MATRIX INTEGRALS AND INTEGRABLE SYSTEMS

BY

Motohico Mulase

Department of Mathematics
University of California, Davis
Davis, CA 95616
momulase@ucdavis.edu

1. Hermitian matrix integrals.

Let $\mathcal{H}_n$ be the space of all $n \times n$ Hermitian matrices. This is an n^2-dimensional real Euclidean space. We denote by dX the usual Lebesgue measure on $\mathcal{H}_n$. The unitary group $U(n)$ acts on this space by conjugation. We call a function $f(X)$ on $\mathcal{H}_n$ *invariant* if it satisfies

$$f(X) = f(U \cdot X \cdot U^{-1}) \tag{1.1}$$

for every unitary matrix $U \in U(n)$. Since a Hermitian matrix is diagonalizable by a unitary matrix, (1.1) means that

$$f(X) = f(k_0, k_1, \cdots k_{n-1}) \tag{1.2}$$

is a symmetric function in the eigenvalues k_0, k_1, $\cdots$, k_{n-1} of $X \in \mathcal{H}_n$. The main object of this article is a matrix integral

$$Z_n(t, f) = \int_{\mathcal{H}_n} \exp\left(\operatorname{trace} \sum_{\alpha=1}^{m} t_\alpha X^\alpha\right) f(X)\, dX\ , \tag{1.3}$$

where $f(X)$ is an invariant function on $\mathcal{H}_n$. In this paper, we are interested in the case when there are n functions $\phi_j(k)$ in one variable such that

$$f(X) = f(k_0, k_1, \cdots k_{n-1}) = \frac{\det \begin{pmatrix} \phi_0(k_0) & \phi_1(k_0) & \dots & \phi_{n-1}(k_0) \\ \phi_0(k_1) & \phi_1(k_1) & \dots & \phi_{n-1}(k_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(k_{n-1}) & \phi_1(k_{n-1}) & \dots & \phi_{n-1}(k_{n-1}) \end{pmatrix}}{\Delta\big(k_0, k_1 \cdots, k_{n-1}\big)}\ , \tag{1.4}$$

where

$$\Delta\big(k_0, k_1 \cdots, k_{n-1}\big) = \det \begin{pmatrix} 1 & k_0 & k_0^2 & \dots & k_0^{n-1} \\ 1 & k_1 & k_1^2 & \dots & k_1^{n-1} \\ 1 & k_2 & k_2^2 & \dots & k_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & k_{n-1} & k_{n-1}^2 & \dots & k_{n-1}^{n-1} \end{pmatrix}$$

is the Vandermonde determinant. Note that the ratio in (1.4) is symmetric in the k-variables. It is natural to ask why the integral (1.3) is interesting. The answer we have is the following:

(1) First of all, let us take $f(X) = 1$. Then the *asymptotic* expansion of $Z_n(t, 1)$ with respect to the t-variables, after setting $t_1 = 0$ and $t_2 = -1/2$, gives the generating function of the order of the automorphism groups of arbitrary *ribbon* or *fat* graphs (Bessis-Itzykson-Zuber [2]).

(2) Under the same setting of (1), let $t_j = -1/j \cdot (i\lambda)^{j-2}$, $j \geq 3$, and take the limit $m \to \infty$. Then the asymptotic expansion of the logarithm of the integral (1.3) in λ and n (the size of the matrices) gives the generating function of the Euler characteristic of the moduli spaces $\mathfrak{M}_{g,s}$ of algebraic curves of genus g with s smooth marked points for arbitrary $g \geq 0$ and $s \geq 1$ satisfying $2 - 2g - s < 0$ (Penner [3]).

(3) If $f(X)$ satisfies (1.4), then the integral $Z_n(t, f)$ is a *continuum limit* of the soliton solutions of the total hierarchy of the KP equations (the KP system). In particular, it satisfies the KP system itself (Kontsevich-Mulase-Shiota [9]).

(4) Take $f(X) = \chi_s(k_0)\chi_s(k_1) \cdots \chi_s(k_{n-1})$, where $\chi_s(k)$ is the characteristic function of the interval $(-\infty, s] \subset \mathbb{R}$. Again we set $t_1 = 0$ and $t_2 = -1/2$. Then (1.3) gives the distribution of the largest eigenvalue of a *random* Hermitian matrix X with respect to the potential $t_3X^3 + \cdots t_mX^m$. This is a very special case of the general theory established by Tracy-Widom [15]. They showed that the s-dependence of the matrix integral is governed by a nonlinear integrable system. In particular, their system reduces to the Painlevé IV for the Gaussian case (i.e., when $t_3 = \cdots = t_m = 0$). Since the deformation of the potential from 0 to $t_3X^3 + \cdots t_mX^m$ is controlled by the KP equations by (3), the distribution for an arbitrary potential can be obtained by solving the KP system with a solution of the Painlevé equation as its initial value.

Therefore, the matrix integral (1.3) connects combinatorics of graph theory, topology of the moduli spaces of algebraic curves, and the two different types of nonlinear integrable systems represented by the KP equations and the Painlevé equations, respectively. A similar relation is known by Witten [22] and Kontsevich [18] for the intersection theory of certain cohomology classes of $\mathfrak{M}_{g,s}$ via the Kontsevich integral

$$Z_n(\Lambda) = \int_{\mathcal{H}_n} \exp\left(\operatorname{trace}\left(-\frac{1}{2}X^2\Lambda + \frac{\sqrt{-1}}{6}X^3\right)\right) dX \,, \tag{1.5}$$

where Λ is a positive-definite real diagonal matrix of size n.

The purpose of this article is to give an explanation of the above (1)–(3). For the Tracy-Widom Theory, we refer to [14] and [15].

2. Feynman diagram expansions.

The key point of the connection between topology of the moduli spaces of algebraic curves and the matrix integrals of (1.3) and (1.5) is the asymptotic expansion of these integral in terms of the Feynman diagrams. The technique of Feynman diagram expansion was invented by Feynman for reducing the infinite-dimensional integral (the Feynman path integral) appeared in quantum electrodynamics to an infinite-series of finite-dimensional integrals. The infinite-series is a summation over all graphs with certain properties representing the physical process, such as the collision pattern of elementary particles.

