: E' \rightarrow E'$ by $a_t(e \oplus \tilde{e}) = e \oplus t\tilde{e}$, and $b_t: F' \rightarrow F'$ by $b_t(f \oplus \tilde{f}) = f \oplus t\tilde{f}$. Then set

$$(1.3.19) \quad w_t(s) = \bar{\partial}s + b_{\bar{t}}\rho_{F'}(\overline{a_t s})^{r-1}.$$

It is clear that $w_1 = w'$ and $w_0 = w''$. To show injectivity, it is enough to show that

$$(1.3.20) \quad |w_t(s)|^2 = |\bar{\partial}(s)|^2 + |b_{\bar{t}}\rho_{F'}(\overline{a_t s})^{r-1}|^2.$$

For in that case, $w_t(s) = 0$ implies $\bar{\partial}s = 0$, whence $s \in E$, so $a_t s = s$, $\bar{s}^{r-1} \in F$, and $b_{\bar{t}}\rho_{F'}(\overline{a_t s})^{r-1} = \bar{s}^{r-1}$, which vanishes only if $s = 0$. To establish equation (1.3.20), we need vanishing of the cross terms

$$(1.3.21) \quad 0 = (\bar{\partial}s, b_{\bar{t}}\rho_{F'}(\overline{a_t s})^{r-1}).$$

The adjoint of $b_{\bar{t}}\rho_{F'}$ is $b_t\rho_{F'}$. Also $\rho_{F'}\bar{\partial}s = \bar{\partial}s$, $b_t\bar{\partial}s = \bar{\partial}a_t s$, so the right hand side of equation (1.3.21) is

$$(1.3.22) \quad \int_{\Sigma} \bar{\partial}(\overline{a_t s}) \cdot (\overline{a_t s})^{r-1} = \frac{1}{r} \int_{\Sigma} \bar{\partial}((\overline{a_t s})^r) = 0,$$

as desired.

Finally to indicate very briefly the quantum field theoretic origin of this definition, let me note that the twisted superconformal field theories which as originally advocated by K. Li [17] are relevant to this problem have various realizations. Apart from the realization by gauged WZW models used in [22], they can be realized as twisted Landau-Ginzburg models. In a version considered by K. Ito [13], the bosonic field can be interpreted as a section of \mathcal{T} , say s , and the bosonic part of the Lagrangian is $|w(s)|^2$ with w as defined above. In this version, one can see—for instance, by adapting arguments given by Atiyah and Jeffrey in [2]—that the Feynman path integral is a device for computing what we have called $c_D(\mathcal{F}, \mathcal{E}; w)$.

A purely algebraic version of the above definition of $c_D(\mathcal{V}')$ has been described recently by Faltings.

1.4 The Mumford-Morita-Miller classes. The other ingredients that we need to formulate the main conjecture can be defined more immediately. Each of the s marked points x_i of Σ has a cotangent space $T^*\Sigma|_{x_i}$, which, as Σ varies in $\bar{\mathcal{M}}'_{g,s}$, varies as the fiber of a complex line bundle \mathcal{L}_i . Henceforth, we associate with each marked point x_i a non-negative integer n_i (in addition to the m_i , $0 \leq m_i \leq r - 1$ introduced earlier). The objects of principal interest in this paper will be the intersection numbers

$$(1.4.1) \quad \frac{1}{r^g} \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i} \cdot c_D(\mathcal{V}), \bar{\mathcal{M}}'_{g,s} \right).$$

These numbers vanish, of course, unless a certain dimensional condition is obeyed, namely

$$(1.4.2) \quad \sum_i n_i + D = 3g - 3 + s.$$

We will present a precise conjectural formula for these numbers in the next subsection. As we will see, this formula reduces for $r = 2$ to a formula for the quantities

$$(1.4.3) \quad \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i}, \bar{\mathcal{M}}_{g,s} \right)$$

that was described in reference [21] and proved by Kontsevich in reference [14]. As was explained in [21], the numbers (1.4.3) contain the same information as the intersection numbers of the Mumford-Morita-Miller stable cohomology classes on $\bar{\mathcal{M}}_{g,s}$.

In an appropriate quantum field theory, described in [22], the intersection numbers just introduced appear as the expectation values of certain “operators” $\tau_{n,m}$ with respect to a suitable Feynman path integral measure. In this context it is natural to denote the intersection numbers of equation (1.4.3) as

$$(1.4.4) \quad \langle \tau_{n_1, m_1} \tau_{n_2, m_2} \cdots \tau_{n_s, m_s} \rangle.$$

One often denotes $\tau_{n,m}$ as $\tau_n(U_m)$, the n^{th} “gravitational descendant” of a “primary field” U_m .

The intersection numbers of interest are then naturally denoted

$$(1.4.5) \quad \left\langle \prod_{n,m} \tau_{n,m}^{d_{n,m}} \right\rangle$$

where the $d_{n,m}$ are non-negative integers, almost all zero, and (1.4.5) is to be understood as follows. Given a set of $d_{n,m}$'s, set

$$(1.4.6) \quad s = \sum_{n,m} d_{n,m}$$

Consider Riemann surfaces of genus g with s marked points, labeled as follows: for each (n, m) , precisely $d_{n,m}$ of the points are labeled by $(n_i, m_i) = (n, m)$. Determine g and D to obey equations (1.4.2) and (1.3.2). If the g so determined is not a non-negative integer, set (1.4.5) to zero; otherwise set

$$(1.4.7) \quad \left\langle \prod_{n,m} \tau_{n,m}^{d_{n,m}} \right\rangle = \frac{1}{r^g} \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i} \cdot c_D(\mathcal{V}), \overline{\mathcal{M}}'_{g,s} \right).$$

(This expression is considered to vanish for $g = 0, s \leq 2$ and $g = 1, s = 0$ where the moduli space of stable curves is empty.)

The intersection numbers that we have introduced are conveniently arranged in a generating functional. Introduce variables $t_{n,m}$, $n \in \mathbb{N}$, $0 \leq m \leq r - 1$. Set

$$(1.4.8) \quad F(t_{0,0}, t_{0,1}, \dots) = \sum_{d_{n,m}} \left\langle \prod_{n,m} \tau_{n,m}^{d_{n,m}} \right\rangle \prod_{n,m} \frac{\tau_{n,m}^{d_{n,m}}}{d_{n,m}!}.$$

F is known as the free energy.

If one restricts the sum in equation (1.4.8) to $d_{n,m}$'s such that the genus (determined as above) has a given value g , one gets what we will call the genus g contribution F_g to the free energy. Thus, $F = \sum_{g \geq 0} F_g$. We will compute F_0 in section 3.

If one considers $t_{n,m}$ to be of degree $1 - n - \gamma m$, then it follows from (1.4.2) and (1.3.2) that F_g is homogeneous of degree $(1 - g)(2 + 2\gamma)$. In particular, F_0 is the piece of highest degree, and in the context of the Gelfand-Dikii equations, which we will introduce presently, F_0 can be computed by systematically keeping only the highest degree terms. This can be done by replacing the commutators of differential operators in those equations by the Poisson brackets of corresponding symbols. This will be useful in section 3.3.

1.5 The Gelfand-Dikii equations. The generalized KdV hierarchies of Gelfand and Dikii [10] are conveniently described in terms of formal

pseudo-differential operators in one dimension. We consider an operator defined by a series

$$(1.5.1) \quad K = \sum_{i=-\infty}^n k_i D^i,$$

where the coefficients are functions $k_i(x)$ of a variable x , and

$$(1.5.2) \quad D = \frac{i}{\sqrt{r}} \frac{\partial}{\partial x}.$$

One makes the decomposition

$$(1.5.3) \quad K = K_+ + K_-, \quad K_+ = \sum_{i=0}^n k_i D^i.$$

One defines the residue of K as the function multiplying D^{-1} :

$$(1.5.4) \quad \text{res } K = k_{-1}.$$

Now consider differential operators of the form

$$(1.5.5) \quad Q = D^r - \sum_{i=0}^{r-2} u_i(x) D^i.$$

As a pseudodifferential operator, Q has a unique r^{th} root given by a series of the form $Q^{1/r} = D + \sum_{i>0} w_i D^{-i}$ with the w_i being differential polynomials in the u_i . The expression $Q^{n+m/r}$ will denote the pseudodifferential operator $(Q^{1/r})^{nr+m}$. As Q commutes with its own power $Q^{n+m/r}$, we have

$$(1.5.6) \quad [Q_+^{n+m/r}, Q] = -[Q_-^{n+m/r}, Q].$$

This commutator is a differential operator (obvious on the left hand side) of order at most $r-2$ (obvious on the right hand side). The coefficients of this differential operator are differential polynomials in the u_i . Hence one can introduce the Gelfand-Dikii equations

$$(1.5.7) \quad i \frac{\partial Q}{\partial t_{n,m}} = [Q_+^{n+(m+1)/r}, Q] \cdot \frac{c_{n,m}}{\sqrt{r}},$$

where $c_{n,m}$ are the constants

$$(1.5.8) \quad c_{n,m} = \frac{(-1)^n r^{n+1}}{(m+1)(r+m+1) \dots (nr+m+1)},$$

introduced for convenience. Concretely, the equations (1.5.7) are differential equations for the coefficients u_i of Q . These equations take the general form

$$(1.5.9) \quad \frac{\partial u_i}{\partial t_{n,m}} = R_{i;n,m}(u_j, \partial_x u_k, \partial_x^2 u_l, \dots),$$

where the $R_{i;n,m}$ are polynomials in the u_j and their x derivatives. The equation for $n = m = 0$ is just $\partial u_i / \partial t_{0,0} = \partial u_i / \partial x$, so it is natural to identify $t_{0,0}$ and x . It is possible to prove that the Gelfand-Dikii equations are compatible in the sense that the flows corresponding to $\partial / \partial t_{n,m}$ and $\partial / \partial t_{n',m'}$ commute. Notice that

$$(1.5.10) \quad \frac{\partial Q}{\partial t_{n,r-1}} = 0, \quad n = 0, 1, 2 \dots$$

since for $m = r - 1$, $Q_+^{n+(m+1)/r} = Q^{n+1}$ commutes with Q .

