Conservation laws of random matrix theory

NICHOLAS M. ERCOLANI

This paper presents an overview of the derivation and significance of recently derived conservation laws for the matrix moments of Hermitian random matrices with dominant exponential weights that may be either even or odd. This is based on a detailed asymptotic analysis of the partition function for these unitary ensembles and their scaling limits. As a particular application we derive closed form expressions for the coefficients of the genus expansion for the associated free energy in a particular class of dominant even weights. These coefficients are generating functions for enumerating g-maps, related to graphical combinatorics on Riemann surfaces. This generalizes and resolves a 30+ year old conjecture in the physics literature related to quantum gravity.

1. Introduction

We present an overview of some recent developments in the application of random matrix analysis to the *topological combinatorics* of surfaces. Such applications have a long history about which we should say a few words at the outset. The combinatorial objects of interest here are *maps*. A map is an embedding of a graph into a compact, oriented and connected surface X with the requirement that the complement of the graph in X should be a disjoint union of simply connected open sets. If the genus of X is g, this object is referred to as a *g-map*. The notion of g-maps was introduced by Tutte [1968] and his collaborators in the 1960s as part of their investigations of the four color conjecture.

In the early 1980s Bessis, Itzykson and Zuber, a group of physicists studying 't Hooft's diagrammatic approaches to large-N expansions in quantum field theory, discovered a profound connection between the problem of enumerating g-maps and random matrix theory [Bessis et al. 1980]. That seminal work was the basis for bringing asymptotic analytical methods into the study of maps and other related combinatorial problems.

Subsequently other physicists [Douglas and Shenker 1990; Gross and Migdal 1990] realized that the matrix model diagrammatics described by Bessis et al. provide a natural means for discretizing the Einstein–Hilbert action in two dimensions. From that and a formal double scaling limit, they were able to put

forward a candidate for so-called *2D quantum gravity*. This generated a great deal of interest in the emerging field of string theory. We refer to [Di Francesco et al. 1995] for a systematic review of this activity and to [Mariño 2005] for a description of more recent developments related to topological string theory.

All of these applications were based on the postulated existence of a $1/n^2$ asymptotic expansion of the free energy associated to the random matrix partition function, where n denotes the size of the matrix, as n becomes large. The combinatorial significance of this expansion is that the coefficient of $1/n^{2g}$ should be the generating function for the enumeration of g-maps (ordered by the cardinality of the map's vertices). In [Ercolani and McLaughlin 2003] the existence of this asymptotic expansion and several of its important analytical properties were rigorously established. This analysis was based on a Riemann–Hilbert problem originally introduced by Fokas, Its and Kitaev [Fokas et al. 1992] to study the 2D gravity problem.

The aim of this paper is to outline how the results of [Ercolani and McLaughlin 2003] and its sequel [Ercolani et al. 2008] have been used to gain new insights into the map enumeration problem. In particular, we will be able to prove and significantly extend a conjecture made in [Bessis et al. 1980] about the closed form structure of the generating functions for map enumeration.

Over time combinatorialists have made novel use of many tools from analysis including contour integrals and differential equations. In this work we also introduce nonlinear partial differential equations, in particular a hierarchy of conservation laws reminiscent of *Burgers equation* (see (4-53)) and the *shallow water wave equations* [Whitham 1974] (see (6-80)). This appears to make contact with the class of *differential posets* introduced by Stanley [1988] (see Remark 4.8).

2. Background

The general class of matrix ensembles we analyze has probability measures of the form

$$d\mu_{t_j} = \frac{1}{Z^{(n)}(g_s, t_j)} \exp\left\{-\frac{1}{g_s} \mathrm{Tr}[V_j(M, t_j)]\right\} dM, \qquad (2\text{-}1)$$

where

$$V_j(\lambda; t_j) = \frac{1}{2}\lambda^2 + \frac{t_j}{j}\lambda^j, \qquad (2\text{-}2)$$

defined on the space $\mathcal{H}_n$ of $n \times n$ Hermitian matrices, M, and where g_s is a positive parameter, referred to as the *string coefficient*. The normalization factor $Z^{(n)}(g_s, t_j)$, which serves to make μ_t a probability measure, is called the *partition function* of this unitary ensemble.

Remark 2.1. In previous treatments [Ercolani and McLaughlin 2003; Ercolani et al. 2008; Ercolani 2011; Ercolani and Pierce 2012], we have used the parameter $1/N$ instead of g_s. This was in keeping with notational usages in some areas of random matrix theory; however, since here we are trying to make a connection to some applications in quantum gravity, we have adopted the notation traditionally used in that context. This also is why we have scaled the time parameter t_j by $1/j$ in the first three sections of this paper.

For general polynomial weights V it is possible to establish the following fundamental asymptotic expansion [Ercolani and McLaughlin 2003; Ercolani et al. 2008] of the logarithm of the *free energy* associated to the partition function. More precisely, those papers consider weights of the form

$$V(\lambda) = \frac{1}{2}\lambda^2 + \sum_{\ell=1}^{j} \frac{t_\ell}{\ell}\lambda^\ell, \tag{2-3}$$

with j even.

We introduce a renormalized partition function, which we refer to as a *tau function* representation:

$$\tau_{n,g_s}^2(\vec{t}) = \frac{Z^{(n)}(g_s, \vec{t})}{Z^{(n)}(g_s, 0)}, \tag{2-4}$$

where $\vec{t} = (t_1, \ldots t_j) \in \mathbb{R}^j$. The principal object of interest is the *large n* asymptotic expansion of this representation for which one has the result [Ercolani and McLaughlin 2003; Ercolani et al. 2008]

$$\log \tau_{n,g_s}^2(\vec{t})$$
$$= n^2 e_0(x, \vec{t}) + e_1(x, \vec{t}) + \frac{1}{n^2}e_2(x, \vec{t}) + \cdots + \frac{1}{n^{2g-2}}e_g(x, \vec{t}) + \cdots \tag{2-5}$$

as $n \to \infty$ while $g_s \to 0$ with $x = ng_s$, called the 't Hooft parameter, held fixed. Moreover, for

$$\mathcal{T} = (1 - \epsilon, 1 + \epsilon) \times \left(\{|\vec{t}| < \delta\} \cap \{t_j > 0\}\right)$$

and for some $\epsilon > 0$, $\delta > 0$, we have:

(i) The expansion is uniformly valid on compact subsets of $\mathcal{T}$.

(ii) $e_g(x, \vec{t})$ extends to be complex analytic in

$$\mathcal{T}^{\mathbb{C}} = \left\{(x, \vec{t}) \in \mathbb{C}^{j+1} \,\middle|\, |x - 1| < \epsilon, |\vec{t}| < \delta\right\}.$$

(iii) The expansion may be differentiated term by term in $(x, \vec{t})$ with uniform error estimates as in (i).

The meaning of (i) is that for each g there is a constant, K_g, depending only on $\mathcal{T}$ and g such that

$$\left| \log \tau_{n,g_s}^2(\vec{t}) - n^2 e_0(x,\vec{t}) - \cdots - \frac{1}{n^{2g-2}} e_g(x,\vec{t}) \right| \leq \frac{K_g}{n^{2g}}$$

for $(x,\vec{t})$ in a compact subset of $\mathcal{T}$. The estimates referred to in (iii) have a similar form with τ_{n,g_s}^2 and $e_j(x,\vec{t})$ replaced by their mixed derivatives (the same derivatives in each term) and with a possibly different set of constants K_g.

Remark 2.2. These results were extended to the case where j is odd in [Ercolani and Pierce 2012].

To explain the topological significance of the $e_g(x,\vec{t})$ as generating functions, we begin with a precise definition of the objects they enumerate. A *map* Σ on a compact, oriented and connected surface X is a pair $\Sigma = (K(\Sigma), [\iota])$ where

- $K(\Sigma)$ is a connected 1-complex;

- ι is an embedding of $K(\Sigma)$ into X;

- the complement of $K(\Sigma)$ in X is a disjoint union of open cells (faces);

- the complement of the vertices in $K(\Sigma)$ is a disjoint union of open segments (edges).

Two such maps, Σ_1 and Σ_2, are isomorphic if there is an orientation-preserving homeomorphism of X to itself which maps the associated embedding of $K(\Sigma_1)$ homeomorphically to that of $K(\Sigma_2)$. When the genus of X is g one refers to the map as a *g-map*. What Bessis et al. [1980] effectively showed was that the partial derivatives of $e_g(1,\vec{t})$ evaluated at $\vec{t} = 0$ "count" a geometric quotient of a certain class of *labeled g-maps*.

As a means to reduce from enumerating these labeled g-maps to enumerating g-maps, it is natural to try taking a geometric quotient by a "relabeling group" more properly referred to as a *cartographic group* [Bauer and Itzykson 1996].

This labeling has two parts. First, the vertices of the same valence, ℓ, have an order labeling $1, \ldots, n_\ell$. Second, at each vertex one of the edges is distinguished. Given that X is oriented, this second labeling gives a unique ordering of the edges around each vertex. The fact that the coefficients of the free energy expansion (2-5) enumerate this class of labeled g-maps is a consequence of statements (i)–(iii) on the previous page, which enable one to evaluate a mixed partial derivative of e_g in terms of the *Gaussian unitary ensemble* (GUE) where the calculation of correlation functions of matrix coefficients all reduce to calculating just the quadratic correlation functions. (A precise description of how this enumeration works may be found in [Ercolani and McLaughlin 2003]).