To see why and how the graph expansion appears in the asymptotic series of an integral, let us compute a toy model. So let (x, y) denote the usual Euclidean inner product of vectors x and y of $\mathbb{R}^n$. For a function $f(x)$ on $\mathbb{R}^n$, we define

$$\langle f\rangle = \frac{\int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}(x,x)\right) f(x)\, dx}{\int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}(x,x)\right) dx}\ . \tag{2.1}$$

Then we have

$$\langle x_i x_j\rangle = \delta_{ij}\ ,$$

which picks up the coefficient of $x_i x_j$ in the quadratic form (x, x) used in (2.1). Wick's Lemma tells us

$$\langle x_{j_1} x_{j_2} \cdots x_{j_{2m-1}} x_{j_{2m}}\rangle = \frac{1}{2^m \cdot m!} \sum_{\sigma\in\mathfrak{S}_{2m}} \langle x_{j_{\sigma(1)}} x_{j_{\sigma(2)}}\rangle \cdots \langle x_{j_{\sigma(2m-1)}} x_{j_{\sigma(2m)}}\rangle\ ,$$

where the summation runs over the symmetric group $\mathfrak{S}_{2m}$ of $2m$ letters. We can interpret the sum with the normalization factor as summing over all different *pairings* of indices j_1, $\cdots$, j_{2m}. For example,

$$\langle x_i x_j x_k x_\ell\rangle = \delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk} \tag{2.2}$$

shows the three different parings of indices i, j, k and ℓ. Let us draw a little cross sign $+$ with each of the 4 lines labeled by these indices. Using (2.2) we can connect the 4 endpoints of the cross $+$ to make a shape like ∞. All three parings give different *indexed* 4-valent graphs with only one vertex at the center, but topologically they are the same graph ∞. (When we say a *graph* in this article, we do not allow any vertex of valency less than 3.)

Similarly, the quantity $\langle x_i x_j x_k x_\ell x_p x_q x_r x_s\rangle$ tells how to connect the 8 endpoints of the two crosses $++$ indexed by (i, j, k, ℓ) and (p, q, r, s). Thus we obtain indexed 4-valent graphs with two vertices. Note that these graphs are not necessarily topologically connected.

Let's study the integral (2.1) further for $n = 1$. By definition, we have $\langle x^4 \rangle = 3$, which is the number of indexed 4-valent graphs with one vertex. To obtain the number of *un-indexed* graphs, we have to divide $\langle x^4 \rangle$ by 4!, but we then get 1/8. What is this 8? It turns out that this is the order of the automorphism group of the figure ∞: we can rotate the figure 180° around the center, and flip each circle up and down independently. In this way we can establish a general formula

$$\left\langle \exp\left(\frac{t}{4!}x^4\right)\right\rangle = \sum_{m=0}^{\infty}\sum_{\Gamma}\frac{t^m}{\#\mathrm{Aut}(\Gamma)}\ , \tag{2.3}$$

where the second summation is taken over all 4-valent graphs with m vertices.

It was G. 'tHooft who discovered that replacing the integral (2.1) by a Hermitian matrix integral forces the graphs to be drawn on oriented surfaces. To see this, let us define

$$\langle f(X)\rangle = \frac{\int_{\mathcal{H}_n}\exp\left(-\frac{1}{2}\mathrm{trace}\ X^2\right)f(X)\,dX}{\int_{\mathcal{H}_n}\exp\left(-\frac{1}{2}\mathrm{trace}\ X^2\right)dX}\ , \tag{2.4}$$

where $f(X)$ is a function on $\mathcal{H}_n$. Let x_{ij} be the ij-entry of the matrix X. Since the quadratic form we use in (2.4) is

$$\mathrm{trace}\ X^2 = \sum_{i,j} x_{ij}x_{ji}\ ,$$

we have

$$\langle x_{ij}x_{k\ell}\rangle = \delta_{i\ell}\delta_{jk}\ .$$

Therefore,

$$\begin{aligned}
\left\langle \frac{1}{4}\mathrm{trace}\ X^4\right\rangle &= \left\langle \frac{1}{4}\sum_{i,j,k,\ell} x_{ij}x_{jk}x_{k\ell}x_{\ell i}\right\rangle\\
&= \frac{1}{4}\sum_{i,j,k,\ell}\big(\langle x_{ij}x_{jk}\rangle\langle x_{k\ell}x_{\ell i}\rangle + \langle x_{ij}x_{k\ell}\rangle\langle x_{jk}x_{\ell i}\rangle + \langle x_{ij}x_{\ell i}\rangle\langle x_{jk}x_{k\ell}\rangle\big)\\
&= \frac{1}{4}\sum_{i,j,k,\ell}\big(\delta_{ik}\delta_{jj}\delta_{ki}\delta_{\ell\ell} + \delta_{i\ell}\delta_{kj}\delta_{ji}\delta_{\ell k} + \delta_{ii}\delta_{j\ell}\delta_{\ell j}\delta_{kk}\big)\\
&= \frac{1}{4}(2n^3+n)\ .
\end{aligned}$$

To illustrate the combinatorics of the graph counting, Penner introduces the word *fat graph* for matrix integrals. We consider a **fat**, oriented cross sign $+$, with each of the 4 lines labeled by a double index, say ij, jk, $k\ell$ and ℓi. The orientation can be indicated by assigning a directional arrow to each index. So we choose outward arrow for the first index and inward arrow for the second index. Then the first term of the third line of the above computation says that we have to connect i-out to k-in, j-out to j-in, k-out to i-in, and ℓ-out to ℓ-in. The result is of course a **fat**, oriented ∞. Note that it has three indexed circuits: the $i = k$-circuit, the j-circuit and the ℓ-circuit. If we attach an oriented disk to each circuit, then we obtain an oriented sphere. In other words, our graph ∞ is drawn on a sphere.