Instead of regarding the Gelfand-Dikii equations as equations for the u_i , they can equally well be regarded as equations for the objects

$$(1.5.11) \quad v_i = -\frac{r}{i+1} \text{res}(Q^{(i+1)/r}), \quad 0 \leq i \leq r-2.$$

Indeed, the v_i are differential polynomials in u_j , and the expressions for the v_i in terms of the u_j , being “triangular,” can be inverted to express the u_j as differential polynomials in the v_i .

1.6 The conjecture. We can finally formulate the conjectured connection of intersection theory on $\overline{\mathcal{M}}'_{g,s}$ with the generalized KdV hierarchies. Identify $t_{0,0}$ with x , and set

$$(1.6.1) \quad \frac{\partial^2 F}{\partial t_{0,0} \partial t_{0,i}} = v_i, \quad \text{for } 0 \leq i \leq r-2.$$

(The v_i were defined in equation (1.5.11).) The main conjecture is then that the differential operator Q constructed with these v_i obeys the Gelfand-Dikii equations (1.5.7) which we repeat:

$$(1.6.2) \quad i \frac{\partial Q}{\partial t_{n,m}} = [Q_+^{n+(m+1)/r}, Q] \cdot \frac{c_{n,m}}{\sqrt{r}}.$$

In addition, one can prove (as we will see in section 2) that F obeys the “string equation”:

$$(1.6.3) \quad \frac{\partial F}{\partial t_{0,0}} = \frac{1}{2} \sum_{i,j=0}^{r-2} \eta^{ij} t_{0,i} t_{0,j} + \sum_{n=1}^{\infty} \sum_{m=0}^{r-2} t_{n+1,m} \frac{\partial F}{\partial t_{n,m}}.$$

Here $\eta^{ij} = \delta_{i+j,r-2}$. These equations uniquely determine F , except for a possible additive constant,¹ essentially since equation (1.6.2) determines the dependence of F on all the $t_{n,m}$ except $t_{0,0}$, and equation (1.6.3) gives the “initial conditions” that determine the dependence on $t_{0,0}$. (The elementary argument that the Gelfand-Dikii and string equations together uniquely determine F is written out in detail—for $r = 2$, but the general case is no different—in section (2a) of [21].)

In view of equation (1.5.10), a special case of the conjecture is that

$$(1.6.4) \quad \frac{\partial F}{\partial t_{n,r-1}} = 0, \quad n \geq 0.$$

Using the string equation, the Gelfand-Dikii equations can be integrated once to give

$$(1.6.5) \quad \frac{\partial^2 F}{\partial t_{0,0} \partial t_{n,m}} = -c_{n,m} \operatorname{res}(Q^{n+\frac{(m+1)}{r}}).$$

For $n = 0$ this reduces to the equation (1.6.1) that defines the relation between F and Q . It is often convenient to take (1.6.3) and (1.6.5) as the basic equations.

1.7 Relation to the original conjecture; illustration of the definitions.

Now, let us explain how the present conjecture is related to the earlier one discussed in [21, 14]. The present conjecture reduces for $r = 2$ to the older one if equation (1.6.4) is valid (which we expect in general but will later prove only in genus zero). In this case, the only nonzero correlation functions are those of the $\tau_{n,0}$. With $r = 2$ and all $m_i = 0$, \mathcal{T} reduces to a square root of K , $\bar{\mathcal{M}}'_{g,s}$ reduces to the 2^{2g} fold cover of $\bar{\mathcal{M}}_{g,s}$ given by a choice of such square root, and $D = 0$. We have to interpret the factor of $c_0(\mathcal{V})$ that appears in the definition (1.4.7) of the correlation functions.

¹ One can consider this undetermined additive constant to correspond to the contribution of the degenerate moduli space of genus one with zero marked points. It was discussed from that standpoint in [21]. For our purposes one can just set $F(0, 0, \dots) = 0$.

$c_0(\mathcal{V}) = c_0(\mathcal{F}, \mathcal{E}; w)$ can be computed by restricting to a generic point $x \in \bar{\mathcal{M}}'_{g,s}$. In the case of an even spin structure, $H^0(\mathcal{T})$ and $H^1(\mathcal{T})$ vanish at generic x ; we can take the finite dimensional subbundles of \mathcal{F} and \mathcal{E} in the definition of $c_0(\mathcal{F}, \mathcal{E}; w)$ to be $E = F = 0$. Hence $c_0(\mathcal{F}, \mathcal{E}; w) = 1$. In the case of an odd spin structure, $H^0(\mathcal{T})$ and $H^1(\mathcal{T})$ are generically one dimensional. Over a generic point x in moduli space, one can take F and E to be one dimensional spaces of harmonic forms. The map $w: E \rightarrow \pi^*F$ is then $s \rightarrow \bar{s}$. This map is of degree -1 , so $c_0(\mathcal{F}, \mathcal{E}; w) = -1$.

The other classes in (1.4.7) are pullbacks from $\bar{\mathcal{M}}_{g,s}$. The even spin structures give a cover of $\bar{\mathcal{M}}_{g,s}$ of degree $(2^{2g} + 2^g)/2$. The odd spin structures give a cover of degree $(2^{2g} - 2^g)/2$. In view of the evaluation of $c_0(\mathcal{V})$, the sum over spin structures gives the difference of these numbers or 2^g ; this factor cancels the explicit denominator in (1.4.7), which thus reduces to

$$(1.7.1) \quad \left\langle \prod_n \tau_{n,0}^{d_{n,0}} \right\rangle = \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i}, \bar{\mathcal{M}}_{g,s} \right).$$

This is precisely the definition of the correlation functions given in [21]. That the conjectured formula for these correlation functions given here reduces to the formula proposed in [21] is more obvious, since the Gelfand-Dikii hierarchy reduces at $r = 2$ to the KdV hierarchy.

For a similar but perhaps more striking illustration of the definition of $c_D(\mathcal{V})$, let us compute for arbitrary r the object $\langle \tau_{1,0} \rangle$. On dimensional grounds this receives a contribution only from genus one, so $D = 0$ and

$$(1.7.2) \quad \langle \tau_{1,0} \rangle = \frac{1}{r} (c_1(\mathcal{L}) \cdot c_0(\mathcal{V}), \bar{\mathcal{M}}_{1,1}).$$

As explained in [21], \mathcal{L} when regarded as a line bundle over $\bar{\mathcal{M}}_{1,1}$ has degree $1/24$; its first Chern class is Poincaré dual to $[x]/24$, where x is a generic point in $\bar{\mathcal{M}}_{1,1}$. (We must recall that moduli space is an orbifold, not a manifold; Chern classes need not be integral.) We must take account of the cover of degree r^2 corresponding to the choice of an r^{th} root \mathcal{T} of the (trivial) canonical bundle K . For $r^2 - 1$ of the possible choices, \mathcal{T} is a non-trivial line bundle, and $H^0(\mathcal{T}) = H^1(\mathcal{T}) = 0$. In the definition of $c_0(\mathcal{F}, \mathcal{E}; w)$, we can take $F = E = 0$, so $c_0(\mathcal{F}, \mathcal{E}; w) = 1$. For the other choice, \mathcal{T} is trivial, and $H^0(\mathcal{T})$ and $H^1(\mathcal{T})$ are each one dimensional. Taking F and E to be the one dimensional spaces of harmonic forms, the map $w: E \rightarrow \pi^*F$ is $s \rightarrow \bar{s}^{r-1}$, which is of degree $-(r-1)$. The sum over the choices of \mathcal{T} thus gives a factor $(r^2 - 1) \cdot 1 - 1 \cdot (r-1) = r(r-1)$. Using this in (1.7.2), we get

$\langle \tau_{1,0} \rangle = (r-1)/24$, which can be shown to agree with the prediction from the Gelfand-Dikii equations.

The remainder of this paper is devoted to explaining what is known about the conjecture for $r > 2$. In section 2, we discuss some general issues, and in section 3, we verify that the conjecture is valid in genus zero.

2. Some further properties

2.1 Decoupling of the Ramond sector. Our first goal is to verify that (1.6.4) holds in genus zero. It appears that deeper arguments would be needed to prove this property for arbitrary genus.