To help fix these ideas we consider the case of a j-regular g-map (i.e., every vertex has the same valence, j) of size m (i.e., the map has m vertices) which is the main interest of this paper. The cartographic group in this case is generated by the symmetric group S_m which permutes the vertex labels and m factors of the cyclic group C_j, which rotates the distinguished edge at a given vertex in the direction of the holomorphic (counterclockwise) orientation on X. The order of the cartographic group here is the same as that of the product of its factors which is $m! j^m$. On the other hand the generating function for g-maps in this setting is given by

$$e_g(t_j) = e_g\big(x = 1, \vec{t} = (0, \dots, 0, t_j)\big) \tag{2-6}$$

$$= \sum_{m \geq 1} \frac{1}{m! j^m} (-t_j)^m \kappa_j^{(g)}(m), \tag{2-7}$$

where $\kappa_j^{(g)}(m)$ is the number of labeled j-regular g-maps on m vertices. The fractional factor in the sum perfectly cancels the order of the cartographic group, making this series appear to indeed be the ordinary generating function for pure g-maps. However, for some g-maps the cartographic action may have nontrivial isotropy and this can create an "over-cancellation" of the labeling. This happens when a particular relabeling of a given map can be transformed back to the original labeling by a diffeomorphism of the underlying Riemann surface X. In this event the two labelings are indistinguishable and the diffeomorphism induces an automorphism of the underlying map. In addition, the element of the cartographic group giving rise to this situation is an element of the isotropy group of the given map. Hence, as a generating function for the geometric quotient, (2-6) is expressible as

$$e_g(t_j) = \sum_{g\text{-maps}\,\Sigma} \frac{1}{|\mathrm{Aut}(\Sigma)|} (-t_j)^{m(\Sigma)}, \tag{2-8}$$

$$E_g(x, t_j) = e_g(x, \vec{t} = (0, \dots, 0, t_j)) \tag{2-9}$$

$$= \sum_{g\text{-maps}\,\Sigma} \frac{1}{|\mathrm{Aut}(\Sigma)|} (-t_j)^{m(\Sigma)} x^{f(\Sigma)} = x^{2-2g} e_g(x^{j/2-1} t_j),$$

where $m(\Sigma)$ is the number of vertices of Σ, $f(\Sigma)$ is the number of faces of Σ and $\mathrm{Aut}(\Sigma)$ is the automorphism group of the map Σ. We have included the x-dependent form, (2-9), of e_g since that will play an important role later on and also to observe that this is in fact a *bivariate* generating function for enumerating g-maps with a fixed number of vertices and faces. Moreover, in this j-regular setting, one sees that the bivariate function is self-similar. This is

a direct consequence of Euler's relation:

$$2 - 2g = \#\,\text{vertices} - \#\,\text{edges} + \#\,\text{faces} = m(\Sigma) - \frac{j}{2} m(\Sigma) + f(\Sigma),$$

$$t_j^{m(\Sigma)} x^{f(\Sigma)} = x^{2-2g} (t_j x^{j/2-1})^{m(\Sigma)}. \tag{2-10}$$

The presence of geometric factors such as $1/|\text{Aut}(\Sigma)|$ is not uncommon in enumerative graph theory, a classical example being that of Erdős–Rényi graphs [Janson et al. 1993]. In the quantum gravity setting these factors also have a natural interpretation in terms of the discretization of the reduction to conformal structures via a quotient of metrics by the action of the diffeomorphism group. We refer to [Di Francesco et al. 1995; Bauer and Itzykson 1996] for further details on this attractive set of ideas.

In [Bessis et al. 1980], e_0, e_1 and e_2 were explicitly computed for the case of valence $j = 4$. We quote, from the same paper, the following conjecture (some notation has been changed to be consistent with ours):

It would of course be very interesting to obtain $e_g(t_4)$ in closed form for any value of g. The method of this paper enabled us to do so up to $g = 2$, but works in the general case, although it requires an increasing amount of work. We conjecture a general expression of the form

$$e_g = \frac{(1-z_0)^{2g-1}}{(2-z_0)^{5(g-1)}} P^{(g)}(z_0), \quad g \geq 2,$$

with $P^{(g)}$ a polynomial in z_0, the degree of which could be obtained by a careful analysis of the above procedure.

Here $z_0 = z_0(t_4)$ is equal, up to a scaling, to the generating function for the Catalan numbers; below it will signify $z_0(t_{2\nu})$, which is similarly related to the generating function for the higher Catalan numbers (4-62).

The main purpose of this paper is to show how this conjecture can be verified and significantly extended. In particular, we will show, for the case of even valence, $j = 2\nu$:

Theorem 2.3. *For $g \geq 2$ and for $\nu \geq 2$,*

$$e_g(z_0) = C^{(g)} + \frac{c_0^{(g)}(\nu)}{(\nu - (\nu-1)z_0)^{2g-2}} + \cdots + \frac{c_{3g-3}^{(g)}(\nu)}{(\nu - (\nu-1)z_0)^{5g-5}} \tag{2-11}$$

$$= \frac{(z_0 - 1)^r Q_{5g-5-r}(z_0)}{(\nu - (\nu-1)z_0)^{5g-5}}, \tag{2-12}$$

$$r = \max\left\{ 1, \left\lfloor \frac{2g-1}{\nu-1} \right\rfloor \right\}, \tag{2-13}$$

The top coefficient and the constant term are respectively given by

$$c_{3g-3}^{(g)}(\nu) = \frac{1}{(5g-5)(5g-3)\nu^2} a_{3g-1}^{(g)}(\nu) \neq 0, \tag{2-14}$$

$$C^{(g)} = -2(2g-3)! \left[\frac{1}{(2g+2)!} - \frac{1}{(2g)!\,12} \right.$$
$$\left. + \frac{(1-\delta_{2,g})}{(2g-1)!} \sum_{k=2}^{g-1} \frac{(2-2k)_{2g-2k+2}}{(2g-2k+2)!} C^{(k)} \right], \tag{2-15}$$

where $a_{3g-1}^{(g)}(\nu)$ *(see Theorem 4.7) is proportional to the g-th coefficient in the asymptotic expansion at infinity of the ν-th equation in the Painlevé I hierarchy* [Ercolani 2011] *and* $(r)_m = r(r-1)\ldots(r-m+1)$. (Explicit expressions for e_1 and e_0 are given in (5-70) and (5-71).)

Our methods can be extended to the case of j odd and the derivation of the analogue to Theorem 2.3 is in progress (see Section 6).

The route to getting these results passes through nonlinear PDE, in particular a class of nonlinear evolution equations known as conservation laws which come from studying scaling limits of the recursion operators for orthogonal polynomials whose weights match those of the matrix models.

This appeal to orthogonal polynomials also motivated the approaches of [Bessis et al. 1980; Douglas and Shenker 1990]. However to give a rigorous *and* effective treatment to the problem of finding closed form expressions for the coefficients of the asymptotic free energy, (2-5), requires essential use of Riemann–Hilbert analysis on the Riemann–Hilbert problem for orthogonal polynomials that was introduced in [Fokas et al. 1992]. Though we will not review this analysis here, we will state the consequences of it needed for our applications and reference their sources.

In Section 3 we present the necessary background on orthogonal polynomials and introduce the main equations governing their recurrence operators: the *difference string equations* and the *Toda lattice equations*. In Section 4 we describe how (2-5) can be used to derive and solve (in the case of even valence) the continuum limits of these equations which relates to the nonlinear evolution equations alluded to earlier. In Section 5 we outline the proof Theorem 2.3 and in Section 6 we describe the extension of this program to the case of odd valence and briefly mention what has been accomplished in that case thus far. This will also help to illuminate the full picture behind the idea of conservation laws for random matrices.

Over the years there have been a number of efforts to systematically address the question of graphical enumeration on Riemann surfaces by studying the

resolvent of the random matrix and associated Schwinger–Dyson equations [Ambjørn et al. 1993; Chekhov and Eynard 2006; Eynard 2004; Eynard and Orantin 2009]. These approaches have led to interesting direct extensions of the equations that Tutte originally introduced to study maps. They also have particular relevance for the topics described in the concluding remark Section 7.2. Our methods take a different approach, based on orthogonal polynomials, but the development of a unified perspective that incorporates the Toda, string and Schwinger–Dyson equations would undoubtedly yield valuable insights.

3. The role of orthogonal polynomials and their asymptotics

Let us recall the classical relation between orthogonal polynomials and the space of square-integrable functions on the real line, $\mathbb{R}$, with respect to exponentially weighted measures. In particular, we want to focus attention on weights that correspond to the random matrix weights $V(\lambda)$, (2-2), with j even. (Recently this relation has been extended to the cases of j odd [Bleher and Deaño 2012; Ercolani and Pierce 2012], with the orthogonal polynomials generalized to the class of so-called *non-Hermitian* orthogonal polynomials; however, for this exposition we will stick primarily with the even case.) To that end we consider the Hilbert space $H = L^2\big(\mathbb{R}, e^{-g_s^{-1}V(\lambda)}\big)$ of weighted square integrable functions. This space has a natural polynomial basis, $\{\pi_n(\lambda)\}$, determined by the conditions that

$$\pi_n(\lambda) = \lambda^n + \text{ lower order terms,}$$
$$\int \pi_n(\lambda)\pi_m(\lambda)e^{-g_s^{-1}V(\lambda)}d\lambda = 0 \quad \text{for } n \neq m.$$

For the construction of this basis and related details we refer the reader to [Deift 1999].

With respect to this basis, the operator of multiplication by λ is representable as a semiinfinite tridiagonal matrix:

$$\mathcal{L} = \begin{pmatrix} a_0 & 1 & & \\ b_1^2 & a_1 & 1 & \\ & b_2^2 & a_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}. \tag{3-16}$$

We commonly refer to $\mathcal{L}$ as the *recursion operator* for the orthogonal polynomials and to its entries as *recursion coefficients*. (When V is an even potential, it follows from symmetry that $a_j = 0$ for all j.) We remark that often a basis of orthonormal, rather than monic orthogonal, polynomials is used to make this representation. In that case the analogue of (3-16) is a symmetric tridiagonal matrix. As long

as the coefficients $\{b_n\}$ do not vanish, these two matrix representations can be related through conjugation by a semiinfinite diagonal matrix of the form $\mathrm{diag}\big(1, b_1^{-1}, (b_1 b_2)^{-1}, (b_1 b_2 b_3)^{-1}, \dots\big)$.

Similarly, the operator of differentiation with respect to λ, which is densely defined on H, has a semiinfinite matrix representation, $\mathcal{D}$, that can be expressed in terms of $\mathcal{L}$ as

$$\mathcal{D} = \frac{1}{g_s}(\mathcal{L} + t\mathcal{L}^{j-1})_-, \tag{3-17}$$

where the minus subscript denotes projection onto the strictly lower part of the matrix.

From the canonical (Heisenberg) relation on H, one sees that

$$[\partial_\lambda, \lambda] = 1,$$

where here λ in the bracket and 1 on the right-hand side are regarded as multiplication operators. With respect to the basis of orthogonal polynomials this may be reexpressed as

$$\big[\mathcal{L}, (\mathcal{L} + t\mathcal{L}^{j-1})_-\big] = g_s I. \tag{3-18}$$

The relations implicit in (3-18) have been referred to as *string equations* in the physics literature. In fact the relations that one has, row by row, in (3-18) are actually successive differences of consecutive string equations in the usual sense. However, by continuing back to the first row one may recursively decouple these differences to get the usual equations. To make this distinction clear we will refer to the row by row equations that one has directly from (3-18) as *difference string equations*.

$\mathcal{L}$ depends smoothly on the coupling parameter t_j in the weight $V(\lambda)$; see (2-2). The explicit dependence can be determined from the fact that multiplication by λ commutes with differentiation by t_j. This yields our second fundamental relation on the recurrence coefficients,

$$g_s \frac{\partial}{\partial t_j}\mathcal{L} = [(\mathcal{L}^j)_-, \mathcal{L}], \tag{3-19}$$

which is equivalent to the j-th equation of the semiinfinite Toda lattice hierarchy. The Toda equations for $j = 1$ are

$$-g_s \frac{da_{n,g_s}}{dt_1} = b_{n+1,g_s}^2 - b_{n,g_s}^2, \tag{3-20}$$

$$-g_s \frac{db_{n,g_s}^2}{dt_1} = b_{n,g_s}^2(a_{n,g_s} - a_{n-1,g_s}). \tag{3-21}$$

Hirota equations. One may apply standard methods of orthogonal polynomial theory [Szegő 1939] to deduce the existence of a semiinfinite lower unipotent matrix A such that

$$\mathcal{L} = A^{-1}\epsilon A,$$

where

$$\epsilon = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & 0 & 0 & \ddots & \\ & & \ddots & \ddots & \end{pmatrix}.$$

(For a description of the construction of such a unipotent matrix we refer to Proposition 1 of [Ercolani and McLaughlin 2001].)