Similarly, the second term of the third line of the above computation says that we have to connect i-out to ℓ-in, k-out to j-in, j-out to i-in and ℓ-out to k-in. Then the resulting fat graph has only one circuit labeled by $i = j = k = \ell$. If we attach an oriented disk to this circuit, then we obtain an oriented torus. Therefore, our graph ∞ is drawn on a torus.

Conversely, if we have a graph drawn on an oriented surface, then we can determine a unique *fattening* of the graph. Combinatorially, a fat graph is a graph with a cyclic ordering of edges at each vertex. Kontsevich called this object a *ribbon graph*.

Now we can see

$$\left\langle \exp\left(\frac{t}{4}\text{trace } X^4\right)\right\rangle = \sum_{m\geq 0}\sum_{\Gamma}\frac{n^{s(\Gamma)}}{\#\text{Aut}(\Gamma)}\cdot t^m ,$$

where the second sum runs over all ribbon graphs with m vertices, and $s(\Gamma)$ is the number of indexed circuits of a ribbon graph Γ. We are almost done to explain Statement (1) of Section 1:

$$\left\langle \exp\left(\sum_{j\geq 3} t_j \cdot \text{trace } X^j\right)\right\rangle = \sum_{m=0}^{\infty}\sum_{m_3+m_4+\cdots=m}\sum_{\Gamma}\frac{n^{s(\Gamma)}}{\#\text{Aut}(\Gamma)}\cdot\prod_{j\geq 3}(jt_j)^{m_j} . \tag{2.6}$$

This time the graph summation is taken over all ribbon graphs with m vertices, where m_j of them are j-valent. Except for the normalization factor, which depends on the matrix size n, (2.6) is the same as (1.3) with $t_1 = 0, t_2 = -\frac{1}{2}$, and $f(X) = 1$.

3. The Penner model.

A compact Riemann surface of genus $g > 1$ has three different appearances: (i) as a nonsingular algebraic curve, (ii) as a complex manifold of dimension 1, and (iii) as a hyperbolic surface with constant negative curvature. According to each of the three different points of view, there are several different ways of presenting the *moduli space* $\mathfrak{M}_g$ of these structures [16]. In this paper, we use the point of view of complex geometry, and define $\mathfrak{M}_{g,s}$ as the moduli space of nonsingular complex manifolds of dimension 1 and genus g (a *complex curve*) with s marked points. To find the connection between these moduli spaces and the ribbon graphs discussed in the previous section, we need a classical theory of *quadratic differentials*.

We are interested in a special quadratic differential on a complex curve C with a pole of order 2 at each marked point. A holomorphic quadratic differential is a holomorphic section of the line bundle $K^{\otimes 2}$, where $K = T^*C$ is the cotangent bundle (= the canonical bundle) of C. Let us denote by $(x_1, x_2, \cdots, x_s)$ the (ordered) set of s marked points of C. The quadratic differential we are interested in is an element of

$$H^0\left(C, K^{\otimes 2} \otimes \bigotimes_{j=1}^{s} \mathcal{O}(2x_j)\right) . \tag{3.1}$$

We want to choose a unique element of this cohomology group, but of course it has dimension $3g - 3 + 2s$. So we need a very strong condition to reduce the dimension to 0.

With respect to a local coordinate z of C, a quadratic differential q is represented locally by $q = h(z)dz^{\otimes 2}$, where $h(z)$ is a holomorphic function. A parameterized real curve

$$\gamma : I \longrightarrow C$$

on C defined on an interval $I \subset \mathbb{R}$ is called a *horizontal trajectory* of a quadratic differential q if

$$h(\gamma(t))\left(\frac{d\gamma(t)}{dt}\right)^2 > 0 \tag{3.2}$$

for all $t \in I$. The local expression of q changes, under a coordinate transformation $z = z(w)$, following the rule

$$q = h(z)dz^{\otimes 2} = h(z(w))\left(\frac{dz}{dw}\right)^2 dw^{\otimes 2} . \tag{3.3}$$

Take a point z_0 such that $q(z_0) \neq 0$, and define a local coordinate around z_0 by

$$w = w(z) = \int_{z_0}^{z} \sqrt{h(z)}\, dz . \tag{3.4}$$

Since $h(z) \neq 0$ near z_0, w is a holomorphic local coordinate around z_0, which is called the *canonical coordinate*. Because of (3.3), the local expression of a quadratic differential becomes

$$q = dw^{\otimes 2}$$

in the canonical coordinate (3.4). The horizontal trajectories in this coordinate are simply the horizontal lines $\gamma(t) = t + ic$ for $c \in \mathbb{R}$. For a generic element of the cohomology (3.1), horizontal trajectories behave wildly on C. So we impose the following conditions on $q = h(z)dz^{\otimes 2}$:

(1) A generic horizontal trajectory of q is a closed real curve on C,
(2) nonclosed horizontal trajectories sweep out a measure zero subset of C, and
(3) for a small simple loop α_j around the marked point x_j with the compatible orientation, we have

$$\frac{1}{2\pi i}\int_{\alpha_j} \sqrt{h(z)}\, dz = a_j > 0 ,$$

where a_j is a prescribed positive number.

To see the implication of the condition (3), let us study an example. Consider

$$q = -\frac{a^2}{z^2}\, dz^{\otimes 2}$$

around the origin with $a > 0$. Then a family

$$z = \gamma(t) = re^{it}, \quad r > 0$$

of concentric circles gives horizontal trajectories, because we have

$$h(\gamma(t))\left(\frac{d\gamma(t)}{dt}\right)^2 = -\frac{a^2}{r^2}e^{-2it}(ire^{it})^2 = a^2 > 0\ .$$

Note that the number a_j is determined by the quadratic differential itself and does not depend on the choice of a loop α_j. This is called the *perimeter* of the punctured disk around the point x_j.