It will be helpful to know that \mathcal{V} is always a D dimensional vector bundle in genus zero. Indeed, \mathcal{V} is a vector bundle of this dimension provided $H^0(\Sigma, \mathcal{T}) = 0$, which will be so if the degree of \mathcal{T} is negative. The degree of \mathcal{T} is actually

$$(2.1.1) \quad \gamma \cdot \left(2g - 2 - \sum_i m_i \right).$$

This is indeed always negative for $g = 0$. (It does not matter if Σ degenerates, since if so each component has genus zero, and the above argument can be applied to each component separately.)

Consider a Riemann surface Σ of genus g with s marked points x_1, \dots, x_s labeled by (n_i, m_i) . Equation (1.6.4) is the assertion that the intersection numbers (1.4.1) vanish if $m_i = r-1$ for some i . This will be true if $c_D(\mathcal{V}) = 0$ at least rationally. Given that \mathcal{V} is a vector bundle of dimension D , it suffices to find a surjective map $\mathcal{V} \rightarrow \mathbb{C}$. Undetermined r^{th} roots of unity in the following discussion mean that our conclusions will be valid only rationally.

By Serre duality, \mathcal{V} is the vector bundle whose fiber is $H^0(\Sigma, K \otimes \mathcal{T}^{-1})$ for Σ a smooth curve (and $H^0(\Sigma, \text{Hom}(\mathcal{T}, K))$ in general; this refinement will not affect the argument since the x_i never coincide with the double points). A section of $K \otimes \mathcal{T}^{-1}$ is, intuitively, a $(1-\gamma)$ -differential with a possible pole of order m_j/r at each x_j . In particular, at x_i the possible order of the pole is $1-\gamma$, and by extracting the residue, that is the coefficient of

$$(2.1.2) \quad \left(\frac{dx}{x-x_i} \right)^{1-\gamma},$$

one gets a linear map $H^0(\Sigma, K \otimes \mathcal{T}^{-1}) \rightarrow \mathbb{C}$. As the pair (Σ, \mathcal{T}) varies in moduli space, this residue map varies as a morphism $\mathcal{V} \rightarrow \mathbb{C}$. It suffices to prove that this is surjective in genus zero.

This is an exercise from the Riemann-Roch theorem. One has the exact sequence of sheaves on Σ

$$(2.1.3) \quad 0 \rightarrow K \otimes \mathcal{T}^{-1} \otimes \mathcal{O}(x_i)^{-1} \rightarrow K \otimes \mathcal{T}^{-1} \rightarrow (K \otimes \mathcal{T}^{-1})|_{x_i} \rightarrow 0.$$

Note that for $m_i = r - 1$, $(K \otimes \mathcal{T}^{-1})|_{x_i}$ has a canonical identification with \mathbb{C} , this being the residue map. The exact sequence in cohomology is

$$(2.1.4) \quad \cdots H^0(\Sigma, K \otimes \mathcal{T}^{-1}) \xrightarrow{\text{Res}_{x_i}} \mathbb{C} \rightarrow H^1(\Sigma, K \otimes \mathcal{T}^{-1} \otimes \mathcal{O}(x_i)^{-1}) \rightarrow \cdots$$

The residue map is therefore surjective if

$$(2.1.5) \quad H^1(\Sigma, K \otimes \mathcal{T}^{-1} \otimes \mathcal{O}(x_i)^{-1}) = 0.$$

That will be so if the degree of $K \otimes \mathcal{T}^{-1} \otimes \mathcal{O}(x_i)^{-1}$ is $\geq 2g - 1$. That degree is actually

$$(2.1.6) \quad D' = (2g - 2)(1 - \gamma) + \gamma \sum_j m_j - 1.$$

Bearing in mind that D' must be an integer, this formula implies that if $m_i = r - 1$, then $D' \geq -1$ for $g = 0$, completing the proof that $c_D(\mathcal{V}) = 0$ (rationally) in that case.

2.2 Recursion relations of topological gravity. We will now show that the correlation functions of (1.4.1) obey the general genus zero recursion relations of topological gravity, described in [21]. As explained there, these are the closest analogs of the Gelfand-Dikii equations that seem easy to understand algebrogeometrically. Most of the discussion is in parallel with that of [21], but there are a couple of special points. The relations we will obtain are known [7] to agree with the Gelfand-Dikii equations.

Consider a genus zero curve Σ with s marked points x_1, \dots, x_s , labeled by (n_i, m_i) . We wish to compute

$$(2.2.1) \quad \langle \tau_{n_1, m_1} \cdots \tau_{n_s, m_s} \rangle = \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i} \cdot c_D(\mathcal{V}), \bar{\mathcal{M}}'_{0,s} \right).$$

Actually, in genus zero, the r^{th} root of a line bundle is unique if it exists, so the cover of moduli space that we have introduced is trivial and $\bar{\mathcal{M}}'_{0,s} = \bar{\mathcal{M}}_{0,s}$.

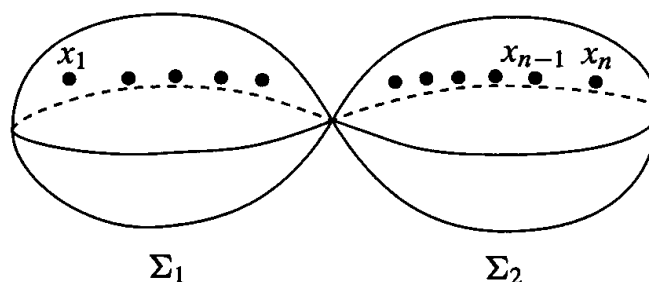


Figure 1.

(2.2.1) can be simplified inductively by picking suitable sections of the line bundles \mathcal{L}_i . For instance, suppose Σ is a smooth genus zero curve, which can be identified with the Riemann sphere, and consider the differential form on Σ

$$(2.2.2) \quad \omega = \frac{dx}{x - x_{s-1}} - \frac{dx}{x - x_s}.$$

It can be described invariantly as the unique differential form on Σ whose only singularities are simple poles at x_{s-1} and x_s , with residues 1 and -1 . If Σ is a smooth curve, ω has no zeroes. For a genus zero stable curve with several components (see Figure 1), there is still a unique ω whose only singularities¹ are the simple poles at x_{s-1} and x_s of prescribed residues. However, ω may have zeroes—in fact it will vanish identically on certain components. In complex codimension one, which is the important case in the present discussion, Σ can degenerate at most to a curve with two components, Σ_1 and Σ_2 , sharing a double point P (as in the Figure). A differential form on such a Σ is permitted to have a simple pole at P , with equal and opposite residues on the two branches. If x_{s-1} and x_s are on the same component, say Σ_2 , then ω vanishes identically on Σ_1 (otherwise it would be a differential on Σ_1 with at most a single pole at P , which is impossible). If x_{s-1} and x_s are on opposite branches, say Σ_1 and Σ_2 , respectively, then ω has no zeroes. (On Σ_1 it has poles at x_{s-1} and P , and on Σ_2 it has poles at x_s and P .)

A section of \mathcal{L}_1 can be obtained by evaluating ω at x_1 :

$$(2.2.3) \quad s = \frac{dx_1}{x_1 - x_{s-1}} - \frac{dx_1}{x_{s-1} - x_s}.$$

¹ Apart from simple poles at the double points, with equal and opposite residues at the two branches; these are permitted in the definition of the canonical bundle of a curve with double points.

From what has been said, it is easy to determine the divisor of s . s has no poles (since x_1 never coincides with the only possible poles at x_{s-1} and x_s). s vanishes precisely when Σ degenerates to two branches with x_1 on one branch, say Σ_1 , and x_{s-1} and x_s on the other branch. So let S be the finite set $\{2, 3, \dots, s-2\}$. For any decomposition of S as a union of disjoint subsets X and Y , let $D_{X,Y}$ be the divisor in $\bar{\mathcal{M}}_{0,s}$ parameterizing curves that degenerate to two components, one containing x_1 and x_j , $j \in X$, and the other containing x_{s-1} , x_s , and x_j , $j \in Y$. The divisor of s is the sum of the $D_{X,Y}$. Replacing $c_1(\mathcal{L}_1)^{n_1}$ by $c_1(\mathcal{L}_1)^{n_1-1}$ times the divisor of s , we can rewrite (2.2.1) as follows:

$$(2.2.4) \quad \langle \tau_{n_1, m_1} \dots \tau_{n_s, m_s} \rangle = \sum_{S=X \cup Y} \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i - \delta_{i,1}} \cdot c_D(\mathcal{V}), D_{X,Y} \right).$$

To proceed further, we need to understand the restriction of \mathcal{V} to $D_{X,Y}$. Let $\bar{\Sigma}$ be the normalization of Σ , which is the disjoint union of Σ_1 and Σ_2 . Let P' and P'' be the inverse images of P on Σ_1 and Σ_2 . As we have seen in equations (1.2.6) and (1.2.8), the P' and P'' behave rather like “ordinary” marked points labeled with integers m' and m'' in the usual range, with either $m' \leq r-2$ and $m'' = r-2-m'$, or $m' = m'' = r-1$. m' is determined by considering the degeneration of the sheaf \mathcal{T} ; indeed

$$(2.2.5) \quad -2 - m_1 - \sum_{j \in X} m_j - m'$$

must be divisible by r . Similarly, m'' is such that

$$(2.2.6) \quad -2 - \sum_{j \in Y} m_j - m_{s-1} - m_s - m''$$

is divisible by r .