This is related to the Hankel matrix

$$\mathcal{H} = \begin{pmatrix} m_0 & m_1 & m_2 & \dots \\ m_1 & m_2 & m_3 & \dots \\ m_2 & m_3 & m_4 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where

$$m_k = \int_{\mathbb{R}} \lambda^k e^{-g_s^{-1}V(\lambda)}\, d\lambda$$

is the k-th moment of the measure, by

$$ADA^{\dagger} = \mathcal{H},$$

$$D = \mathrm{diag}\{d_0, d_1 \dots\},$$

with

$$d_n = \frac{\det \mathcal{H}_{n+1}}{\det \mathcal{H}_n},$$

where $\mathcal{H}_n$ denotes the $n \times n$ principal submatrix of $\mathcal{H}$ whose determinant may be expressed as (see [Szegő 1939]) as

$$\det \mathcal{H}_n = n!\hat{Z}^{(n)}(t_1, t_{2\nu}),$$

$$\hat{Z}^{(n)}(t_1, t_{2\nu}) = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left\{ -g_s^{-2}\left[g_s \sum_{m=1}^{n} V(\lambda_m; t_1, \ t_{2\nu}) \right.\right.$$

$$\left.\left. -g_s^2 \sum_{m \neq \ell} \log|\lambda_m - \lambda_\ell| \right] \right\} d^n\lambda, \quad \text{(3-22)}$$

where $V(\lambda; t_1, \ t_{2\nu+1}) = \frac{1}{2}\lambda^2 + t_1\lambda + \frac{t_{2\nu}}{2\nu}\lambda^{2\nu}$. We set $\det \mathcal{H}_0 = 1$.

Remark 3.1. One sometimes needs to extend the domain of the tau functions to include other parameters, such as t_1, as we have done here. Doing this presents no difficulties in the prior constructions.

The diagonal elements may in fact be expressed as $d_n = \dfrac{\tau_{n+1,g_s}^2}{\tau_{n,g_s}^2} d_n(0)$, where

$$\tau_{n,g_s}^2 = \frac{\hat{Z}^{(n)}(t_1, t_{2v})}{\hat{Z}^{(n)}(0, 0)} \tag{3-23}$$

$$= \frac{Z^{(n)}(t_1, t_{2v})}{Z^{(n)}(0, 0)}, \tag{3-24}$$

which agrees with the definition of the tau function given in (2-4). The second equality follows by reducing the unitarily invariant matrix integrals in (3-24) to their diagonalizations, which yields (3-23) [Ercolani and McLaughlin 2003]. Tracing through these connections, from $\mathcal{L}$ to D, one may derive the fundamental identity relating the random matrix partition function to the recurrence coefficients:

$$b_{n,g_s}^2 = \frac{d_n}{d_{n-1}} = \frac{\tau_{n+1,g_s}^2 \tau_{n-1,g_s}^2}{\tau_{n,g_s}^4} b_{n,g_s}^2(0), \tag{3-25}$$

which is the basis for our analysis of continuum limits in the next section. (Note that $b_{0,g_s}^2(0) = 0$ and therefore $b_{0,g_s}^2 \equiv 0$.) We will also need a differential version of this relation:

Lemma 3.2 (Hirota).

$$a_{n,g_s} = -g_s \frac{\partial}{\partial t_1} \log\left[\frac{\tau_{n+1,g_s}^2}{\tau_{n,g_s}^2}\right] = -g_s \frac{\partial}{\partial t_1} \log\left[\frac{Z^{(n+1)}(t_1, t_{2v})}{Z^{(n)}(t_1, t_{2v})}\right], \tag{3-26}$$

$$b_{n,g_s}^2 = g_s^2 \frac{\partial^2}{\partial t_1^2} \log \tau_{n,g_s}^2 = g_s^2 \frac{\partial^2}{\partial t_1^2} \log Z^{(n)}(t_1, t_{2v}). \tag{3-27}$$

(A derivation of this lemma may be found in [Bleher and Its 2005].) It follows from (3-27), (2-5) and (2-9) that:

Corollary 3.3.

$$b_{n,g_s}^2(x; t_{2v}) = x\left(z_0(x; t_{2v}) + \cdots + \frac{1}{n^{2g}} z_g(x; t_{2v}) + \cdots\right), \tag{3-28}$$

$$z_g(x; t_{2v}) = \frac{d^2}{dt_1^2} e_g(x; t_1, t_{2v})|_{t_1=0} \tag{3-29}$$

$$= x^{1-2g} z_g(x^{v-1} t_{2v}) \tag{3-30}$$

is a uniformly valid asymptotic expansion (in the sense of (iii) on page 165).

Path weights and recurrence coefficients. In order to effectively utilize the relations (3-18)–(3-19) it will be essential to keep track of how the matrix entries of powers of the recurrence operator, $\mathcal{L}^j$, depend on the original recurrence coefficients. That is best done via the combinatorics of weighted walks on the index lattice of the orthogonal polynomials. For the case of even potentials, the relevant walks are *Dyck paths* which are walks, P, on $\mathbb{Z}$ which, at each step, can either increase by 1 or decrease by 1. Set

$$\mathcal{P}^j(m_1, m_2) = \text{the set of all Dyck paths of length } j \text{ from } m_1 \text{ to } m_2. \quad (3\text{-}31)$$

Then step weights, path weights and the (m_1, m_2)-entry of $\mathcal{L}^j$ are, respectively, given by

$$\omega(p) = \begin{cases} 1 & \text{if the } p\text{-th step moves from } n \text{ to } n+1 \text{ on the lattice,} \\ b_n^2 & \text{if the } p\text{-th step moves from } n \text{ to } n-1, \end{cases}$$

$$\omega(P) = \prod_{\text{steps } p \in P} \omega(p),$$

$$\mathcal{L}^j_{m_1, m_2} = \sum_{P \in \mathcal{P}^j(m_1, m_2)} \omega(P). \quad (3\text{-}32)$$

Dyck representation of the difference string equations. The *difference string equations* are given (for the 2ν-valent case) by (3-18):

$$[\mathcal{L}, (\mathcal{L} + t\mathcal{L}^{2\nu-1})_-] = g_s I. \quad (3\text{-}33)$$

By parity considerations, when the potential V is even, the only nontautological equations come from the diagonal entries of (3-34): *the (n, n) entry* gives

$$g_s = (\mathcal{L} + t\mathcal{L}^{2\nu-1})_{n+1, n} - (\mathcal{L} + t\mathcal{L}^{2\nu-1})_{n, n-1}. \quad (3\text{-}34)$$

In terms of Dyck paths this becomes

$$\frac{x}{n} = b_{n+1}^2 - b_n^2 + t \sum_{P \in \mathcal{P}^{2\nu-1}(1,0)} \left(\prod_{m=1}^{\nu} b_{n+\ell_m(P)+1}^2 - \prod_{m=1}^{\nu} b_{n+\ell_m(P)}^2 \right), \quad (3\text{-}35)$$

where $\ell_m(P)$ denotes the lattice location of the path P after the m-th downstep and we have used the relation $x = ng_s$ on the left-hand side of the equation.

We illustrate this more concretely for the case of $j = 2\nu = 4$. Referring to (3-31), the relevant path classes here are

$\mathcal{P}^1(1, 0) = $ a descent by one step,

$\mathcal{P}^3(1, 0) = $ paths with exactly one upstep and two downsteps (Figure 1).
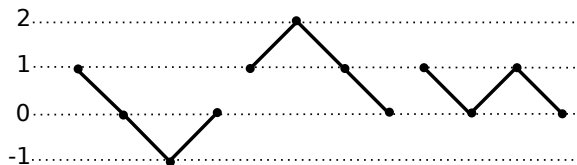
Figure 1. Elements of $\mathcal{P}^3(1,0)$.

Note that the structure of the path classes does not actually depend upon n. This is a reflection of the underlying spatial homogeneity of these equations. Thus, for the purpose of describing the path classes, one can translate n to 0.

Now applying (3-32) the difference string equation becomes, for $n > 0$,

$$\frac{1}{n} = (b_{n+1}^2 - b_n^2) + t\{b_{n+1}^2(b_n^2 + b_{n+1}^2 + b_{n+2}^2) - b_n^2(b_{n-1}^2 + b_n^2 + b_{n+1}^2)\},$$

where, for this example, we have set the parameter x equal to 1.

Dyck representation of the Toda equations. We now pass to a more explicit form of the *Toda equation* (3-19) in the case $j = 2\nu$:

$$-\frac{x}{2\nu n}\frac{db_n^2}{dt_{2\nu}} = (\mathcal{L}^{2\nu})_{n+1,n-1} - (\mathcal{L}^{2\nu})_{n,n-2}$$

$$= \sum_{P \in \mathcal{P}^{2\nu}(2,0)} \left(\prod_{m=1}^{\nu+1} b_{n+\ell_m(P)+1}^2 - \prod_{m=1}^{\nu+1} b_{n+\ell_m(P)}^2 \right).$$

Once again we illustrate these equations in the tetravalent case ($\nu = 2$). The relevant path class is

$$\mathcal{P}^4(1,-1) = \text{paths with exactly one upstep and three downsteps.}$$

Applying (3-32), the tetravalent Toda equations become

$$-\frac{1}{4n}\frac{db_n^2}{dt} = b_{n+1}^2 b_n^2(b_{n-1}^2 + b_n^2 + b_{n+1}^2 + b_{n+2}^2)$$

$$-b_n^2 b_{n-1}^2(b_{n-2}^2 + b_{n-1}^2 + b_n^2 + b_{n+1}^2), \quad (3\text{-}36)$$

where we have again used the relation $x = ng_s$ and then set the parameter $x = 1$.

4. Continuum limits

The continuum limits of the difference string and Toda equations will be described in terms of certain scalings of the independent variables, both discrete and continuous. As indicated at the outset, the positive parameter g_s sets the scale for the potential in the random matrix partition function and is taken to be small.

The discrete variable n labels the lattice *position* on $\mathbb{Z}^{\geq 0}$ that marks, for instance, the n-th orthogonal polynomial and recurrence coefficients. We also always take n to be large and in fact to be of the same order as $1/g_s$; that is, as n and g_s tend to ∞ and 0 respectively, they do so in such a way that their product

$$x \doteq g_s \, n \qquad (4\text{-}37)$$

remains fixed at a value close to 1.