Let us study one more example to see the effect of a zero of a quadratic differential. Consider, this time, a quadratic differential with a zero of order m at the origin:

$$q = z^m\, dz^{\otimes 2}\ .$$

Then the $m+2$ radial rays

$$\gamma(t) = \exp\left(\frac{2\pi i}{m+2}\right)\cdot t\ , \quad t > 0,$$

are the only horizontal trajectories coming out of the zero of q. Now we have

Strebel's Theorem [30]**.** *Let C be a nonsingular complex curve of genus g with s marked points $x_1, \cdots, x_s$, such that $2 - 2g - s < 0$ and $s > 0$. Choose a set of s arbitrary positive numbers $a_1, \cdots, a_s$. Then there exists a unique element q of the cohomology* (3.1) *satisfying the three conditions listed above.*

The amazing consequence of the theorem is that the union of the horizontal trajectories of q passing through its zeros (the *critical trajectory*) is a connected ribbon graph Γ on C. The vertices of this graph are the zeros of q and the edges are the horizontal trajectories connecting the zeros. The graph has a vertex of $m + 2 \geq 3$ valency at each zero of order $m \geq 1$. Because of the positivity condition (3.2), we can define the length of an edge by

$$\int \sqrt{h(\gamma(t))}\ \gamma'(t)\ dt\ ,$$

where we choose an orientation of the edge so that its length is positive. Thus we obtain a unique *metrized* ribbon graph Γ out of the data C, $(x_1, \cdots, x_s)$ and $(a_1, \cdots, a_s)$. Since the complement of the critical trajectory of C is a disjoint union of s open disks with perimeter a_j, the number of closed circuits of Γ is s.

Let Γ be a connected ribbon graph. We denote by $v(\Gamma)$ the number of vertices of the underlying graph and $e(\Gamma)$ the number of edges of the underlying graph. As before, $s(\Gamma)$ is the number of closed circuits. As we have seen in the previous section, we can attach s oriented disks to Γ to make it into a compact oriented surface. We can determine the genus g of the surface by

$$v(\Gamma) - e(\Gamma) + s(\Gamma) = 2 - 2g \ .$$

The condition $2-2g-s < 0$ we need in Strebel's Theorem is equivalent to say $v(\Gamma) < e(\Gamma)$, which is necessary to have a graph with no vertices of valency less than 3.

Let us denote by $RG^{met}_{g,s}$ the space of all metrized connected ribbon graphs with s closed circuits which are drawn on an oriented surface of genus g. The Strebel theory gives us a canonical map

$$\mathbb{R}^s_+ \times \mathfrak{M}_{g,s} \longrightarrow RG^{met}_{g,s} \ . \tag{3.5}$$

It is known that (3.5) gives an isomorphism of orbifolds (Penner [20], Kontsevich [18]).

Since our matrix integral (2.6) counts all ribbon graphs that are not necessarily connected, we need to find a generating function for connected graphs. So let c_m be the number of connected graphs with m vertices, and a_m the number of arbitrary graphs with m vertices. If a graph with m vertices consists of m_j connected graphs with j vertices for each $j = 1, 2, 3, \cdots$, then we have

$$a_m = \sum_{m_1+2m_2+3m_3+\cdots=m} \frac{c_1^{m_1} \cdot c_2^{m_2} \cdot c_3^{m_3} \cdot \cdots}{m_1! \cdot m_2! \cdot m_3! \cdot \cdots} \ . \tag{3.6}$$

Therefore, if we define

$$Z(t) = \sum_{m=0}^{\infty} a_m \cdot t^m$$
$$F(t) = \sum_{m=1}^{\infty} c_m \cdot t^m \ ,$$

then they satisfy

$$F(t) = \log Z(t) \ .$$

Applying this consideration to our situation, we obtain

$$\log \left\langle \exp \left(\sum_{j\geq 3} t_j \cdot \text{trace}\, X^j \right) \right\rangle = \sum_{v\geq 1} \sum_{v_3+v_4+\cdots=v} \sum_{\Gamma} \frac{n^{s(\Gamma)}}{\#\text{Aut}(\Gamma)} \cdot \prod_{j\geq 3} (jt_j)^{v_j} \ ,$$

where the third summation runs over all *connected* ribbon graphs with v vertices out of which v_j of them are j-valent. At this stage, Penner substitutes

$$t_j = -\frac{1}{j}\,(i\lambda)^{j-2}, \quad j = 3,4,5,\cdots .$$

Then we have
(3.7)

$$\begin{aligned}\log\left\langle \exp\left(\sum_{j\geq 3} t_j \cdot \text{trace } X^j\right)\right\rangle &= \sum_{v\geq 1}\sum_{v_3+v_4+\cdots=v}\sum_{\Gamma}\frac{n^{s(\Gamma)}}{\#\text{Aut}(\Gamma)}\cdot\prod_{j\geq 3}(-1)^{v_j}(i\lambda)^{v_j\cdot(j-2)}\\ &= \sum_{\Gamma}\frac{n^{s(\Gamma)}}{\#\text{Aut}(\Gamma)}\cdot(-1)^{v(\Gamma)}(i\lambda)^{2e(\Gamma)-2v(\Gamma)}\\ &= \sum_{\Gamma}\frac{n^{s(\Gamma)}}{\#\text{Aut}(\Gamma)}\cdot(-1)^{v(\Gamma)-\chi(\Gamma)}\lambda^{-2\chi(\Gamma)}\\ &= \sum_{\Gamma}\frac{(-1)^{e(\Gamma)}}{\#\text{Aut}(\Gamma)}\cdot\lambda^{-2\chi(\Gamma)}n^{s(\Gamma)} ,\end{aligned}$$

where we used a formula $3v_3+4v_4+5v_5+\cdots = 2e(\Gamma)$, and $\chi(\Gamma) = v(\Gamma)-e(\Gamma)$ is the Euler characteristic of the topological underlying graph of a ribbon graph Γ. Of course in the last three lines of the above formula, the summation is taken over all nonempty, connected ribbon graphs Γ.