As discussed in connection with equation (1.2.6), if $m', m'' \leq r-2$, then \mathcal{T} is the direct image of a locally free sheaf \mathcal{T}' on $\bar{\Sigma}$. In this case, $\mathcal{V} = H^0(\Sigma, K \otimes \mathcal{T}^{-1})$ has a direct sum decomposition

$$(2.2.7) \quad \mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2,$$

where

$$(2.2.8) \quad \mathcal{V}_i = H^0(\Sigma_i, K \otimes \mathcal{T}'^{-1}), \quad i = 1, 2.$$

Hence

$$(2.2.9) \quad c_D(\mathcal{V}) = c_{D_1}(\mathcal{V}_1) \cdot c_{D_2}(\mathcal{V}_2),$$

with $D_i = \dim(\mathcal{V}_i)$. Moreover, as $D_{X,Y} \cong \bar{\mathcal{M}}_{0,2+n_X} \times \bar{\mathcal{M}}_{0,3+n_Y}$ (n_X and n_Y are the cardinalities of X and Y), we get

$$(2.2.10) \quad \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i - \delta_{i,1}} \cdot c_D(\mathcal{V}), D_{X,Y} \right) \\ = \left(c_1(\mathcal{L}_1)^{n_1-1} \prod_{j \in X} c_1(\mathcal{L}_j)^{n_j} \cdot c_{D_1}(\mathcal{V}_1), \bar{\mathcal{M}}_{0,2+n_X} \right) \\ \cdot \left(\prod_{j \in Y} c_1(\mathcal{L}_j)^{n_j} \cdot c_1(\mathcal{L}_{s-1})^{n_{s-1}} c_1(\mathcal{L}_s)^{n_s} c_{D_2}(\mathcal{V}_2), \bar{\mathcal{M}}_{0,3+n_Y} \right).$$

(This vanishes if X is empty since $\bar{\mathcal{M}}_{0,2}$ is empty.)

In view of equation (1.2.7), the situation is different if $m' = m'' = r - 1$. In this case, it is the subsheaf \mathcal{T}' of \mathcal{T} that is the direct image of a locally free sheaf \mathcal{T}'' on the normalization. Setting $\mathcal{V}_i = H^0(\Sigma_i, \mathcal{T}'')$, $i = 1, 2$, we get from (1.2.7) and (1.2.8) an exact sequence

$$(2.2.11) \quad 0 \rightarrow \mathcal{V}_1 \oplus \mathcal{V}_2 \rightarrow \mathcal{V} \xrightarrow{\text{Res}} \mathcal{O} \rightarrow 0,$$

instead of (2.2.7). Here Res is the map that extracts the residue at the double point. Consequently, $c_D(\mathcal{V}) = 0$ if the residue map is always surjective. This can be proved just as we proved a similar statement in establishing the decoupling of the Ramond sector. This vanishing of $c_D(\mathcal{V})$ means that the contributions to (2.2.4) with $m' = m'' = r - 1$ can be dropped. (Physicists describe what we have just proved by saying that at least in genus zero, the decoupling of the Ramond sector is compatible with factorization.)

By using (2.2.10) to evaluate the terms with $m', m'' \leq r - 2$, and dropping the terms with $m' = m'' = r - 1$, (2.2.4) can be rewritten

$$(2.2.12) \quad \left\langle \prod_{i=1}^s \tau_{n_i, m_i} \right\rangle = \sum_{S=X \cup Y} \sum_{m', m''=0}^{r-2} \left\langle \tau_{n_1-1, m_1} \prod_{j \in X} \tau_{n_j, m_j} \tau_{0, m'} \right\rangle \\ \cdot \eta^{m' m''} \cdot \left\langle \tau_{0, m''} \prod_{j \in Y} \tau_{n_j, m_j} \cdot \tau_{n_{s-1}, m_{s-1}} \tau_{n_s, m_s} \right\rangle.$$

Obviously, by repeated use of this formula one can express all genus zero correlation functions in terms of the correlation functions

$$(2.2.13) \quad \langle \tau_{0,m_1} \tau_{0,m_2} \dots \tau_{0,m_s} \rangle = (c_D(\mathcal{V}), \bar{\mathcal{M}}_{0,s}).$$

of “primary fields.” (2.2.13) vanishes on dimensional grounds unless

$$(2.2.14) \quad s - 3 = D = -1 + 2\gamma + \gamma \sum_i m_i.$$

From this it follows that (2.2.13) vanishes if $s > r + 1$, so for each r there are only finitely many correlation functions of primaries to be determined. They will be evaluated in section 3.

If we set $s = 4$, and shift $n_1 \rightarrow n_1 + 1$, (2.2.12) implies

$$(2.2.15) \quad \begin{aligned} & \langle \tau_{n_1+1,m_1} \tau_{n_2,m_2} \tau_{n_3,m_3} \tau_{n_4,m_4} \rangle \\ &= \sum_{m',m''} \langle \tau_{n_1,m_1} \tau_{n_2,m_2} \tau_{0,m'} \rangle \cdot \eta^{m'm''} \cdot \langle \tau_{0,m''} \tau_{n_3,m_3} \tau_{n_4,m_4} \rangle. \end{aligned}$$

As the left hand side is invariant under permutations of 2, 3, 4, we get

$$(2.2.16) \quad \begin{aligned} & \langle \tau_{n_1,m_1} \tau_{n_2,m_2} \tau_{0,m'} \rangle \eta^{m'm''} \langle \tau_{0,m''} \tau_{n_3,m_3} \tau_{n_4,m_4} \rangle \\ &= \langle \tau_{n_1,m_1} \tau_{n_3,m_3} \tau_{0,m'} \rangle \eta^{m'm''} \langle \tau_{0,m''} \tau_{n_2,m_2} \tau_{n_4,m_4} \rangle \end{aligned}$$

As explained in [21], these formulas have the following interpretation. First of all, let

$$(2.2.17) \quad \langle\langle \tau_{n_1,m_1} \dots \tau_{n_s,m_s} \rangle\rangle = \frac{\partial}{\partial t_{n_1,m_1}} \dots \frac{\partial}{\partial t_{n_s,m_s}} F(t_{0,0}, \dots).$$

Thus, the left hand side of (2.2.17) is a function of the $t_{n,m}$ which, at $t_{n,m} = 0$, reduces to $\langle \tau_{n_1,m_1} \dots \tau_{n_s,m_s} \rangle$. In deriving (2.2.16), we have set $s = 4$ in (2.2.12), but as is explained in [21] (see the derivation of equations (2.72) and (3.28) of that paper), by considering the equations (2.2.12) for arbitrary s , one learns that the $\langle\langle \dots \rangle\rangle$ objects obey the analog of (2.2.16):

$$(2.2.18) \quad \begin{aligned} & \langle\langle \tau_{n_1,m_1} \tau_{n_2,m_2} \tau_{0,m'} \rangle\rangle \eta^{m'm''} \langle\langle \tau_{0,m''} \tau_{n_3,m_3} \tau_{n_4,m_4} \rangle\rangle \\ &= \langle\langle \tau_{n_1,m_1} \tau_{n_3,m_3} \tau_{0,m'} \rangle\rangle \eta^{m'm''} \langle\langle \tau_{0,m''} \tau_{n_2,m_2} \tau_{n_4,m_4} \rangle\rangle. \end{aligned}$$

The special case of this equation with all $n_i = 0$ can be interpreted as follows. Set $t_i = t_{0,i}$, and let

$$(2.2.19) \quad c_{ijk} = \frac{\partial^3 F}{\partial t_i \partial t_j \partial t_k}$$

and $c_{ij}{}^k = c_{ijs} \eta^{sk}$. Consider an algebra generated by objects ϕ_i with multiplication law

$$(2.2.20) \quad \phi_i \phi_j = \sum_k c_{ij}{}^k \phi_k.$$

Then (2.2.18) says that this is a commutative, associative algebra for every value of the $t_{n,m}$, which moreover is compatible with the metric η (in the sense that $\eta(ab, c) = \eta(b, ac)$) since the c_{ijk} are completely symmetric. Thus, the function F has the property that its third derivatives at any point provide the structure constants of a commutative, associative algebra, compatible with the metric $ds^2 = \eta^{ij} dt_i dt_j$.