In addition to the *global* or *absolute* lattice variable n, we also introduce a *local* or *relative* lattice variable denoted by k. It varies over integers but will always be taken to be small in comparison to n and independent of n. The Dyck lattice paths naturally introduce the composite discrete variable $n + k$ into the formulation of the difference string and Toda equations which we think of as a small discrete variation around a large value of n. The spatial homogeneity of those equations manifests itself in their all having the same form, independent of what n is, while k in those equations varies over $\{-\nu-1, \ldots, -1, 0, 1, \ldots, \nu+1\}$, the *bandwidth* of the (2ν)-th Toda/difference string equations. Taking $\nu + 1 \ll n$ will insure the necessary separation of scales between k and n. We define

$$w \doteq (n + k)g_s \qquad (4\text{-}38)$$

$$= x + g_s \, k = x\left(1 + \frac{k}{n}\right) \qquad (4\text{-}39)$$

as a *spatial* variation close to x which will serve as a continuous analogue of the lattice location along a Dyck path relative to the starting location of the path.

We also introduce the self-similar scalings

$$s_1 \doteq x^{-\frac{1}{2}} t_1, \qquad (4\text{-}40)$$

$$s_{2\nu} \doteq x^{\nu-1} 2\nu t_{2\nu}, \qquad (4\text{-}41)$$

$$\widetilde{w} \doteq \left(1 + \frac{k}{n}\right), \qquad (4\text{-}42)$$

that are natural given (2-10). In terms of these scalings, (3-28) may be rewritten [Ercolani et al. 2008] as

$$b_{n,g_s}^2(s_{2\nu}) = x\left(x z_0(s_{2\nu}) + \cdots + \frac{x^{1-2g}}{n^{2g}} z_g(s_{2\nu}) + \cdots\right), \qquad (4\text{-}43)$$

$$z_g(s_{2\nu}) = \frac{d^2}{ds_1^2} e_g(s_1, s_{2\nu})|_{s_1=0}, \qquad (4\text{-}44)$$

and

$$b_{n+k,g_s}^2(s_{2\nu}) = x\left(f_0(s_{2\nu}, w) + \cdots + \frac{1}{n^{2g}} f_g(s_{2\nu}, w) + \cdots\right), \qquad (4\text{-}45)$$

$$f_g(s_{2\nu}, w) = w^{1-2g} z_g(s_{2\nu}\widetilde{w}^{\nu-1}). \qquad (4\text{-}46)$$

Remark 4.1. The variables s_j as defined above differ slightly from their usage in related works [Ercolani et al. 2008; Ercolani 2011] where $s_j = -\alpha_j t_j$ for appropriate parameters $\alpha_j > 0$.

We also introduce a shorthand notation to denote the expansion of the coefficients of $f(s_1, s_{2\nu+1}, w)$ around $w = x$.

Definition 4.2. For $w = x + g_s k$,

$$f(s_1, s_{2\nu}, w) = \sum_{j=0}^{\infty} \frac{f_{w^{(j)}}|_{w=x}}{j!}\left(\frac{kx}{n}\right)^j, \qquad (4\text{-}47)$$

where the subscript $w^{(j)}$ denotes the operation of taking the j-th derivative with respect to w of each coefficient of f:

$$f_{w^{(j)}} = \sum_{g\geq 0} \frac{\partial^j}{\partial w^j} f_g(s_1, s_{2\nu}, w) \frac{1}{n^{2g}}.$$

As valid asymptotic expansions, these representations denote the asymptotic series whose successive terms are gotten by collecting all terms with a common power of $1/n$ in (4-47).

In what follows we will frequently abuse notation and drop the evaluation at $w = x$. In particular, we will write:

$$b_{n+k,g_s}^2 = \sum_{j=0}^{\infty} \frac{f_{w^{(j)}}}{j!}(g_s k)^j = \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{g\geq 0} \frac{\partial^j}{\partial w^j} f_g(s_{2\nu}, w) \frac{1}{n^{2g}}(g_s k)^j.$$
$$(4\text{-}48)$$

In doing this these series must now be regarded as formal but whose orders are still defined by collecting all terms in $1/n$ and g_s of a common order. (Recall that $g_s \sim 1/n$, so that $n^{-\alpha} g_s^{\beta} = \mathbb{O}(n^{-(\alpha+\beta)})$). They will be substituted into the difference string and the Toda equations to derive the respective continuum equations. At any point in this process, if one evaluates these expressions at $w = x$ and $g_s = x/n$ one may recover valid asymptotic expansions in which the b_{n+k,g_s}^2 have their original significance as valid asymptotic expansions of the recursion coefficients.

The continuum limit of the Toda equations. One is now in a position to study the Toda lattice equations (3-36) expanded on the asymptotic series (4-48):

$$-\frac{1}{n}\frac{d}{ds}f(s,w) =$$

$$\sum_{P \in \mathcal{P}^{2\nu}(1,-1)} \left(\prod_{m=1}^{\nu+1} \sum_{j=0}^{\infty} \frac{f_{w^{(j)}}}{j!}(g_s\,(\ell_m(P)+1))^j - \prod_{m=1}^{\nu+1} \sum_{j=0}^{\infty} \frac{f_{w^{(j)}}}{j!}(g_s\,\ell_m(P))^j \right).$$

From now on we will take $s_1 = 0$, since its role in determining the structure of the asymptotic expansions of the b_{n+k} is now completed, and set $s_{2\nu} = s$.

Collecting terms in these equations order by order in orders of $1/n$ one has a hierarchy of equations that, in principle, allows one to recursively determine the coefficients of (4-43). We will refer to this hierarchy as the *continuum Toda equations*. (Note that one has such a hierarchy for each value of ν.) Of course this is a standard procedure in perturbation theory. The equations we will derive are PDEs in the form of evolution equations in which w, now regarded as a continuous variable, is the independent *spatial* variable and $s_{2\nu}$ is the *temporal* variable. One must still determine, at each level of the hierarchy, which solution of the PDE is the one that corresponds to the expressions given for f_g in (4-46). This amounts to a kind of solvability condition. This process was carried out fully in [Ercolani et al. 2008; Ercolani 2011]. We will now state the results of that analysis.

Theorem 4.3 [Ercolani et al. 2008]. *The continuum limit, to all orders, of the Toda Lattice equations as $n \to \infty$ is given by the following infinite order partial differential equation for $f(s,w)$:*

$$-\frac{df}{ds} = F^{(\nu+1)}(g_s; f, f_w, \dots, f_{w^{(j)}}, \dots)$$

$$\doteq \sum_{g \geq 0} g_s^{2g}\, F_g^{(\nu+1)}(f, f_w, f_{w^{(2)}}, \cdots, f_{w^{(2g+1)}})$$

$$= c_\nu f^\nu f_w + g_s^2 F_1^{(\nu+1)}(f, f_w, f_{ww}, f_{www}) + \cdots \qquad (4\text{-}49)$$

for (s,w) near $(0,1)$ and initial data given by $f(0,w) = w$,

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4\text{-}50)$$

$$F_g^{(\nu+1)} = \sum_{\substack{\lambda \in \Lambda^{2g+1} \\ \ell(\lambda) \leq \nu+1}} \left(\sum_{\vec{\ell} \in \mathcal{P}^{2\nu}(1,-1)} \left(m_\lambda(\ell_1+1, \dots, \ell_{\nu+1}+1) \right. \right.$$
$$\left. \left. - m_\lambda(\ell_1, \dots, \ell_{\nu+1}) \right) \right) \frac{f_{w^\lambda}}{\lambda!}, \quad (4\text{-}51)$$

where Λ^{2g+1} denotes the set of number partitions $\lambda = (\lambda_1, \lambda_2, \dots)$ such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots$ and whose size $|\lambda| = \sum_i \lambda_i$ equals $2g + 1$; we call $\ell(\lambda) = \sum_j r_j(\lambda)$ (where $r_j(\lambda) = \#\{\lambda_i \mid \lambda_i = j\}$) the length of λ. The w-derivatives of f have been expressed in multiindex notation so that

$$f_{w^\lambda} = f_{w^{(\lambda_1)}} \cdots f_{w^{(\lambda_{\nu+1})}}.$$

Recall also that the monomial symmetric function [Macdonald 1995] *associated to the partition* λ, *appearing inside the inner sum of* (4-51), *is given by*

$$m_\lambda(x_1, \ldots, x_{\nu+1}) = \frac{1}{r(\lambda)!} \sum_{\sigma \in S_{\nu+1}} x_{\sigma(1)}^{\lambda_1} \cdots x_{\sigma(\nu+1)}^{\lambda_{\nu+1}}, \qquad (4\text{-}52)$$

where division by the factor $r(\lambda)! = \prod_j r_j(\lambda)!$ *ensures that there is no redundancy among the* (*monic*) *terms appearing in the symmetric polynomial. Finally, since a Dyck path in* $\mathscr{P}^{2\nu}(1, -1)$ *is uniquely determined by the locations* ℓ_m *of its downsteps, the inner sum of* (4-51) *is well defined.*

To help see how the expression (4-51) for the forcing term $F_g^{(\nu+1)}$ works, let us illustrate it in the simplest case where $g = 0$. Then there is only one partition, $\lambda = (1)$ in Λ^1 and

$$m_{(1)}(x_1, \ldots, x_{\nu+1}) = x_1 + \cdots + x_{\nu+1},$$
$$m_{(1)}(\ell_1 + 1, \ldots, \ell_{\nu+1} + 1) - m_{(1)}(\ell_1, \ldots, \ell_{\nu+1}) = \nu + 1.$$

Hence, $c_\nu = (\nu + 1)\binom{2\nu}{\nu+1}$.

We now introduce the global *conservation law* structure (4-53) of these continuum Toda equations.

Proposition 4.4. *Equation* (4-49) *may be rewritten as*

$$-\frac{df}{ds} = F^{(\nu+1)} = \partial_w \hat{F}^{(\nu+1)}$$
$$\doteq \partial_w \sum_{g \geq 0} g_s^{2g} \hat{F}_g^{(\nu+1)}, \qquad (4\text{-}53)$$

where

$$\hat{F}_g^{(\nu+1)}$$
$$= \sum_{\substack{\lambda \in \Lambda^{2g} \\ \ell(\lambda) \leq \nu+1}} \sum_{\vec{\ell} \in \mathscr{P}^{2\nu}(1,-1)} \left(\sum_{\mu \prec \lambda} \frac{1}{|\lambda - \mu| + 1} \binom{\lambda}{\mu} m_\mu(\ell_1, \ldots, \ell_{\nu+1}) \right) \frac{1}{r(\lambda)!} \frac{f_{w^\lambda}}{\lambda!}.$$

Here $\mu \prec \lambda$ *means* μ *is a partition such that* $\mu_1 \leq \lambda_1, \ldots, \mu_{\nu+1} \leq \lambda_{\nu+1}$ *and* $|\lambda - \mu| = \sum_i (\lambda_i - \mu_i) = |\lambda| - |\mu|$. *The combinatorial coefficient*

$$\binom{\lambda}{\mu} = \binom{\lambda_1}{\mu_1} \cdots \binom{\lambda_{\nu+1}}{\mu_{\nu+1}}$$

together with the factor $|\lambda - \mu| + 1$ *account for the multiplicities induced by* w-*differentiation* (*see Remark 4.8*).