To associate (3.7) with topology of the moduli spaces $\mathfrak{M}_{g,s}$, we use the stratification of the space $RG^{met}_{g,s}$. Since the metric of a ribbon graph lives on its edges, each stratum of $RG^{met}_{g,s}$ containing Γ has dimension $e(\Gamma)$. Actually, the stratum is equal to

$$\mathbb{R}_+^{e(\Gamma)}/\text{Aut}(\Gamma) . \tag{3.8}$$

Thurston [21] introduced the notion of Euler characteristic of an *orbifold* like (3.8). It is defined by

$$\chi\left(\mathbb{R}_+^{e(\Gamma)}/\text{Aut}(\Gamma)\right) = \frac{(-1)^{e(\Gamma)}}{\#\text{Aut}(\Gamma)} .$$

Since

$$RG^{met}_{g,s} = \coprod_{\Gamma;\ s(\Gamma)=s,\chi(\Gamma)=2-2g-s}\frac{\mathbb{R}_+^{e(\Gamma)}}{\#\text{Aut}(\Gamma)} ,$$

we compute

$$\chi\left(RG^{met}_{g,s}\right) = \sum_{\Gamma;\ s(\Gamma)=s,\chi(\Gamma)=2-2g-s}\frac{(-1)^{e(\Gamma)}}{\#\text{Aut}(\Gamma)} .$$

We have to note that Thurston's Euler characteristic is not a topological invariant. It is an invariant of the orbifold structure, which is an algebraic notion. Since $\chi(\Gamma) = 2-2g-s(\Gamma)$, we conclude that

$$\log\left\langle \exp\left(\sum_{j\geq 3}\left(-\frac{1}{j}\,(i\lambda)^{j-2}\right)\cdot \text{trace}\, X^j\right)\right\rangle = \sum_{g\geq 0, s\geq 1; 2-2g-s<0} \chi\left(RG_{g,s}^{met}\right)\cdot \lambda^{-4+4g+2s}\cdot n^s\ .$$

Using the Strebel theory and the fact that

$$\chi\big(\mathbb{R}_+^s\big) = (-1)^s\ ,$$

we obtain the generating function of the Euler characteristic of the moduli spaces $\mathfrak{M}_{g,s}$:

$$\log\left(\frac{\int \exp\left(-\frac{1}{2}\,\text{trace}\, X^2\right)\cdot \exp\left(\sum_{j\geq 3}\left(-\frac{1}{j}\,(i\lambda)^{j-2}\right)\cdot \text{trace}\, X^j\right)\,dX}{\int \exp\left(-\frac{1}{2}\,\text{trace}\, X^2\right)\,dX}\right) \tag{3.9}$$

$$= \sum_{g\geq 0, s\geq 1; 2-2g-s<0} (-1)^s \chi\left(\mathfrak{M}_{g,s}\right)\cdot \lambda^{-4+4g+2s}\cdot n^s\ .$$

4. τ-functions of the KP system.

When we hear that the matrix integral $Z_n(t,f)$ of (1.3) with an $f(X)$ of the form in (1.4) is a solution to the KP equations (i.e., a *τ-function*), we might wonder why it should be so. The explanation we are giving in this section says that it is because the matrix integral of (1.3) is a continuum limit of the *soliton solutions* of the KP equations.

Among the several definitions of τ-functions, we use the one given by the *bosonization* in this paper. This definition makes it clear that the KP equations that characterize the τ-functions are equivalent to the classical Plücker relations. Let

$$V = \mathbb{C}((z))$$

denote the field of formal Laurent series in z with finite order poles at $z=0$. We use

$$V^{(n)} = z^{-n}\cdot \mathbb{C}[[z]], \quad n\in\mathbb{Z}$$

as a basis for open subsets of V around 0 to introduce the Krull topology in V. For every vector subspace $W\subset V$, let γ_W denote the composition of natural linear maps

$$\gamma_W : W \longrightarrow V \longrightarrow V/V^{(-1)} \cong \mathbb{C}[z^{-1}]\ .$$

The *Grassmannian Gr* is defined to be the space of all closed vector subspaces $W\subset V$ such that the map γ_W is Fredholm of index 0. An ordered set $(v_0, v_1, v_2, \cdots)$ of vectors in

V is said to be *admissible* if there are positive integers α and $M > \alpha$ depending on the ordered set such that

$$v_n = z^{-n} + \sum_{j=0}^{\infty} c_{-\alpha+j,-n} z^{-\alpha+j}$$

for all $n \geq M$. For every admissible ordered set $(v_0, v_1, v_2, \cdots)$, we can define a *semi-infinite wedge product*

$$v_0 \wedge v_1 \wedge v_2 \wedge \cdots \ .$$

The linear span of all admissible semi-infinite wedge products is denoted by $\bigwedge^{\frac{\infty}{2}}(V)$.

Let $W \in Gr$ be a point of the Grassmannian, which we identify with the vector subspace of V it represents. By the definition of the Grassmannian, W has an admissible basis. If we denote by $GL(W)$ the set of all base-change matrices of admissible bases of W, then we have

$$GL(W) \cong GL\big(\mathbb{C}[z^{-1}]\big) = \varinjlim GL_n \ .$$

In particular, the determinant of an element of $GL(W)$ is a well-defined finite number. Now, take an admissible basis $(w_0, w_1, w_2, \cdots)$ for $W \in Gr$. Then

$$w_0 \wedge w_1 \wedge w_2 \wedge \cdots \in \bigwedge^{\frac{\infty}{2}}(V)$$

defines a unique line depending only on the point W. Therefore, we can define an embedding, known as the *Plücker embedding*, of the Grassmannian into an infinite-dimensional projective space:

$$Gr \longrightarrow \mathbb{P}\left(\bigwedge^{\frac{\infty}{2}}(V)\right) \ .$$

To define the bosonization, let $e_n = z^{-n}$, $n \in \mathbb{Z}$, be a formal basis for V, and V^* its formal dual with the dual formal basis e_n^* defined by

$$\langle e_m^*, e_n \rangle = \delta_{mn} \ .$$

We can define the interior product

$$i(e_n^*) : \bigwedge^{\frac{\infty}{2}}(V) \longrightarrow \bigwedge^{\frac{\infty}{2}-1}(V)$$

as usual. If we apply the interior product consecutively for all e_n^*, $n \geq 0$, then we obtain a linear map

$$\cdots \cdot i(e_2^*) \cdot i(e_1^*) \cdot i(e_0^*) : \bigwedge^{\frac{\infty}{2}}(V) \longrightarrow \mathbb{C} \ .$$