It may sound well-nigh impossible for a non-cubic function to have this property, so here are a few examples. If $t_{n,m} = 0$ for $n > 0$, then F reduces to a polynomial of degree $r + 1$ in the $t_m = t_{0,m}$, $0 \leq m \leq r - 2$. These polynomials will be determined in section 3. Here are the first few. For $r = 2$,

$$(2.2.21) \quad F = \frac{t_0^3}{6}.$$

For $r = 3$,

$$(2.2.22) \quad F = \frac{t_0^2 t_1}{2} + \frac{t_1^4}{72}.$$

And for $r = 4$,

$$(2.2.23) \quad F = \frac{t_0^2 t_2 + t_0 t_1^2}{2} + \frac{t_1^2 t_2^2}{16} + \frac{t_2^5}{8 \cdot 5!}.$$

2.3 The string equation. Now, as in [6, 21], we will explain the basis for the string equation, which we recall:

$$(2.3.1) \quad \frac{\partial F}{\partial t_{0,0}} = \frac{1}{2} \sum_{i,j=0}^{r-2} \eta^{ij} t_{0,i} t_{0,j} + \sum_{n=1}^{\infty} \sum_{m=0}^{r-2} t_{n+1,m} \frac{\partial F}{\partial t_{n,m}}.$$

To justify it, let Σ be a curve of genus g with $s+1$ marked points x_0, \dots, x_s , with x_0 labeled by $(0, 0)$ and the other x_i , $i > 0$ labeled by (n_i, m_i) . The corresponding correlation function is

$$(2.3.2) \quad \left\langle \tau_{0,0} \cdot \prod_{i=1}^s \tau_{n_i, m_i} \right\rangle = \frac{1}{r^g} \left(\prod_{i=1}^s c_1(\mathcal{L}_i)^{n_i} \cdot c_D(\mathcal{V}), \bar{\mathcal{M}}'_{g, s+1} \right).$$

Now (except for a few low values of g and s) there is a map $\pi: \bar{\mathcal{M}}'_{g, s+1} \rightarrow \bar{\mathcal{M}}'_{g, s}$ that corresponds to forgetting about x_0 .¹ (At the level of the curves parameterized by the $\bar{\mathcal{M}}$'s, this operation is more subtle, as we will recall momentarily.) If it were the case that \mathcal{L}_i and \mathcal{V} were pullbacks of similar objects on $\bar{\mathcal{M}}'_{g, s}$, then (2.3.2) would vanish trivially, on dimensional grounds. In that case, one would get not the string equation (2.3.1), but the simpler formula

$$(2.3.3) \quad \frac{\partial F}{\partial t_{0,0}} = 0.$$

Actually, the \mathcal{L}_i are not pullbacks of the analogous objects on $\bar{\mathcal{M}}'_{g, s}$. This depends on the following. Let $\bar{\mathcal{C}}'_{g, s+1}$ be the universal curve parameterized by $\bar{\mathcal{M}}'_{g, s+1}$. The map $\pi: \bar{\mathcal{C}}'_{g, s+1} \rightarrow \bar{\mathcal{C}}'_{g, s}$ that induces the forgetful map of the $\bar{\mathcal{M}}$'s does not consist just of forgetting x_0 ; it also contracts to a point any genus zero component that is left with only two marked points (or double points) if x_0 is forgotten. Because of the latter step, the \mathcal{L}_i are not pullbacks of the analogous objects on $\bar{\mathcal{M}}'_{g, s}$. Rather, there is a computable correction, explained in references [6, 21], which leads to the term $\sum_{n, m} t_{n+1, m} \partial F / \partial t_{n, m}$ on the right hand side of the string equation.

Also, for $g = 0, s = 2$ and $g = 1, s = 0$, the forgetful map $\bar{\mathcal{M}}'_{g, s+1} \rightarrow \bar{\mathcal{M}}'_{g, s}$ does not exist. The former case leads to the $\eta^{ij} t_{0, i} t_{0, j} / 2$ term in the string equation, as explained in [21], and the latter does not contribute since for $g = 1, s = 0$, the correlation function of equation (2.3.2) vanishes on dimensional grounds.

These considerations suffice to prove the string equation if the $c_D(\mathcal{V})$ factor is deleted in (2.3.1) (which was the case considered in [21]). To

¹ This depends on the fact that $m_0 = 0$, so that if Σ is irreducible, an r^{th} root \mathcal{T} of $K \otimes \mathcal{O}(x_i)^{-m_i}$ can be chosen independent of x_0 . If Σ is reducible, and a genus zero component of Σ is contracted upon forgetting about x_0 , the possible choices of \mathcal{T} on Σ are still in natural 1-1 correspondence with their direct images \mathcal{T}_0 on the contracted curve Σ_0 .

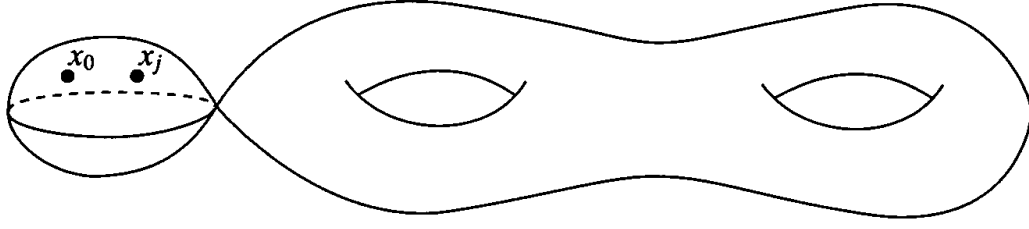


Figure 2.

justify the string equation with that factor present, we must also show that the sheaves \mathcal{V}' and \mathcal{V} on $\bar{\mathcal{M}}'_{g,s+1}$ are pullbacks of the corresponding objects on $\bar{\mathcal{M}}'_{g,s}$. To see this, the main point is that if (as in Figure 2) Σ_0 is a genus zero component of Σ which contains x_0 and will be contracted if x_0 is forgotten, then $H^1(\Sigma, \mathcal{T})$ has no elements supported on Σ_0 . (This is an easy consequence of the fact that such a Σ_0 has only three marked or double points, one of which, x_0 , is labeled by $m = 0$.) Hence if $\tilde{\Sigma}, \tilde{\mathcal{T}}$ is the result of contracting Σ_0 , then $H^1(\Sigma, \mathcal{T}) \cong H^1(\tilde{\Sigma}, \tilde{\mathcal{T}})$.

3. Analysis in genus zero

3.1 The first cases. In section 2.2, we obtained a recursion relation which determines all correlation functions $\langle \tau_{n_1, m_1} \dots \tau_{n_s, m_s} \rangle$ in genus zero in terms of the correlation functions of the primary fields $\tau_{0, m} = \tau_m$. In this section, we will complete the description in genus zero by determining the quantities

$$(3.1.1) \quad \langle \tau_{m_1} \dots \tau_{m_s} \rangle = (c_D(\mathcal{V}), \bar{\mathcal{M}}_{0,s}).$$

Here

$$(3.1.2) \quad D = -1 + 2\gamma + \gamma \sum_i m_i = s - 3.$$

The correlation function of equation (3.1.1) vanishes for $s < 3$, since $\bar{\mathcal{M}}_{0,s}$ is empty. For $s = 3$, the moduli space of stable curves consists of a single point. The dimensional condition (3.1.2) requires $m_1 + m_2 + m_3 = r - 2$. In that case, \mathcal{V} is zero dimensional, and $c_D(\mathcal{V}) = 1$, so we get

$$(3.1.3) \quad \langle \tau_{m_1} \tau_{m_2} \tau_{m_3} \rangle = \delta_{m_1+m_2+m_3, r-2}.$$

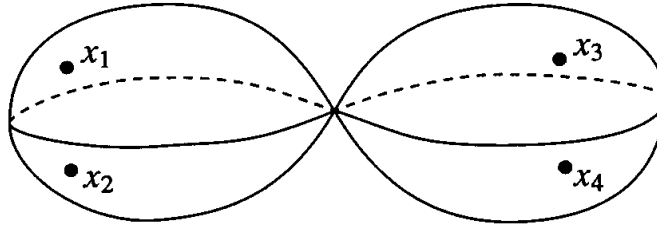


Figure 3.

Now we move on to $s = 4$. The moduli space is one dimensional, so the dimensional condition is $\sum_{i=1}^4 m_i = 2r - 2$, and \mathcal{V} is a line bundle. A smooth curve Σ of genus zero with four marked points can be identified as the complex plane (plus infinity) with four distinct marked points x_1, \dots, x_4 . This identification is unique up to an $SL(2, \mathbb{C})$ transformation of the x_i . Configurations of four distinct points, up to $SL(2, \mathbb{C})$, give the open moduli space $\mathcal{M}_{0,4}$. Compactification is achieved by adding three points at infinity where Σ degenerates (as in Figure 3) to two components Σ_1 and Σ_2 , sharing a double point and each containing precisely two of the m_i .

We recall that away from infinity, $\mathcal{V} = H^0(\Sigma, K \otimes \mathcal{T}^{-1})$, where \mathcal{T} is a line bundle with an isomorphism $\psi: \mathcal{T}^{\otimes r} \cong K \otimes_{i=1}^4 \mathcal{O}(x_i)^{-m_i}$. Such a \mathcal{T} is unique up to isomorphism (and the isomorphism is unique up to multiplication by an r^{th} root of unity). In general, $\mathcal{V} = H^0(\Sigma, \text{Hom}(\mathcal{T}, K))$, where the behavior of \mathcal{T} at infinity was explained in section 1.2. The first Chern class $c_1(\mathcal{V})$ can be measured by computing the divisor of a section. To this aim, we first pick a section s that trivializes \mathcal{V} over the finite part of moduli space, and then we determine the behavior at infinity.