It is straightforward to check the (key) second equality in (4-53) by direct differentiation. We can again take the special case of $g = 0$ as an example for which $\mu = \lambda = \varnothing$ and $m_\varnothing \equiv 1$, so that

$$\hat{F}_0^{(\nu+1)} = \binom{2\nu}{\nu+1} f^{\nu+1},$$

$$\partial_w \hat{F}_0^{(\nu+1)} = (\nu+1)\binom{2\nu}{\nu+1} f^\nu f_w = c_\nu f^\nu f_w = F_0^{(\nu+1)}.$$

One is now in a position to deduce the form of the Toda hierarchy. This is done by setting $x = 1$ so that $g_s = 1/n$. One then collects *all* terms of order n^{-2g} in the resulting expansion of (4-49) and this will be a partial differential equation in s and w that we refer to as the g-th equation in the continuum Toda hierarchy.

At leading order in the hierarchy one observes that, for general ν, the continuum Toda equation is an inviscid Burgers equation [Whitham 1974]:

$$\frac{d}{ds} f_0 = -\frac{c_\nu}{\nu+1} \partial_w (f_0)^{\nu+1}, \tag{4-54}$$

with initial data $f_0 = w$. A solution exists and is unique for sufficiently small values of s. It may be explicitly calculated by the method of characteristics, also known as the *hodograph* method in the version we now present. Consider the (hodograph) relation among the independent variables (s, w, f_0):

$$w = f_0 + c_\nu s f_0^\nu. \tag{4-55}$$

Lemma 4.5. *A local solution of* (4-54) *is implicitly defined by* (4-55).

Proof. The annihilator of the differential of (4-55),

$$(1 + \nu c_\nu s f_0^{\nu-1})\, df_0 - dw + c_\nu f_0^\nu\, ds,$$

is a two-dimensional distribution locally on the space (s, w, f_0). An initial curve over the w-axis (parametrized as the graph of a function $f_0(w)$), transverse to the locus where $1 + \nu c_\nu s f_0^{\nu-1} = 0$ locally determines a unique integral surface foliated by the integral curves of the *characteristic* vector field

$$\frac{df_0}{ds} = 0, \tag{4-56}$$

$$\frac{dw}{ds} = c_\nu f_0^\nu, \tag{4-57}$$

$$f_0(0, w) = f_0(w). \tag{4-58}$$

Equation (4-56) requires that along an integral curve of the characteristic vector field, f_0 is constant; i.e.,

$$0 = \frac{df_0}{ds}(s, w(s)) = \frac{\partial f_0}{\partial s} + \frac{\partial f_0}{\partial w}\frac{dw}{ds} = \frac{\partial f_0}{\partial s} + c_\nu f_0^\nu \frac{\partial f_0}{\partial w},$$

by (4-57), which is equivalent to (4-54). Using (4-58) to set $f_0(0, w) = w$ pins down our solution uniquely. $\qquad\square$

Remark 4.6. We note that the numerical coefficients appearing in these Burgers equations depend only on the total number of Dyck paths in $\mathscr{P}^{2\nu}(1, -1)$.

From (4-55) and the self-similar form of f_0,

$$f_0(s, w) = w z_0(s w^{\nu-1}), \tag{4-59}$$

one finds [Ercolani et al. 2008] that

$$z_0(s) = \sum_{j \geq 0} c_\nu^j \zeta_j s^j, \tag{4-60}$$

where

$$c_\nu = 2\nu \binom{2\nu-1}{\nu-1} = (\nu+1)\binom{2\nu}{\nu+1}, \tag{4-61}$$

$$\zeta_j = \frac{1}{j}\binom{\nu j}{j-1} = \frac{1}{(\nu-1)j+1}\binom{\nu j}{j}. \tag{4-62}$$

When $\nu = 2$, ζ_j is the j-th Catalan number. For general ν these are the *higher Catalan numbers* which play a role in a wide variety of enumerative combinatorial problems [Pierce 2007].

Continuum limits of the difference string equations. The *continuum difference string* hierarchies may be derived from the difference string equations (3-35) in a manner completely analogous to what was done with the Toda equations in the previous subsection.

Expanding (3-35) on the asymptotic series (4-48) one arrives at the following asymptotic equations:

$$\frac{1}{n} = \sum_{j=1}^{\infty} \frac{f_{w^{(j)}}}{j!}\left(\frac{1}{n}\right)^j$$

$$+ 2\nu s \sum_{P \in \mathscr{P}^{2\nu-1}(0,-1)} \left(\prod_{m=1}^{\nu}\sum_{j=0}^{\infty}\frac{f_{w^{(j)}}}{j!}\left(\frac{\ell_m(P)+1}{n}\right)^j - \prod_{m=1}^{\nu}\sum_{j=0}^{\infty}\frac{f_{w^{(j)}}}{j!}\left(\frac{\ell_m(P)}{n}\right)^j\right).$$

The equations at leading order, $\mathbb{O}(n^{-1})$, are

$$1 = \partial_w f_0 + 2\nu s \sum_{P \in \mathscr{P}^{2\nu-1}(1,0)} \nu f_0^{\nu-1} \partial_w f_0 = \partial_w f_0 + 2\nu \binom{2\nu-1}{\nu} s\nu f_0^{\nu-1} \partial_w f_0,$$

or, equivalently,

$$\partial_w (w - f_0 - c_\nu s f_0^\nu) = 0, \tag{4-63}$$

which one directly recognizes as the spatial derivative of the hodograph solution (4-55). Evaluating that solution at $w = 1$ yields

$$c_\nu s z_0^\nu + z_0 - 1 = 0, \tag{4-64}$$

which is the functional equation for the generating function of the ν-th higher Catalan numbers, mentioned in the previous subsection.

The terms of the equations at $\mathbb{O}(n^{-2g-1})$ can be computed directly and are found to have the form

$$\partial_w[f_g + c_\nu s\nu f_0^{\nu-1} f_g] + 2s\left(c_\nu \partial_w \sum_{\substack{0 \le k_j < g \\ k_1 + \cdots + k_\nu = g}} f_{k_1} \cdots f_{k_\nu}\right) + \partial_w \sum_{k=0}^{g-1} \frac{f_{kw^{(2g-2k)}}}{(2g-2k+1)!}$$

$$+ 2\nu s\left(F_1^{(\nu)}[2g-2] + F_2^{(\nu)}[2g-4] + \cdots + F_g^{(\nu)}[0]\right) = 0, \tag{4-65}$$

where $F_g^{(\nu)}[2m]$ denotes the coefficient of n^{-2m} in

$$F_g^{(\nu)} = \sum_{\substack{\lambda \in \Lambda^{2g+1} \\ \ell(\lambda) \le \nu}} \left(\sum_{\vec{\ell} \in \mathscr{P}^{2\nu-1}(0,-1)} \left(m_\lambda(\ell_1 + 1, \ldots, \ell_\nu + 1) - m_\lambda(\ell_1, \ldots, \ell_\nu)\right)\right) \frac{f_{w^\lambda}}{\lambda!}$$

$$= \partial_w \sum_{\substack{\lambda \in \Lambda^{2g} \\ \ell(\lambda) \le \nu}} \sum_{\vec{\ell} \in \mathscr{P}^{2\nu-1}(0,-1)} \left(\sum_{\mu \prec \lambda} \frac{1}{|\lambda - \mu| + 1} \binom{\lambda}{\mu} m_\mu(\ell_1, \ldots, \ell_\nu)\right) \frac{1}{r(\lambda)!} \frac{f_{w^\lambda}}{\lambda!}$$

$$= \partial_w \hat{F}_g^{(\nu)}.$$

The above relations are derived in exactly the same manner as those which lead to Theorem 4.3 and Proposition 4.4.

As a consequence of this result one sees that the continuum difference string equation is directly integrable:

$$f_g = \frac{-2s}{1 + c_\nu s\nu f_0^{\nu-1}} \left\{ \left(c_\nu \sum_{\substack{0 \le k_j < g \\ k_1 + \cdots + k_\nu = g}} f_{k_1} \cdots f_{k_\nu}\right) \right.$$

$$\left. + \nu\left(\hat{F}_1^{(\nu-1)}[2g-2] + \hat{F}_2^{(\nu-1)}[2g-4] + \cdots + \hat{F}_g^{(\nu-1)}[0]\right) + \frac{1}{2s} \sum_{k=0}^{g-1} \frac{f_{kw^{(2g-2k)}}}{(2g-2k+1)!} \right\}.$$

Setting $w = 1$ and applying (4-64) to eliminate s this reduces to

$$z_g = \frac{2z_0(z_0 - 1)}{(\nu - (\nu - 1)z_0)}\left\{\left(\sum_{\substack{0 \le k_j < g \\ k_1 + \cdots + k_\nu = g}} \frac{z_{k_1}}{z_0} \cdots \frac{z_{k_\nu}}{z_0}\right) + \frac{1}{2(1 - z_0)}\sum_{k=0}^{g-1} \frac{f_{kw^{(2g-2k)}}|_{w=1}}{(2g - 2k + 1)!}\right.$$

$$\left. + \frac{\nu}{c_\nu z_0^\nu}\left(\widehat{F}_1^{(\nu-1)}[2g - 2] + \widehat{F}_2^{(\nu-1)}[2g - 4] + \cdots + \widehat{F}_g^{(\nu-1)}[0]\right)\Big|_{w=1}\right\}.$$

It is immediate from this representation that z_g is a rational function of z_0. A priori this *antiderivative* should also include a constant term (in w; it could depend on s). This would lead to a term of the form $c(s)/(\nu - (\nu - 1)z_0)$. However, in [Ercolani 2011] it is shown, by an independent argument, that the pole order in z_0 at $\nu/(\nu - 1)$ is always greater than one. Hence the constant of integration must be zero. With further effort this can be refined to:

Theorem 4.7 [Ercolani 2011].

$$z_g(z_0) = \frac{z_0(z_0 - 1)P_{3g-2}(z_0)}{(\nu - (\nu - 1)z_0)^{5g-1}}$$

$$= z_0\left\{\frac{a_0^{(g)}(\nu)}{(\nu - (\nu - 1)z_0)^{2g}} + \frac{a_1^{(g)}(\nu)}{(\nu - (\nu - 1)z_0)^{2g+1}} + \cdots + \frac{a_{3g-1}^{(g)}(\nu)}{(\nu - (\nu - 1)z_0)^{5g-1}}\right\},$$

where P_{3g-2} is a polynomial of degree $3g - 2$ in z_0 whose coefficients are rational functions of ν over the rational numbers $\mathbb{Q}$ and $a_{3g-1}^{(g)}(\nu) \ne 0$.