Physicists use a more elegant notation:

$$\cdots \cdot i(e_2^*) \cdot i(e_1^*) \cdot i(e_0^*)(v_0 \wedge v_1 \wedge v_2 \wedge \cdots) = \langle vac | v_0 v_1 v_2 \cdots \rangle \ .$$

To introduce the *time evolution* of the KP equations, we need a ring $\mathbb{C}[[t_1, t_2, t_3, \cdots]]$ of formal power series in infinitely many variables. This is a complete topological ring with respect to the topology which is defined by taking the ideals generated by homogeneous polynomials of degree n as the basis for open neighborhoods of 0. Here, we define

$$\deg t_m = m\ .$$

In the same way we can define a ring

$$\mathbb{C}((z))[[t_1, t_2, t_3, \cdots]]$$

by assigning degree 0 to all the nonzero elements of $\mathbb{C}((z))$. Then the *time evolution operator*

$$H = \exp\left(\sum_{m=1}^{\infty} t_m \cdot z^{-m}\right) = \sum_{n=0}^{\infty} z^{-n} \cdot p_n(t) \in \mathbb{C}((z))[[t_1, t_2, t_3, \cdots]]$$

becomes a well-defined element, where

$$p_n(t) = \sum_{n_1+2n_2+3n_3+\cdots=n} \frac{t_1^{n_1} \cdot t_2^{n_2} \cdot t_3^{n_3} \cdot \cdots}{n_1! \cdot n_2! \cdot n_3! \cdot \cdots} \tag{4.1}$$

is a homogeneous polynomial of degree n. Note that we have encountered the same formula in (3.6). Let $v_0 \wedge v_1 \wedge v_2 \wedge \cdots$ be an element of $\bigwedge^{\frac{\infty}{2}}(V)$. Then the interior product of $i(e_n^*)$ can be extended linearly to the formal expression

$$H \cdot (v_0 \wedge v_1 \wedge v_2 \wedge \cdots) = Hv_0 \wedge Hv_1 \wedge Hv_2 \wedge \cdots\ .$$

It is known by Sato [27] and Date et al. [23] that assigning a quantity

$$\cdots \cdot i(e_2^*) \cdot i(e_1^*) \cdot i(e_0^*)\big(H \cdot (v_0 \wedge v_1 \wedge v_2 \wedge \cdots)\big) = \langle vac|H|v_0 v_1 v_2 \cdots\rangle$$

to $v_0 \wedge v_1 \wedge v_2 \wedge \cdots$ defines a linear isomorphism

$$\beta : \bigwedge^{\frac{\infty}{2}}(V) \xrightarrow{\sim} \mathbb{C}[[t_1, t_2, t_3, \cdots]]\ .$$

This map is called the *bosonization*. We use the same notation for the projective version of the bosonization:

$$\beta : \mathbb{P}\left(\bigwedge^{\frac{\infty}{2}}(V)\right) \xrightarrow{\sim} \mathbb{P}\left(\mathbb{C}[[t_1, t_2, t_3, \cdots]]\right)\ .$$

Definition. [27], [23]. *The image $\beta(Gr)$ of the Grassmannian Gr under the projective bosonization β is called the space of τ-functions of the KP equations.*

The t_m-derivative of the time evolution operator H is equal to the m-th derivative in t_1. Using this fact, we can translate the Plücker relations of the Grassmannian to a set of quadratic nonlinear partial differential equations of the KP τ-functions. It is known that this set of equations coincides with the Hirota bilinear equations for the KP system (Sato [27]).

Let us give an exact formula for the τ-function $\beta(W)$ when $W \in Gr$ is spanned by $w_0, \cdots, w_{n-1} \in \mathbb{C}((z))$ and $e_m = z^{-m}$ for all $m \geq n$. Without loss of generality, we can assume that w_j's have the expansion of the form

$$w_j = \sum_{i=0}^{\infty} \xi_{ij} \cdot z^{-n-1+i} \ .$$

Then the τ-function $\tau(t, W) = \beta(W)$ corresponding to W is given by an $n \times n$ determinant

$$\det \left(\sum_{r=0}^{\infty} p_{r-i}(t) \ \xi_{rj} \right)_{ij} \ , \tag{4.2}$$

where we set $p_m(t) = 0$ for $m < 0$.

Now we can prove that the matrix integral $Z_n(t, f)$ of (1.3) with (1.4) is a τ-function. In the proof, we need a trivial lemma:

Lemma. *Let $\phi_0(k), \cdots, \phi_{n-1}(k)$ and $\psi_0(k), \cdots, \psi_{n-1}(k)$ be $2n$ arbitrary functions in k. Then*

$$\det\left[\phi_i(k_\ell)\right] \cdot \det\left[\psi_j(k_\ell)\right] = \sum_{\sigma \in \mathfrak{S}_n} \det\left[\phi_i(k_{\sigma(j)}) \cdot \psi_j(k_{\sigma(j)})\right] \ ,$$

where σ runs over all permutations of $\mathfrak{S}_n$.

To prove the lemma, we calculate the left-hand side by the usual product formula of the determinant. Then it becomes a summation of n^n terms. Because of the multilinearity of the determinants, only $n!$ of these terms are nonzero. Rearranging the $n!$ terms, we obtain the above formula. In what follows, we use this formula for $\psi_j(k) = k^j$. Since our $f(X)$ is a symmetric function, we can use the classical formula (see Mehta [13])

$$\int_{\mathcal{H}_n} \exp\left(\text{trace} \sum_{\alpha=1}^{m} t_\alpha X^\alpha \right) f(X) \, dX$$

$$= c(n) \int_{\mathbb{R}^n} \exp\left(\sum_{i=0}^{n-1} \sum_{\alpha=1}^{m} t_\alpha \ k_i^\alpha \right) \Delta(k_0, \cdots, k_{n-1})^2 f(k_0, \cdots, k_{n-1}) \, dk_0 \cdots dk_{n-1} \ ,$$

where $c(n)$ is a constant depending on n, which is the volume of the generic orbit of the $U(n)$-action on $\mathcal{H}_n$. For