A trivialization of \mathcal{V} over the finite part of moduli space is given by the section

$$(3.1.4) \quad s = (dx)^{1-\gamma} \cdot \frac{\prod_{i < j} (x_i - x_j)^{\frac{\gamma}{2}(m_i + m_j) - \frac{1}{3}(1-\gamma)}}{\prod_{k=1}^4 (x - x_k)^{\gamma m_k}}.$$

(The meaning of such a formula is of course that under $\psi': (K \otimes \mathcal{T}^{-1})^{\otimes r} \cong K^{\otimes(r-1)} \otimes \mathcal{O}(x_i)^{\otimes m_i}$, $s^{\otimes r}$ is mapped to the r^{th} power of the right hand side.) This expression is $SL(2, \mathbb{C})$ invariant, and so descends from the configurations of four points on the Riemann sphere to the moduli space $\mathcal{M}_{0,4}$. s obviously has no zeroes or poles on $\mathcal{M}_{0,4}$, where the x_i are distinct. We now must consider the behavior under degeneration.

As we have proved in section 2.1, \mathcal{V} is always a vector bundle in genus zero. In particular in the present situation, $V = H^0(\Sigma, \text{Hom}(\mathcal{T}, K))$ is always one dimensional, even if Σ is a degenerate curve with two components

Σ_i , $i = 1, 2$. However, in that case, depending on the values of the m_i , a generator v of V may vanish on one of the Σ_i . Indeed, each component of Σ contains precisely two of the x_i . Suppose that as in the figure x_1 and x_2 are on Σ_1 and x_3 and x_4 are on Σ_2 . If $m_1 + m_2 = m_3 + m_4$, then we are dealing with a “Ramond” degeneration in the language of section 1.2. \mathcal{T} is locally free over the degenerate curve, and v does not vanish on either component. Otherwise, say $m_1 + m_2 \leq r - 2$, $m_3 + m_4 \geq r$. This is a “Neveu-Schwarz” degeneration. \mathcal{T} is not locally free, but is the direct image of a locally free sheaf on the normalization; this sheaf has degree -1 on Σ_1 and -2 on Σ_2 . $\text{Hom}(\mathcal{T}, K)$ is likewise the direct image of a locally free sheaf on the normalization, which has degree -1 on Σ_1 and 0 on Σ_2 . Consequently, $H^0(\Sigma, \text{Hom}(\mathcal{T}, K))$ is one dimensional, as expected, and a generator v has its support on Σ_2 .

Now, let us determine the behavior of the section s of equation (3.1.4) near a degeneration with x_1, x_2 on Σ_1 and x_3, x_4 on Σ_2 . On Σ_1 , the degeneration is $x_3 \rightarrow x_4$, with x_1, x_2 fixed. In this process, s behaves as $(x_3 - x_4)^{\gamma(m_3+m_4)/2 - (1-\gamma)/3}$, so the order of zero (or minus the order of pole) is $\gamma(m_3 + m_4)/2 - (1 - \gamma)/3$. Similarly, on Σ_2 , the degeneration is $x_1 \rightarrow x_2$, with x_3, x_4 fixed, and the order of zero is $\gamma(m_1 + m_2)/2 - (1 - \gamma)/3$. The order of the zero of s is the smaller of these two numbers. (The order of the zero is necessarily larger on a component on which the generating section v of the last paragraph vanishes.) Adding similar contributions from the other points at infinity, the total degree of the divisor of s is

$$(3.1.5) \quad [s] = \frac{\gamma}{2} (\min(m_1 + m_2, m_3 + m_4) + \min(m_1 + m_3, m_2 + m_4) \\ + \min(m_1 + m_4, m_2 + m_3)) - (1 - \gamma).$$

It is elementary (but not very obvious) that this formula is equivalent to $[s] = \gamma \cdot \min(m_1, \dots, m_4, r - 1 - m_1, \dots, r - 1 - m_4)$. This gives the final result,

$$(3.1.6) \quad \langle \tau_{m_1} \dots \tau_{m_4} \rangle = \gamma \cdot \min(m_i, r - 1 - m_i).$$

3.2 Uniqueness. Direct computation of the correlation functions for $s \geq 5$ appears to be considerably more difficult. Luckily, we can proceed by using the associativity formula of section 2.2. Let $F(t_0, \dots, t_{r-2})$ be the generating function of the genus zero correlation functions of primaries; it is the same as the generating function (1.4.8) with $t_{0,m} = t_m, t_{n,m} = 0, n > 0$. (We can drop t_{r-1} since we have established the decoupling of the Ramond

sector in genus zero.) We know that $F(t_0, \dots, t_{r-2})$ is a polynomial of degree at most $r + 1$, and we have determined above the terms of degree ≤ 4 .

As in section 2.2 let

$$(3.2.1) \quad \langle\langle \tau_{m_1} \dots \tau_{m_s} \rangle\rangle = \frac{\partial^s F}{\partial t_{m_1} \dots \partial t_{m_s}}.$$

Then we obtained a formula

$$(3.2.2) \quad \langle\langle \tau_{m_1} \tau_{m_2} \tau_{m'} \rangle\rangle \eta^{m' m''} \langle\langle \tau_{m''} \tau_{m_3} \tau_{m_4} \rangle\rangle = \langle\langle \tau_{m_1} \tau_{m_3} \tau_{m'} \rangle\rangle \eta^{m' m''} \langle\langle \tau_{m''} \tau_{m_2} \tau_{m_4} \rangle\rangle$$

(which is an “associativity” formula in a sense explained in section 2.2). I will now show that there is at most one solution F of this equation that agrees with the known terms of degree ≤ 4 .

Suppose that F and $F + G$ are two such solutions. In particular, then, the lowest order terms in G are at least quintic. If inductively it is known that the lowest terms in G are of order p , we will prove that in fact the p^{th} order terms vanish. (The argument will only work for $p \geq 5$.) This will suffice to prove $G = 0$.

Let G' be the p^{th} order part of G , and for $s \leq p$, let

$$(3.2.3) \quad \{\tau_{m_1} \dots \tau_{m_s}\} = \frac{\partial^s G'}{\partial t_{m_1} \dots \partial t_{m_s}}.$$

Subtracting equation (3.2.2) for F from the same equation for $F + G$, the lowest order term (which is of order $p - 3$) is

$$(3.2.4) \quad \{\tau_{m_1} \tau_{m_2} \tau_{m_3+m_4}\} + \{\tau_{m_1+m_2} \tau_{m_3} \tau_{m_4}\} = \{\tau_{m_1} \tau_{m_3} \tau_{m_2+m_4}\} + \{\tau_{m_1+m_3} \tau_{m_2} \tau_{m_4}\}$$

where one is to set $\tau_m = 0$ if $m > r - 2$. In arriving at this formula, we have used $\eta^{m' m''} = \delta_{m' + m'', r-2}$ and $\langle\langle \tau_{m_1} \tau_{m_2} \tau_{m_3} \rangle\rangle = \delta_{m_1 + m_2 + m_3, r-2}$.

Consider an arbitrary p^{th} derivative of G' , say $\{\tau_{a_1} \dots \tau_{a_p}\}$. Let $z \geq y \geq x$ be the three largest of the a_i , and let b_j be the others. Set $m_1 = x + z - (r - 1)$, $m_2 = r - 1 - z$, $m_3 = y$, and $m_4 = z$. All these m_i obey $0 \leq m_i \leq r - 2$, but $m_3 + m_4$ and $m_2 + m_4$ are $> r - 2$, so if one inserts these values of the m 's in equation (3.2.4), at least two of the terms drop out. If $x + y + z \geq 2r - 2$, then three terms drop out of (3.2.4), which gives just $\{\tau_x \tau_y \tau_z\} = 0$, and hence $\{\tau_{a_1} \dots \tau_{a_p}\}$ (which is a derivative of this) is also zero. In general, (3.2.4) gives

$$(3.2.5) \quad \{\tau_x \tau_y \tau_z\} = \{\tau_{x+y+z-(r-1)} \tau_{r-1-z} \tau_z\}$$

and hence upon differentiation

$$(3.2.6) \quad \{\tau_{a_1} \dots \tau_{a_p}\} = \{\tau_{a'_1} \dots \tau_{a'_p}\}$$

where the a'_p are $x + y + z - (r - 1)$, $r - 1 - z$, z , and the b_j . Let $z' \geq y' \geq x'$ be the three largest a' 's. For $p \geq 5$, it follows from the dimensional formula (3.1.2) that $r - 1 - z$ is not one of x' , y' , z' , and hence that $x' + y' + z' > x + y + z$. Consequently, after repeating this process finitely many times, we learn that $\{\tau_{a_1} \dots \tau_{a_p}\} = 0$, as desired.

This shows that genus zero correlation functions of the $\tau_{0,m}$ are uniquely determined by the associativity equation and the terms of order ≤ 4 . If in addition one has the recursion relation of equation (2.2.12), then genus zero correlation functions of arbitrary $\tau_{n,m}$'s are uniquely determined.