Remark 4.8. A key element in the proof of Proposition 4.4 is the observation that differentiation with respect to w adjusts the multinomial labeling of partial derivatives in the expansion according to the edges of the Hasse–Young graph (Figure 2). This graph describes the adjacency relations between Young diagrams of differing sizes. The edges describe which partitions of size $2g + 1$ are *covered* by a given partition of size $2g$. Conversely it describes which partitions of size $2g$ cover a partition of size $2g + 1$ which in the setting described here acts as an
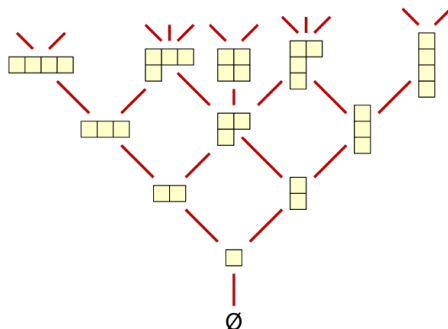
Figure 2. λ Hasse–Young graph (courtesy of D. Eppstein).

antidifferentiation operator. This kind of structure was called a differential poset by Stanley and systematically examined in [Stanley 1988].

5. Determining e_g

The spatial extension of the basic identity (3-25) reads

$$b_{n+k}^2 = \frac{\tau_{n+k+1}^2 \tau_{n+k-1}^2}{\tau_{n+k}^4} b_{n+k}^2(0); \tag{5-66}$$

we have, by taking logarithms,

$$\log \tau_{n+k+1}^2 - 2\log \tau_{n+k}^2 + \log \tau_{n+k-1}^2 = \log(b_{n+k}^2) - \log(b_{n+k}^2)(0), \tag{5-67}$$

where the initial value $b_{n+k}^2(0) = w$ is given by the recursion relations of the Hermite polynomials. As in [Ercolani et al. 2008], we can use formula (5-67) to recursively determine e_g in terms of solutions to the continuum equations. We use the asymptotic expansion of b_{n+k}^2 which has the form (4-45):

$$b_{n+k}^2 = x \sum_{g=0}^{\infty} f_g(s) n^{-2g}. \tag{5-68}$$

Note that the left-hand side of (5-67) has the form of a centered second difference: $\Delta_1 \tau_{n+k}^2 - \Delta_{-1} \tau_{n+k}^2$. It follows that this expression has an expansion for large n involving only even derivatives of the spatial variable w. We have, at order n^{-2g},

$$\frac{\partial^2}{\partial w^2} E_g(s, w) = -\sum_{\ell=1}^{g} \frac{2}{(2\ell+2)!} \frac{\partial^{2\ell+2}}{\partial w^{2\ell+2}} E_{g-\ell}(s, w)$$

$$+ \text{ the } n^{-2g} \text{ terms of } \log\left(\sum_{m=0}^{\infty} \frac{1}{n^{2m}} f_m(s)\right), \tag{5-69}$$

where $E_h(s, w) = w^{2-2h} e_h(w^{\nu-1} s)$. In [Ercolani 2011] it was shown that e_g is rational in z_0 with poles located only at $z_0 = \nu/(\nu-1)$. However we will now prove the more refined result stated in Theorem 2.3.

The proof of this theorem is by induction on g. (The base case of $g = 2$ is established by direct calculation [Ercolani et al. 2008].) We assume that (2-11) holds for all $k < g$. We state here, without proof, some straightforward lemmas and propositions describing the derivatives of (2-11) (details may be found in [Ercolani 2011] where similar lemmas are proved for the z_g).

Lemma 5.1. $(E_k)_{w^{(p)}}(s, w) = \sum_{j=0}^{p} (\nu - 1)^j Q_j^{(p,k)}(\nu) w^{2-2k+(\nu-1)j-p} s^j e_k^{(j)},$

where $e_k^{(j)} = d^j e_k / d\tilde{s}^j$, $\tilde{s} = sw^{\nu-1}$,

$$Q_j^{(p,k)}(\nu) = Q_{j-1}^{(p-1,k)}(\nu) + \{(\nu-1)j - (2k-3+p)\}Q_j^{(p-1,k)}(\nu), \quad 0 < j < p,$$

$$Q_0^{(p,k)}(\nu) = (2-2k)_p, \quad p > 0,$$

$$Q_p^{(p,k)}(\nu) = 1,$$

$$Q_j^{(p,k)}(\nu) = 0, \qquad j > p \text{ or } j < 0.$$

Lemma 5.2. *For $1 < k < g$ and $j > 0$,*

$$e_k^{(j)} = (-1)^j c_\nu^j z_0^{j\nu+1} \left(\sum_{\ell=0}^{3k-4+j} \frac{c_\ell^{(k,j)}(\nu)}{(\nu-(\nu-1)z_0)^{2k+\ell+j-1}} \right),$$

$$c_\ell^{(k,j)}(\nu) = [(j-1)\nu - (2k+\ell+(j-3))]c_\ell^{k,j-1}(\nu)$$
$$+ \nu(2k+\ell+(j-3))c_{\ell-1}^{k,j-1}(\nu),$$

$$c_\ell^{(k,j)}(\nu) = 0 \quad \ell < 0, \ell \geq 3k-3+j,$$

$$c_\ell^{(k,0)}(\nu) = c_\ell^{(k)}(\nu).$$

Lemma 5.3. *For $1 < k < g$ and $j > 0$,*

$$s^j e_k^{(j)} = \frac{z_0}{(\nu-1)^j} \sum_{r=0}^{j} \sum_{m=r}^{3k-4+j+r} \frac{(-1)^{j-r}\binom{j}{r}c_{m-r}^{(k,j)}}{(\nu-(\nu-1)z_0)^{2k+m-1}},$$

where $m = \ell + r$.

Proposition 5.4. *For $1 < k < g$ and $p > 0$,*

$$(E_k)_{w^{(p)}}(s,1) = (2-2k)_p C^{(k)}$$

$$+z_0 \sum_{m=0}^{3k+2p-4} \frac{\sum_{j=0}^{p} Q_j^{(p,k)}(\nu) \sum_{r=0}^{m}(-1)^{j-r}\binom{j}{r}c_{m-r}^{(k,j)}(\nu)}{(\nu-(\nu-1)z_0)^{2k+m-1}}.$$

Lemma 5.5. $\displaystyle\sum_{j=0}^{p} Q_j^{(p,k)}(\nu) \sum_{r=0}^{m}(-1)^{j-r}\binom{j}{r}c_{m-r}^{(k,j)}(\nu) = 0,$

for $m = 0, 1, \ldots, p-1$.

By this vanishing lemma, the minimal pole order of the expansion in Proposition 5.4 is $\geq 2k - 2 + p$. In particular the minimal pole orders coming from terms involving E_k on the right-hand side of (5-69) are all greater than $2k - 2 + (2g + 2 - 2k) = 2g$.

Proposition 5.6. *The n^{-2g} term of* $\log \sum\limits_{m=0}^{\infty} \dfrac{1}{n^{2m}} f_m(s)\Big|_{w=1}$ *is given by*

$$\sum_{|\lambda|=g} \frac{(-1)^{\ell(\lambda)-1}}{\prod_{j\geq 1} r_j(\lambda)!} \prod_{j\geq 1} \left(\frac{z_j}{z_0}\right)^{r_j(\lambda)},$$

which, by Theorem 4.7, equals

$$\sum_{|\lambda|=g} \frac{(-1)^{\ell(\lambda)-1}}{\prod_{j\geq 1} r_j(\lambda)!} \prod_{j\geq 1} \left\{ \frac{a_0^{(j)}(v)}{(v-(v-1)z_0)^{2j}} + \cdots + \frac{a_{3j-1}^{(j)}(v)}{(v-(v-1)z_0)^{5j-1}} \right\}^{r_j(\lambda)}.$$

This result shows that the minimal pole order coming from the log terms in (5-69) is once again greater than $\sum_{j\geq 1} 2jr_j = 2g$.

The preceding lemmas and propositions provide explicit Laurent expansions (in z_0) for all terms on the right-hand side of (5-69), with two exceptions:

$$\frac{\partial^{2g}}{\partial w^{2g}} E_1(s,w) = \frac{\partial^{2g}}{\partial w^{2g}} e_1(w^{v-1}s), \quad e_1 = -\frac{1}{12}\log(v-(v-1)z_0), \quad (5\text{-}70)$$

and

$$\frac{\partial^{2g+2}}{\partial w^{2g+2}} E_0(s,w) = \frac{\partial^{2g+2}}{\partial w^{2g+2}} w^2 e_0(w^{v-1}s),$$

$$e_0 = \frac{1}{2}\log z_0 + \frac{(v-1)^2}{4v(v+1)}(z_0-1)\left(z_0 - \frac{3(v+1)}{v-1}\right). \quad (5\text{-}71)$$

With a small modification (5-70) may be brought into line with Proposition 5.4:

Proposition 5.7.

$(E_1)_{w^{(p)}}(s,1)$

$$= z_0 \sum_{m=0}^{2p-1} \frac{\sum_{j=1}^{p} Q_j^{(p,1)}(v) \sum_{r=0}^{m}(-1)^{j-r}\binom{j}{r}c_{m-r}^{(1,j)}(v)}{(v-(v-1)z_0)^{m+1}}$$

$$= -\frac{(p-1)!}{12}$$

$$+ \frac{1}{v-1} \sum_{m=1}^{2p} \frac{\sum_{j=1}^{p} Q_j^{(p,1)}(v) \sum_{r=0}^{m}(-1)^{j-r}\binom{j}{r}\{vc_{m-r-1}^{(1,j)}(v) - c_{m-r}^{(1,j)}(v)\}}{(v-(v-1)z_0)^{m}},$$

with $c_0^{(1,1)} = \frac{1}{12}(v-1)$. *All other coefficients are then specified by the corresponding recursions stated in Lemmas 5.1–5.3 with k set to 1.*

A variant of the vanishing Lemma 5.5 also holds for $(E_1)_{w^{(p)}}(s,1)$:

Lemma 5.8. $\displaystyle\sum_{j=1}^{p} Q_j^{(p,1)}(v) \sum_{r=0}^{m}(-1)^{j-r}\binom{j}{r}\left(c_{m-r}^{(1,j)}(v) - v c_{m-r-1}^{(1,j)}(v)\right) = 0,$

for $m = 0, 1, \ldots, p-1$.

It follows that the minimal pole order of the expansion in Proposition 5.7 is at least p and so the corresponding contribution to the minimal pole order of (5-69) is $\geq 2g$.