$$f(X) = f(k_0, \cdots, k_{n-1}) = \frac{\det(\phi_j(k_i))}{\Delta(k_0, \cdots, k_{n-1})} ,$$

we have
(4.3)

$$\begin{aligned}
Z_n(t,f) &= \int_{\mathcal{H}_n} \exp\left(\operatorname{trace}\sum_{\alpha=1}^{m} t_\alpha X^\alpha\right) f(X)\cdot dX \\
&= c(n)\int_{\mathbb{R}^n} \exp\left(\sum_{i=0}^{n-1}\sum_{\alpha=1}^{m} t_\alpha\ k_i^\alpha\right) \Delta(k_0,\cdots,k_{n-1}) \det\left(\phi_j(k_i)\right)_{ij} dk_0\cdots dk_{n-1} \\
&= c(n)\int_{\mathbb{R}^n} \exp\left(\sum_{i=0}^{n-1}\sum_{\alpha=1}^{m} t_\alpha\ k_i^\alpha\right) \sum_{\sigma\in\mathfrak{S}_n} \det\left(\phi_j(k_{\sigma(i)})k_{\sigma(i)}^i\right)_{ij} dk_0\cdots dk_{n-1} \\
&= c(n)\int_{\mathbb{R}^n} \exp\left(\sum_{i=0}^{n-1}\sum_{\alpha=1}^{m} t_\alpha\ k_{\sigma(i)}^\alpha\right) \sum_{\sigma\in\mathfrak{S}_n} \det\left(\phi_j(k_{\sigma(i)})k_{\sigma(i)}^i\right)_{ij} dk_0\cdots dk_{n-1} \\
&= c(n)\sum_{\sigma\in\mathfrak{S}_n} \det\left(\int_{\mathbb{R}^n} \exp\left(\sum_{\alpha=1}^{m} t_\alpha\ k_{\sigma(i)}^\alpha\right)\phi_j(k_{\sigma(i)})k_{\sigma(i)}^i\right)_{ij} dk_0\cdots dk_{n-1} \\
&= n!\cdot c(n)\det\left(\int_{-\infty}^{\infty} \exp\left(\sum_{\alpha=1}^{m} t_\alpha\ k_i^\alpha\right)\phi_j(k_i)\ k_i^i\ dk_i\right)_{ij} \\
&= n!\cdot c(n)\det\left(\int_{-\infty}^{\infty} \exp\left(\sum_{\alpha=1}^{m} t_\alpha\ k^\alpha\right)\phi_j(k)\ k^i\ dk\right)_{ij} .
\end{aligned}$$

The above computation makes sense as a complex analytic function in $t_1, \cdots, t_m$ defined on a complex domain whenever the integrals converge. Now we use the formula

$$\exp\left(\sum_{\alpha=1}^{m} t_\alpha\ k^\alpha\right) = \sum_{r=0}^{\infty} p_r(t)\ k^r ,$$

where $p_r(t)$ is the same as (4.1) but depending only on the first m variables $t_1, \cdots, t_m$.

Then we have

$$
\begin{aligned}
Z_n(t,f) &= n! \cdot c(n) \det\left(\int_{-\infty}^{\infty} \sum_{r=0}^{\infty} p_r(t)\ k^r\ \phi_j(k)\ k^i\ dk\right)_{ij} \\
&= n! \cdot c(n) \det\left(\int_{-\infty}^{\infty} \sum_{r=0}^{\infty} p_{r-i}(t)\ k^r\ \phi_j(k)\ dk\right)_{ij} \\
&= n! \cdot c(n) \det\left(\sum_{r=0}^{\infty} p_{r-i}(t) \int_{-\infty}^{\infty} k^r\ \phi_j(k)\ dk\right)_{ij} \\
&= \det\left(\sum_{r=0}^{\infty} p_{r-i}(t)\xi_{rj}\right)_{ij} \quad ,
\end{aligned}
$$

where we have defined

$$
\xi_{rj} = n! \cdot c(n) \int_{-\infty}^{\infty} k^r\ \phi_j(k)\ dk\ .
$$

This determinant is exactly the same as (4.2), and hence we can conclude that $Z_n(t,f)$ satisfies the KP equations with respect to the variables $t_1, \cdots, t_m$. Amazing fact is that the formula we have just established is a continuum version of the famous Hirota soliton solution of the KP equations (Sato [27]). Hirota's solution depends on $nN+N$ parameters c_{ij} and λ_i, where $0 \le i \le N-1$ and $0 \le j \le n-1$. Let

$$
\eta(t,k) = \sum_{\alpha=1}^{\infty} t_\alpha\ k^\alpha\ .
$$

Then Hirota's soliton solution is given by

$$
\sum_{0 \le i_0 < \cdots < i_{n-1} \le N-1} \exp\left(\sum_{j=0}^{n-1} \eta(t, \lambda_{i_j})\right) \Delta(\lambda_{i_0}, \cdots, \lambda_{i_{n-1}}) \det\begin{pmatrix} c_{i_0 0} & \cdots & c_{i_0 n-1} \\ \vdots & & \vdots \\ c_{i_{n-1} 0} & \cdots & c_{i_{n-1} n-1} \end{pmatrix}\ .
$$

This coincides with our $Z_n(t,f)$ in (4.3) if we take

$$
\phi_j(k) = \sum_{i=0}^{N-1} c_{ij}\ \delta(k - \lambda_i)\ .
$$

Therefore, our matrix integral $Z_n(t,f)$ is indeed a *continuum soliton solution* of the KP equations.