3.3 Construction of F . We will now construct an F with the desired properties by analyzing the Gelfand-Dikii and string equations in genus zero. The F that those equations determine is known [7] to obey the algebrogeometric recursion relation of equation (2.2.12), which determines all genus zero correlation functions in terms of those of $t_{0,m}$. We will show that in addition this F obeys the associativity equation and agrees with the algebraic geometry as regards the terms of order ≤ 4 . In view of the uniqueness that we have just seen, these facts suffice to verify the main conjectures of section 1.6 in genus zero.

Background to what follows can be found in [8, 16]; the latter paper contains an elegant analysis, not limited to the small phase space, of the genus zero equations.

First of all, for reasons explained at the end of section 1.4, the genus zero approximation to the Gelfand-Dikii equations is obtained just by replacing commutators by Poisson brackets. Thus, writing $p = D$, we replace the differential operator Q of section 1.5 by a function

$$(3.3.1) \quad W(p, x) = p^r - \sum_{i=0}^{r-2} u_i(x) p^i.$$

The Poisson bracket of two functions A and B is

$$(3.3.2) \quad \{A, B\} = \frac{\partial A}{\partial p} \frac{\partial B}{\partial x} - \frac{\partial A}{\partial x} \frac{\partial B}{\partial p}.$$

The fractional power $W^{n/r}$, for integral n , will now denote the unique n/r power of W that is holomorphic for large $|p|$ and behaves for $p \rightarrow \infty$ as

$W \sim p^n$. A function holomorphic and power behaved for large $|p|$ has a Laurent expansion at infinity, $A(p) = \sum_{i=-\infty}^n A_n p^n$; we write $A = A_+ + A_-$, with $A_+ = \sum_{i=0}^n A_n p^n$. For any function $A(p)$ (holomorphic for large $|p|$), the residue operation is simply

$$(3.3.3) \quad \text{res}(A) = \oint \frac{dp}{2\pi i} A(p)$$

with the integral over a large circle at infinity. A useful fact is $\text{res}(AB_+) = \text{res}(A_-B)$, since $\text{res}(A_+B_+) = \text{res}(A_-B_-) = 0$.

In genus zero, the Gelfand-Dikii equations reduce to

$$(3.3.4) \quad \frac{\partial W}{\partial t_{n,m}} = \frac{c_{n,m}}{r} \cdot \{W_+^{n+(m+1)/r}, W\},$$

with

$$(3.3.5) \quad c_{n,m} = \frac{(-1)^n r^{n+1}}{(m+1)(r+m+1) \dots (nr+m+1)}.$$

W is determined in terms of the free energy F —which is the object we really want—by

$$(3.3.6) \quad \frac{\partial^2 F}{\partial t_{0,0} \partial t_{0,m}} = -\frac{r}{m+1} \text{res}(W^{(m+1)/r}), \quad 0 \leq m \leq r-2.$$

This uniquely determines the coefficients of W as polynomials in the $\partial^2 F / \partial t_{0,0} \partial t_{0,m}$. With W so determined, and with x identified with $t_{0,0}$, the Gelfand-Dikii equations (3.3.4) are equations for F . F is uniquely determined by those equations (or the once integrated version (1.6.5)) together with the string equation:

$$(3.3.7) \quad \frac{\partial F}{\partial t_{0,0}} = \frac{1}{2} \eta^{mm'} t_{0,m} t_{0,m'} + \sum_{n=0}^{\infty} \sum_m t_{n+1,m} \frac{\partial F}{\partial t_{n,m}}.$$

On the small phase space, that is if $t_{n,m} = 0$, $n > 0$, the string equation implies that

$$(3.3.8) \quad \frac{\partial^2 F}{\partial t_{0,0} \partial t_{0,m}} = t_{0,r-2-m}$$

So comparing with (3.3.6), if all $t_{n,m} = 0$, then $\text{res}(W^{(m+1)/r}) = 0$, $0 \leq m \leq r-2$, and hence $W = p^r$ at that point. Also, differentiating (3.3.6) with respect to $x = t_{0,0}$, on the small phase space

$$(3.3.9) \quad \delta_{m,r-2} = -\frac{r}{m+1} \frac{\partial}{\partial x} \text{res}(W^{(m+1)/r}).$$

$\partial W / \partial x$ is a polynomial in p (of order $\leq r-2$); if this polynomial is of order k , then the right hand side of (3.3.9) is non-zero for $m = r - k - 2$. Hence $k = 0$ and

$$(3.3.10) \quad \frac{\partial W}{\partial x} = -1.$$

From this it follows that for $0 \leq m \leq r-2$,

$$(3.3.11) \quad \frac{\partial}{\partial x} W_+^{(m+1)/r} = 0.$$

Hence, after explicitly evaluating the Poisson brackets, the first few Gelfand-Dikii equations reduce to

$$(3.3.12) \quad \frac{\partial W}{\partial t_{0,j}} = -\frac{1}{j+1} \frac{\partial}{\partial p} W_+^{(j+1)/r}.$$

A special case of the basic equation (1.6.5) is

$$(3.3.13) \quad \frac{\partial^2 F}{\partial t_{0,0} \partial t_{1,j}} = \frac{r^2}{(j+1)(r+j+1)} \text{res}(W^{1+(j+1)/r}).$$

The string equation implies that on the small phase space ($t_{n,m} = 0$, $n > 0$)

$$(3.3.14) \quad \frac{\partial^2 F}{\partial t_{0,0} \partial t_{1,j}} = \frac{\partial F}{\partial t_{0,j}},$$

so in fact

$$(3.3.15) \quad \frac{\partial F}{\partial t_{0,j}} = \frac{r^2}{(j+1)(r+j+1)} \text{res}(W^{1+(j+1)/r}).$$

Differentiating with respect to $t_{0,m}$ and using (3.3.12), we get

$$(3.3.16) \quad \frac{\partial^2 F}{\partial t_{0,j} \partial t_{0,m}} = \frac{r}{(j+1)(m+1)} \text{res} \left\{ \frac{\partial}{\partial p} W^{(m+1)/r} \cdot W_+^{(j+1)/r} \right\}.$$

Henceforth we work entirely on the small phase space, $t_{n,m} = 0$, $n > 0$, and we set $t_m = t_{0,m}$ and $\tau_m = \tau_{0,m}$.

For $0 \leq m \leq r-2$, let

$$(3.3.17) \quad \phi_m = \frac{1}{m+1} \frac{\partial}{\partial p} W_+^{(m+1)/r} = -\frac{\partial W}{\partial t_m}.$$

Thus $\phi_m = p^m + \text{lower order terms}$. In fact, the ϕ_m are the monic orthogonal polynomials for the “measure”

$$(3.3.18) \quad \langle A \rangle = r \cdot \text{res} \left\{ \frac{A}{\partial_p W} \right\},$$

in the sense that

$$(3.3.19) \quad \langle \phi_j \phi_m \rangle = \eta_{jm}.$$

Indeed in evaluating

$$(3.3.20) \quad \frac{r}{(j+1)(m+1)} \text{res} \left\{ \frac{\partial_p W_+^{(j+1)/r} \partial_p W_+^{(m+1)/r}}{\partial_p W} \right\},$$

we can replace $W_+^{(m+1)/r}$ by $W^{(m+1)/r}$ without affecting the residue, and then we have $\partial_p W^{(m+1)/r} = \partial_p W \cdot W^{(m+1)/r-1} \cdot (m+1)/r$. So (3.3.20) reduces to

$$(3.3.21) \quad \frac{1}{j+1} \text{res} \left\{ W^{(m+1)/r-1} \partial_p W_+^{(j+1)/r} \right\}.$$

This can be evaluated just from the leading behavior $W = p^r + \dots$, to give the claimed result.

Now we can describe a natural family of commutative, associative algebras, compatible with the metric η , parameterized by the t_m . Indeed, consider the algebra $\mathcal{A} = \mathbb{C}[p]/\partial_p W$. A basis for this commutative, associative algebra as a complex vector space is given by the monic orthogonal polynomials ϕ_m . If we write explicitly

$$(3.3.22) \quad \phi_j \phi_m = c_{jm}^t \phi_t \text{ mod } \partial_p W,$$

then $c_{jms} = c_{jm}^t \eta_{ts}$ is

$$(3.3.23) \quad c_{jms} = r \text{res} \left\{ \frac{\phi_j \phi_m \phi_s}{\partial_p W} \right\}.$$

In particular, c_{jms} is completely symmetric, showing that \mathcal{A} is compatible with the metric η . We will show that

$$(3.3.24) \quad \frac{\partial^3 F}{\partial t_j \partial t_m \partial t_s} = c_{jms}.$$

In particular, this implies that F obeys the associativity equation (3.2.2), which as we have seen almost uniquely determines it.