Finally we observe that for $p \geq 3$, $\dfrac{\partial^p}{\partial w^p} E_0(s, w)$ is a rational function of f_0 and its w-derivatives:

Proposition 5.9.

$$\frac{\partial^p}{\partial w^p} E_0(s, w)$$

$$= \frac{\partial^p}{\partial w^p}\left[w^2 \tfrac{1}{2}\log f_0 + \frac{(v-1)^2}{4v(v+1)}(f_0 - w)\left(f_0 - \frac{3(v+1)}{v-1}w\right) - w^2 \tfrac{1}{2}\log w\right]$$

$$= \binom{p}{2}\left(\frac{f_{0w}}{f_0}\right)_{w^{(p-3)}} + pw\left(\frac{f_{0w}}{f_0}\right)_{w^{(p-2)}} + \frac{w^2}{2}\left(\frac{f_{0w}}{f_0}\right)_{w^{(p-1)}}$$

$$+ \frac{(v-1)^2}{2v(v+1)}\left[\sum_{j=0}^{\lfloor\frac{p}{2}\rfloor}\binom{p}{j} f_{0w^{(j)}} f_{0w^{(p-j)}} - \frac{2v+1}{v-1}\left(w f_{0w^{(p)}} + p f_{0w^{(p-1)}}\right)\right]$$

$$+ (p-3)!\left(-\frac{1}{w}\right)^{p-2}. \tag{5-72}$$

Each line of the proposition can be established directly by induction, starting with the base case $p = 3$. It then follows from Proposition 3.1(iii) of [Ercolani 2011] that the minimal pole order contributed by (5-71) is $\geq 2g$.

We are now in a position to outline the

Proof of Theorem 2.3. In [Ercolani 2011, Theorem 1.3] it was shown that

$$e_g(z_0) = \frac{(z_0 - 1) q_{d(g)}(z_0)}{(v - (v-1)z_0)^{o(g)}}, \tag{5-73}$$

where $q_{d(g)}(z_0)$ denotes a polynomial of degree $d(g)$ in z_0. We first want to determine the relation between this degree and the pole order $o(g)$. To this end we observe from Propositions 5.4, 5.6, 5.7, and 5.9 that the right-hand side of (5-69), evaluated at $w = 1$, is a rational function in z_0 that approaches a finite constant value as $z_0 \to \infty$. From the form of the left-hand side of (5-69) evaluated at $w = 1$ one also sees that its asymptotic order (as $z_0 \to \infty$) is the same as that of e_g. Hence, $d(g) = o(g) - 1$ and this shows that (2-11) is valid up to the determination of the minimal and maximal pole orders at $z_0 = v/(v-1)$, to which we now turn.

In the preceding lemmas and propositions we have seen that, for all terms on the right-hand side of (5-69), the minimal pole order is $\geq 2g$. Furthermore, from these same representations together with Proposition 3.1(iii) of [Ercolani 2011] one sees that, with the possible exception of the genus 0 terms in (5-72), the maximal pole order of the terms on the right-hand side of (5-69) is $5g - 1$. The apparent maximal pole order in (5-72) is $4g + 3$, which exceeds the stated bound when $g = 2, 3$. This maximal order comes from terms containing the factor $f_{0w^{(2g+2)}}$, which are, specifically,

$$\left[\frac{1}{2} \frac{f_{0w^{(2g+2)}}}{f_0} - \frac{(2\nu + 1)(\nu - 1)}{2\nu(\nu + 1)} f_{0w^{(2g+2)}} + \frac{(\nu - 1)^2}{2\nu(\nu + 1)} (f_0 f_{0w^{(2g+2)}}) \right]_{w=1}$$

$$= \frac{f_{0w^{(2g+2)}}}{2f_0} \bigg|_{w=1} \left[1 - \frac{(2\nu + 1)(\nu - 1)}{\nu(\nu + 1)} z_0 + \frac{(\nu - 1)^2}{\nu(\nu + 1)} z_0^2 \right]$$

$$= \frac{f_{0w^{(2g+2)}}}{2\nu(\nu + 1) f_0} \bigg|_{w=1} \{ (\nu - (\nu - 1)z_0) + (\nu - (\nu - 1)z_0)^2 \}.$$

Hence the maximal pole order contributed by the genus 0 terms is $\leq 4g + 2$. Indeed, one may go further with this type of analysis to show that the coefficient of order $\mathbb{O}\big((\nu - (\nu - 1)z_0)^{-4g-2}\big)$ also vanishes. In establishing this the following identity, which is a direct consequence of the quadratic relation satisfied by the generating function for the Catalan numbers $\frac{1}{g+1}\binom{2g}{g}$, proves useful:

$$\frac{(2g)!}{g!} = \sum_{j=1}^{g} \binom{g+1}{j} \frac{(2j - 2)!(2(g - j))!}{(j - 1)!(g - j)!}.$$

It follows that the maximal pole order coming from genus 0 terms is $\leq 4g + 1$ which is $< 5g - 1$ for $g > 2$. This establishes that the pole orders on the right-hand side of (5-69) are always bounded between $2g$ and $5g - 1$. For $g > 2$, the case-by-case checking of terms on the right-hand side of (5-69) that has been carried out in this subsection shows that the maximal pole order is realized by the term in Proposition 5.6 corresponding to the partition λ of $2g$ having minimal length ($= 1$), i.e., the partition whose Young diagram is a single row. This implies that the residue of the maximal order pole is $a_{3g-1}^{(g)}(\nu)$, which is nonzero by Theorem 4.7. (This also holds for $g = 2$, as can be checked by direct calculation; see, for example, [Ercolani 2011, Section 1.4.2].) Hence the maximal order pole is realized. Now, given that e_g has the form (5-73) with $d(g) = o(g) - 1$, it follows that $\partial^2 E_g(s, 1)/\partial w^2$ raises the minimum pole degree by 2 and the maximum pole degree by 4 with the coefficient at this order given by (2-14).

To establish (2-13) first note that by Euler's relation, $2 - 2g = m - \nu m + F$ for a g-map where m is the number of (2ν-valent) vertices and F is the number of faces. Since $F \geq 1$, one immediately sees that the number of vertices of such

a map must satisfy the inequality $m \geq \frac{2g-1}{\nu-1}$. It follows that

$$e_g^{(j)}(s=0) = 0 \quad \text{for } j \leq r = \max\left\{1, \frac{2g-1}{\nu-1}\right\}$$

(e_g must vanish at least simply at $s = 0$ since $\tau_n^2(s = 0) \equiv 1$). Via Cauchy's theorem these conditions may be reexpressed as

$$0 = \frac{1}{2\pi i} \oint_{s\sim 0} \frac{e_g(s)}{s^{j+1}} \, ds = \frac{-1}{2\pi i} \oint_{z\sim 1} \frac{z^{\nu j-1} q_{5g-6}(z)}{(\nu - (\nu-1)z)^{5g-6}(z-1)^j} \, dz,$$

for $j \leq r$, where in the second line we have rewritten e_g as a rational function of z ((5-73) with $o(g) = 5g - 5$) and employed the change of variables

$$\frac{ds}{dz} = -\frac{\nu - (\nu-1)z}{z^{\nu+1}},$$

which may be deduced from the string equation (4-64). This yields a contour integral in z centered at 1. Now one can see that these vanishing conditions are satisfied if and only if $q_{5g-6}^{(j)}(z = 1) = 0$ for $j \leq r$, which in turn proves (2-13).

Finally we turn to the determination of the constant $C^{(g)}$. By Proposition 5.6, contributions to the constant term of e_g come only from the first sum on the right-hand side of (5-69). The parts of this coming from $g = 0$ and $g = 1$ are, by Propositions 5.9 and 5.7, respectively,

$$-2\frac{(2g-1)!}{(2g+2)!} \quad \text{and} \quad 2\frac{(2g-1)!}{(2g)!}\frac{1}{12}.$$

At higher genus, $k < g$, the contribution to the constant term is determined by Lemma 5.1 to be

$$-2\frac{(2-2k)_{2g-2k+2}}{(2g-2k+2)!}C^{(k)}.$$

Hence, by (5-69) we have

$$(2-2g)(1-2g)C^{(g)}$$

$$= -2\frac{(2g-1)!}{(2g+2)!} + 2\frac{(2g-1)!}{(2g)!}\frac{1}{12} - 2\sum_{k=2}^{g-1} \frac{(2-2k)_{2g-2k+2}}{(2g-2k+2)!}C^{(k)},$$

from which (2-15) immediately follows. □

6. The case of odd valence

In the case when j is odd in the weight (2-3) for V, there is clearly a problem in applying the method of orthogonal polynomials as it was outlined in Section 3. Very recently, however, a generalization of the *equilibrium measure* (which governs the leading order behavior of the free energy associated to (2-1)) was

developed and applied to this problem [Bleher and Deaño 2012]. It is based on generalizing to a class of complex valued non-Hermitian orthogonal polynomials on a contour in the complex plane other than the real axis. These extensions were motivated by new ideas in approximation theory related to complex Gaussian quadrature of integrals with high order stationary points [Deaño et al. 2010].

But even when the issue of existence of appropriate orthogonal polynomials has been resolved, there are still a number of significant obstacles to deriving results like Theorem 2.3 that are not present when the valence j is even. For odd valence there is an additional string of recurrence coefficients, the diagonal coefficients a_n of $\mathcal{L}$, whose asymptotics need to be analyzed. This in turn requires that the lattice paths used to define and analyze the Toda and difference string equations must be generalized to the class of *Motzkin paths* which can have segments where the lattice site remains fixed rather than always taking a step (either up or down) as was the case for Dyck paths.

Nevertheless, all these constructions have been carried out in [Ercolani and Pierce 2012] to derive the hierarchies of continuum Toda and difference string equations when the valence j is odd.