5. Comments.

5.1. It is pointed out in Kontsevich [18] that the formula (1.4) for our invariant function $f(X)$ itself gives a τ-function of the KP system in the time variables defined by

$$t'_j = \frac{1}{j} \cdot \text{trace } X^{-j}$$

when we take the limit $n \to \infty$. Kontsevich showed that if one introduces a large square matrix Y satisfying that

$$t_\alpha = \frac{1}{\alpha} \cdot \text{trace } Y^{-\alpha} \ ,$$

then $Z_n(t, f)$, as a function of Y, is given by

$$Z_n(Y, f) = \int \det \left(I - X \otimes Y^{-1} \right)^{-1} f(X) \, dX \ . \tag{5.1}$$

5.2. The above formula (5.1) suggests that the Hermitian matrix integral should be considered as a multi-dimensional contour integral in the n^2-dimensional complex Euclidean space. As long as the integral converges, there is no need to restrict ourselves to the Hermitian matrices. In particular, (3.9) makes sense only after we interpret the integral in terms of a suitable residue calculus. It will be, therefore, an interesting project to study these matrix integrals from the point of view of *algebraic analysis*.

5.3. From the Tracy-Widom Theory [15] one has an expression of $Z_n(t, f)$ as a Fredholm determinant $\det(1 - K)$ with a kernel function K given by

$$K(k_0, k_1) = \frac{\phi_0(k_0)\phi_1(k_1) - \phi_0(k_1)\phi_1(k_0)}{k_1 - k_0} \tag{5.2}$$

if $f(X)$ is defined by a characteristic function of a subset of $\mathbb{R}$. One notices that the form of (5.2) is exactly the same as (1.4) for $n = 2$. I do not know any meaning of this coincidence.

References

Graph theory and matrix integrals:

[1] G. 'tHooft: A planer diagram theory for strong interactions, Nuclear Phys. **B72** (1974) 461–473.

[2] D. Bessis, C. Itzykson and J. B. Zuber: Quantum field theory techniques in graphical enumeration, Advances in Appl. Math. **1** (1980) 109–157.

[3] R. C. Penner: Perturbation series and the moduli space of Riemann surfaces, J. Differ. Geom. **27** (1988) 35–53.

Matrix integrals and integrable systems:

[4] M. Adler and P. van Moerbeke: A matrix integral solution to two-dimensional W_p-gravity, Commun. Math. Phys. **147** (1992) 25–56.

[5] E. Brezin and V. Kazakow: Phys. Lett. **B236** (1990) 144; M. Douglas: Phys. Lett. **B238** (1990) 176; M. Douglas and S. Shenker: Nucl. Phys. **B335** (1990) 635; D. Gross and A. Migdal: Phys. Rev. Lett. **64** (1990) 127.

[6] M. Fukama, H. Kawai and R. Nakayama: Infinite-dimensional Grassmannian structure of two-dimensional quantum gravity, Commun. Math. Phys. **143** (1992) 371.

[7] V. Kac and A. Schwarz: Geometric interpretation of the partition function of 2D gravity, Phys. Lett. **b257** (1991) 329.

[8] S. Kharchev, A. Marshakov, A. Mironov and A. Morozov: Generalized Kontsevich model versus Toda hierarchy and discrete matrix models, Nucl. Phys. **B397** (1993) 339–378.

[9] M. Kontsevich, M. Mulase and T. Shiota: in preparation.

[10] A. S. Schwarz: On some mathematical problems of 2D-gravity and W_h-gravity, Modern Phys. Lett. **A6** (1991) 611.

[11] A. S. Schwarz: On solutions to the string equation, preprint (1992).

[12] P. van Moerbeke: Integrable foundation of string theory, preprint (1993).

Random matrices:

[13] M. L. Mehta: *Random Matrices*, 2nd Ed., Academic Press (1991).

[14] C. A. Tracy and H. Widom: Introduction to random matrices, to appear (1993).

[15] C. A. Tracy and H. Widom: Fredholm determinants, differential equations and matrix models, UCD-ITD Preprint # ITD 92/93–17 (1993).

Topology of moduli spaces:

[16] J. L. Harer: The cohomology of the moduli spaces of curves, in *Theory of Moduli*, Lecture Notes in Math., Springer-Verlag **1337** (1988) 138–221.

[17] J. Harer and D. Zagier: The Euler characteristic of the moduli space of curves, Invent. Math. **85** (1986) 457–485.

[18] M. Kontsevich: Intersection theory on the moduli space of curves and the matrix Airy function, Commun. Math. Phys. **147** (1992) 1–23.

[19] D. Mumford: Towards an enumerative geometry of the moduli spaces of curves, in *Arithmetic and Geometry*, vol. II, Ed. by M. Artin and J. Tate, Birkhäuser (1983) 271–326.

[20] R. C. Penner: The decorated Teichmüler space of punctured surfaces, Commun. Math. Phys. **113** (1987) 299–339.

[21] W. Thurston: Orbifolds and Seifert fiber spaces, Chapter 5 of the Princeton Lecture Notes (1990).

[22] E. Witten: Two dimensional gravity and intersection theory on moduli space, Surveys in Diff. Geom. **1** (1991) 243–310.

KP theory:

[23] E. Date, M. Jimbo, M. Kashiwara and T. Miwa: Transformation groups for soliton equations, in [21], World Scientific (1983) 39–120.

[24] I. M. Gel'fand and L. A. Dikii: Asymptotic behavior of the Sturm-Liouville equations and the algebra of the Korteweg-de Vries equations, Russ. Math. Surv. **30** (1975) 77-113; Fractional powers of operators and Hamiltonian systems, Func. Anal. Appl. **10** (1976) 259–273.

[25] M. Jimbo and T. Miwa: Non-linear integrable systems—Classical theory and quantum theory, World Scientific (1983).

[26] M. Mulase: Algebraic theory of the KP equations, in *Interface between Mathematics and Physics*, Ed. by S. T. Yau, International Press Company (1994) 157–223.

[27] M. Sato: Soliton equations as dynamical systems on an infinite dimensional Grassmannian manifold, Kokyuroku, Res. Inst. Math. Sci., Kyoto Univ. **439** (1981) 30–46.

[28] E. Witten: Quantum field theory, Grassmannians, and algebraic curves, Commun. Math. Phys. **113** (1988) 529–600.

Quadratic differentials:

[29] F. P. Gardiner: *Teichmüler Theory and Quadratic Differentials*, John Wiley and Sons, Wiley-Interscience Publ. (1987).

[30] K. Strebel: *Quadratic Differentials*, Springer-Verlag (1984).