(3.3.24) can be verified by a straightforward but not particularly transparent calculation. By differentiating (3.3.16),

$$(3.3.25) \quad \frac{\partial^3 F}{\partial t_i \partial t_l \partial t_j} = -\frac{r}{l+1} \operatorname{res} \left(\phi_i \partial_j W^{(l+1)/r} \right) - \frac{r}{l+1} \operatorname{res} \left(\partial_j \phi_i \cdot W^{(l+1)/r} \right).$$

Also

$$(3.3.26) \quad \begin{aligned} \partial_j \phi_i &= \frac{1}{i+1} \partial_j \partial_p \left(W^{(i+1)/r} \right)_+ \\ &= \frac{1}{r} \partial_p \left(W^{(i+1)/r-1} \partial_j W \right)_+ \\ &= -\frac{1}{r} \partial_p \left(W^{(i+1)/r-1} \phi_j \right)_+. \end{aligned}$$

On the other hand,

$$(3.3.27) \quad \begin{aligned} r \operatorname{res} \left(\frac{\phi_i \phi_j \phi_l}{\partial_p W} \right) \\ = \operatorname{res} \left(\phi_i \phi_j W^{(l+1)/r-1} \right) - \frac{r}{l+1} \operatorname{res} \left(\frac{\phi_i \phi_j \partial_p W^{(l+1)/r}}{\partial_p W} \right). \end{aligned}$$

(3.3.27) can be related to (3.3.25) using (3.3.26) along with

$$(3.3.28) \quad \begin{aligned} \operatorname{res} \left(\phi_i \partial_j W^{(l+1)/r} \right) &= \frac{l+1}{r} \operatorname{res} \left(\phi_i W^{(l+1)/r-1} \partial_j W \right) \\ &= -\frac{l+1}{r} \operatorname{res} \left(\phi_i \phi_j W^{(l+1)/r-1} \right) \end{aligned}$$

and

$$(3.3.29) \quad \begin{aligned} \operatorname{res} \left(\frac{\phi_i \phi_j \partial_p W^{(l+1)/r}}{\partial_p W} \right) &= \operatorname{res} \left(\left(\frac{\phi_i \phi_j}{\partial_p W} \right)_+ \partial_p W^{(l+1)/r} \right) \\ &= -\frac{1}{r} \operatorname{res} \left(\partial_p \left(W^{(i+1)/r-1} \phi_j \right)_+ W^{(l+1)/r} \right). \end{aligned}$$

The verification of (3.3.24) is completed by assembling these formulas.

To complete the demonstration that the function F determined by the Gelfand-Dikii and string equations agrees in genus zero with the algebrogeometric calculation, we must check the terms of order ≤ 4 in an expansion in powers of the t_m . To begin with, the terms of order ≤ 2 should vanish (since $\bar{\mathcal{M}}_{0,s}$ is empty for $s \leq 2$). (3.3.16) implies that when all $t_{n,m} = 0$ (so that $W = p^r$)

$$(3.3.30) \quad \frac{\partial^2 F}{\partial t_j \partial t_m} = 0$$

so the quadratic terms vanish. The linear terms similarly vanish because of (3.3.15). As noted in the introduction, the Gelfand-Dikii and string equations do not determine an additive constant in F , and one can just set the zeroth order term $F(0, 0, \dots)$ to zero.

There remain the cubic and quartic terms in F . The cubic terms can be read off from equations (3.3.23) and (3.3.24). Setting $t_m = 0$, $\partial_p W = r p^{r-1}$, $\phi_m = p^m$, we get

$$(3.3.31) \quad \langle \tau_j \tau_m \tau_s \rangle = \left. \frac{\partial^3 F}{\partial t_j \partial t_m \partial t_s} \right|_{t_{n,m}=0} = \delta_{j+m+s, r-2},$$

in agreement with the intersection theory. The quartic terms can be computed the same way, but we need the first order corrections to W , which can be determined from (3.3.12). One finds

$$(3.3.32) \quad W = p^r - \sum_{m=0}^{r-2} t_j p^j.$$

Hence

$$(3.3.33) \quad W_+^{(m+1)/r} = p^{m+1} - \frac{m+1}{r} \sum_{i=r-(m+1)}^{r-2} t_i p^{m+1+i-r} + O(t^2)$$

and so

$$(3.3.34) \quad \begin{aligned} \phi_m &= \frac{1}{m+1} \partial_p W_+^{(m+1)/r} \\ &= p^m - \sum_{i=r-(m+1)}^{r-2} \frac{m+1+i-r}{r} t_i p^{m-(r-i)} + O(t^2). \end{aligned}$$

Let $\theta(x)$ be the function that is 1 for $x \geq 0$ and 0 for $x < 0$; and denote the contribution to $\partial^3 F / \partial t_j \partial t_m \partial t_s$ that is linear in the t 's as $\langle \tau_j \tau_m \tau_s \rangle'$. It is straightforward to evaluate the term in (3.3.23) linear in the t 's, with the result

$$(3.3.35) \quad \langle \tau_j \tau_m \tau_s \rangle' = \sum_{u=0}^{r-2} \frac{1}{r} t_u \left(u - (m + u - r - 1)\theta(m + u - r - 1) \right. \\ \left. - (j + u - r - 1)\theta(j + u - r - 1) \right. \\ \left. - (s + u - r - 1)\theta(s + u - r - 1) \right).$$

(The “ u ” term comes from the correction to W , and the terms proportional to $\theta(\dots)$ come from the corrections to the ϕ 's.) Differentiating this with respect to t_u , we get

$$(3.3.36) \quad \langle \tau_j \tau_m \tau_s \tau_u \rangle = \frac{1}{r} \left(u - (m + u - r - 1)\theta(m + u - r - 1) \right. \\ \left. - (j + u - r - 1)\theta(j + u - r - 1) \right. \\ \left. - (s + u - r - 1)\theta(s + u - r - 1) \right).$$

It is elementary—though not very obvious—that this is equivalent to

$$(3.3.37) \quad \langle \tau_{m_1} \dots \tau_{m_4} \rangle = \frac{1}{r} \min(m_i, r - 1 - m_i),$$

in agreement with the formula (3.1.6) that we obtained from the algebro-geometric definition.

I am grateful to P. Deligne for much advice in the course of this work. I also wish to thank R. Dijkgraaf for explanations of the formulas of [8] and related matters.

BIBLIOGRAPHY

1. M. F. Atiyah and I. M. Singer, *The Index Of Elliptic Operators* IV, Ann. Math. **93** (1971), 119-138.
2. M. F. Atiyah and L. Jeffrey, *Topological Lagrangians And Cohomology*, J. Geom. Phys. **7** (1990), 119-136.
3. B. H. Bowditch and D. B. A. Epstein, *Natural Triangulations Associated To A Surface*, Topology **27** (1988), p. 91.
4. E. Brezin and V. A. Kazakov, *Exactly Solvable Field Theories Of Closed Strings*, Phys. Lett. **B236** (1990), 144-150.

5. P. Candelas, X. C. de la Ossa, P. S. Green and L. Parkes, *A Pair of Calabi-Yau Manifolds As An Exactly Soluble Superconformal Field Theory*, Nuclear Physics **B359** (1991), 21-74.
6. P. Deligne, letter to E. Witten (October, 1989).
7. R. Dijkgraaf and E. Witten, *Mean Field Theory, Topological Field Theory, and Multi-Matrix Models*, Nuclear Physics **B342** (1990), 486-522.
8. R. Dijkgraaf, E. Verlinde, and H. Verlinde, *Topological Strings In $d < 1$* , preprint.
9. M. Douglas and S. Shenker, *Strings In Less Than One Dimension*, Nucl. Phys. B. **B335** (1990), 635-654.
10. I. M. Gelfand and L. A. Dikii, *The Resolvent And Hamiltonian Systems*, Funct. Anal. Appl. **11(2)** (1977), 93-105.
11. B. Greene and M. R. Plesser, *Duality in Calabi-Yau Moduli Space*, Nucl. Phys. **B338** (1990), 15-37.
12. D. J. Gross and A. A. Migdal, *A Nonperturbative Treatment of Two Dimensional Quantum Gravity*, Phys. Rev. Lett. **64** (1990), 127-130.
13. Kei Ito, *Topological Phase of $N = 2$ Superconformal Field Theory And Topological Landau-Ginzburg Field Theory*, Harvard preprint (1990).
14. M. Kontsevich, *Intersection Theory On The Moduli Space Of Curves*, preprint, 1990.
15. M. Kontsevich, *Intersection Theory On The Moduli Space Of Curves And The Matrix Airy Function*, (MPI preprint, 1991).
16. I. M. Krichever, *Topological Minimal Models And Soliton Equations*, preprint.
17. K. Li, *Topological Gravity With Minimal Matter; Recursion Relations In Topological Gravity With Minimal Matter*, Cal Tech preprints (1990).
18. H. Neuberger, *Regularized String And Flow Equations*, Rutgers University preprint RU-90-90.
19. R. Penner, *Perturbative Series And The Moduli Space Of Riemann Surfaces*, J. Diff. Geom. **27** (1988), 35-53.
20. R. Seeley and I. M. Singer, *Extending $\bar{\partial}$ To Singular Riemann Surfaces*, J. Geom. Phys. **5** (1988), 121-136.
21. E. Witten, *Two Dimensional Gravity And Intersection Theory On Moduli Space*, Surveys In Differential Geometry **1** (1991), 243-310.
22. E. Witten, *The N Matrix Model And Gauged WZW Models*, preprint IASSNS-HEP-91/26, to appear in Nucl. Phys. B.
23. E. Witten, *On The Kontsevich Model And Other Models Of Two Dimensional Gravity*, preprint IASSNS-HEP-91/24.