The recurrence coefficients again have asymptotic expansions with continuum representations given by

$$a_{n+k,N} = h(s_1, s_{2\nu+1}, w) = x^{1/2} \sum_{g \geq 0} h_g(s_1, s_{2\nu+1}, w) n^{-g} \qquad (6\text{-}74)$$

and

$$h_g(s_1, s_{2\nu+1}, w)$$
$$= -w^{1-g} \sum_{\substack{2g_1+j=g+1 \\ g_1 \geq 0, j > 0}} \frac{1}{j!}$$
$$\times \frac{\partial^{j+1}}{\partial s_1 \partial \tilde{w}^j} \left[\tilde{w}^{2-2g_1} e_{g_1} \left((w\tilde{w})^{-\frac{1}{2}} s_1, (w\tilde{w})^{\nu-\frac{1}{2}} s_{2\nu+1} \right) \right]_{\tilde{w}=1}. \quad (6\text{-}75)$$

The off-diagonal coefficients have corresponding representations which are much as they were in the even valence case,

$$b_{n+k,N}^2 = f(s_1, s_{2\nu+1}, w) = x \sum_{g \geq 0} f_g(s_1, s_{2\nu+1}, w) n^{-2g}, \quad (6\text{-}76)$$

$$f_g(s_1, s_{2\nu+1}, w) = w^{2-2g} \frac{\partial^2}{\partial s_1^2} e_g(w^{-1/2} s_1, w^{\nu-1/2} s_{2\nu+1}). \qquad (6\text{-}77)$$

The coefficients in these expansions have a self-similar structure given by

$$h_g(s_1, s_{2\nu+1}, w) = w^{\frac{1}{2}-g} u_g(s_1 w^{-1/2}, s_{2\nu+1} w^{\nu-\frac{1}{2}}), \qquad (6\text{-}78)$$

$$f_g(s_1, s_{2\nu+1}, w) = w^{1-2g} z_g(s_1 w^{-1/2}, s_{2\nu+1} w^{\nu-\frac{1}{2}}). \qquad (6\text{-}79)$$

At leading order the continuum Toda equations are

$$\frac{\partial}{\partial s}\binom{h_0}{f_0} + (2\nu+1)\begin{pmatrix} B_{11} & B_{12} \\ f_0 B_{12} & B_{11} \end{pmatrix}\frac{\partial}{\partial w}\binom{h_0}{f_0} = 0, \qquad (6\text{-}80)$$

and the leading order continuum difference string equations are

$$\binom{0}{1} = \begin{pmatrix} A_{11} & A_{12} \\ f_0 A_{12} & A_{11} \end{pmatrix}\frac{\partial}{\partial w}\binom{h_0}{f_0}, \qquad (6\text{-}81)$$

where the coefficients of the matrix in (6-80) are specified by

$$B_{11} = \sum_{\mu=1}^{\nu}\binom{2\nu}{2\mu-1,\,\nu-\mu,\,\nu-\mu+1}h_0^{2\mu-1}f_0^{\nu-\mu+1}, \qquad (6\text{-}82)$$

$$B_{12} = \sum_{\mu=0}^{\nu}\binom{2\nu}{2\mu,\,\nu-\mu,\,\nu-\mu}h_0^{2\mu}f_0^{\nu-\mu}, \qquad (6\text{-}83)$$

and those of the matrix in (6-81) by

$$A_{11} = 1 + (2\nu+1)s\sum_{\mu=0}^{\nu-1}\binom{2\nu}{2\mu+1,\,\nu-\mu-1,\,\nu-\mu}(\nu-\mu)h_0^{2\mu+1}f_0^{\nu-\mu-1}, \quad (6\text{-}84)$$

$$A_{12} = (2\nu+1)s\sum_{\mu=0}^{\nu-1}\binom{2\nu}{2\mu,\,\nu-\mu-1,\,\nu-\mu+1}(\nu-\mu+1)h_0^{2\mu}f_0^{\nu-\mu-1}. \quad (6\text{-}85)$$

Remark 6.1. The index μ appearing in the trinomial coefficients corresponds to the number of flat steps, $2\mu-1$ or 2μ, in the Motzkin paths giving rise to that term.

It is straightforward to see that (6-80) may be rewritten in conservation law form as

$$\frac{\partial}{\partial s}\binom{h_0}{f_0+\frac{1}{2}h_0^2} + \frac{\partial}{\partial w}\binom{\Psi_1}{\Psi_2+h_0\Psi_1} = 0, \qquad (6\text{-}86)$$

where the coefficients in the flux vector are given by

$$\Psi_1 = \sum_{\mu=0}^{\nu}\binom{2\nu+1}{2\mu,\,\nu-\mu,\,\nu-\mu+1}h_0^{2\mu}f_0^{\nu-\mu+1}, \qquad (6\text{-}87)$$

$$\Psi_2 = \sum_{\mu=0}^{\nu}\binom{2\nu+1}{2\mu+1,\,\nu-\mu-1,\,\nu-\mu+1}h_0^{2\mu+1}f_0^{\nu-\mu+1}. \qquad (6\text{-}88)$$

Recently, we have determined that the equations (6-81) are in fact a differentiated form of the generalized hodograph solution of the conservation law (6-86).

This hodograph solution is given by

$$\Phi_1 \doteq h_0 + (2\nu + 1)sB_{12} = 0, \qquad (6\text{-}89)$$
$$\Phi_2 \doteq f_0 + (2\nu + 1)sB_{11} = w. \qquad (6\text{-}90)$$

Analogous to what was done in Theorem 2.3 we expect to determine closed form expressions for all the coefficients in the topological expansion with odd weights. The first few of these, for the trivalent case, are [Ercolani and Pierce 2012]

$$e_0(t_3) = \frac{1}{2}\log(z_0) + \frac{1}{12}\frac{(z_0 - 1)(z_0^2 - 6z_0 - 3)}{(z_0 + 1)},$$

$$e_1(t_3) = -\frac{1}{24}\log\left(\frac{3}{2} - \frac{z_0^2}{2}\right),$$

$$e_2(t_3) = \frac{1}{960}\frac{(z_0^2 - 1)^3(4z_0^4 - 93z_0^2 - 261)}{(z_0^2 - 3)^5} \qquad (6\text{-}91)$$

$$= \frac{1}{240} - \frac{3}{64Z} - \frac{29}{32Z^2} - \frac{191}{48Z^3} - \frac{55}{8Z^4} - \frac{21}{5Z^5}, \quad Z = z_0 - 3,$$

where z_0 is implicitly related to t_3 by the polynomial equation

$$1 = z_0^2 - 72t_3^2 z_0^3.$$

In fact, $z_0(t_3)$ is the generating function for a *fractional* generalization of the Catalan numbers. Its m-th coefficient counts the number of connected, noncrossing symmetric graphs on $2m + 1$ equidistributed vertices on the unit circle (V. U. Pierce, private communication).

7. Concluding remarks

7.1. Spectrum. Nothing has been said, in this article, about the eigenvalues of the random matrix M although this is at the heart of the Riemann–Hilbert analysis underlying all of our results. The essential link comes through the *equilibrium measure* [Ercolani and McLaughlin 2003; Ercolani et al. 2008], or density of states, for these eigenvalues. When $t_j = 0$ in (2-2), this equilibrium measure reduces to the well known Wigner semicircle law. As t_j changes this measure deforms; but, for t_j satisfying the bounds implicit in (2-5) (i), its support remains a single interval, $[\alpha, \beta]$. The *edges of the spectrum* $\alpha(t_j), \beta(t_j)$ evolve dynamically with t_j. In fact can show that α and β are the Riemann invariants of the hyperbolic system (6-80). In the case of even valence these invariants collapse to $\pm 2\sqrt{z_0}$ where z_0 is the generating function upon which all the e_g

are built, as described in Theorem 2.3. So the edge of the spectrum is indeed directly related to the genus expansion (2-5).

7.2. Random surfaces. It follows from Corollary 3.3 and the map-theoretic interpretation of e_g given by (2-8) that $z_g(t_j)$ is a generating function for enumerating j-regular g-maps with, in addition, two *legs*. A *leg* is a univalent vertex; that is, a vertex with just one adjacent edge, connected to some other vertex of the map. In particular $z_0(t)$ enumerates such maps on the Riemann sphere or what are more commonly referred to as two-legged *planar maps*. In a remarkable paper, [Schaeffer 1997], building on prior work of Cori and Vauquelin [1981], Schaeffer found a constructive correspondence between two-legged 2ν-regular planar maps and 2ν-valent *blossom trees*. A 2ν-valent blossom tree is a rooted 2ν-valent tree, with each external vertex taken to be a leaf that is colored either black or white and such that each internal (nonroot) vertex is adjacent to exactly $\nu - 1$ black leaves. This gives another interpretation of $z_0(t_{2\nu})$ as the generating function for the enumeration of blossom trees. It seems reasonable to hope that the arithmetic data implicit in the coefficients $a_m^{(g)}(\nu)$ in Theorem 4.7 (resp. $c_m^{(g)}(\nu)$ in Theorem 2.3) might provide a means, such as *sewing rules*, for constructing two-legged 2ν-regular g-maps (resp. pure 2ν-regular g-maps) from blossom trees.

In another direction Bouttier, Di Francesco and Guitter [Di Francesco 2006] have studied the combinatorics of geodesic distance for planar maps. They define the geodesic distance of a two-legged graph to be the minimum number of edges crossed by a continuous path between the two legs and study $r_d(t_{2\nu})$, the generating function for enumerating all two-legged 2ν-valent planar maps whose geodesic distance is $\leq d$. They find surprising and elegant closed form expressions for the $r_d(t_{2\nu})$. The statistics of planar maps is a natural stepping off point for the study of random surfaces. There has been a lot of recent activity in this direction by Le Gall and his collaborators related to the work of Schaeffer and Bouttier et al. See, for example, [Le Gall 2010].

7.3. Enumerative geometry of moduli spaces. A different (from matrix models) representation of 2D quantum gravity may be given in terms of intersection theory on the moduli space of stable curves (Riemann surfaces), $\overline{\mathcal{M}}_{g,n}$ and from this alternate perspective Witten conjectured that a generating function for the intersection numbers of tautological bundles on $\overline{\mathcal{M}}_{g,n}$ should be "given by" a double-scaling limit of the differentiated free energy (3-28) for matrix models. (A precise description of this double-scaling limit may be found in [Ercolani 2011].) He further conjectured that this intersection theoretic generating function should, with respect to appropriate choices of parameter variables, satisfy the Korteweg–de Vries (KdV) equation. Subsequently, Kontsevich [1992] was able

to outline a proof of Witten's KdV conjecture based on a combinatorial model of intersection theory on $\overline{\mathcal{M}}_{g,n}$. This model expresses tautological intersections in terms of sums over trivalent graphs on a genus g Riemann surface. He was then able to recast this sum in terms of a special matrix integral involving cubic weights on which the proof of Witten's KdV conjecture is based. For a readable overview on the above circle of ideas we refer to [Okounkov and Pandharipande 2009].

However, the first Witten conjecture, on relating the intersection-theoretic free energy to the matrix model free energy (see (2-5)), remains open. With the results described in this paper it may now be possible to determine if, and in precisely what sense, this conjecture might be true and to see if this leads to connections between the KdV equation in the Witten–Kontsevich model and the conservation laws given by (6-80). In addition, given the recent results on matrix models with odd dominant weights [Bleher and Deaño 2012; Ercolani and Pierce 2012], it may now be possible to give a rigorous treatment of Kontsevich's matrix integral which, up to now, has been formal.

More recently there have been other, perhaps more natural, approaches to the proof of Witten's KdV conjecture [Kazarian and Lando 2007; Mirzakhani 2007; Goulden et al. 2009], in terms of coverings of the Riemann sphere and *Hurwitz numbers* for which the generating functions specified in Theorem 2.3 should also have a natural interpretation.

7.4. Analytical deformations and critical parameters. In [Ercolani 2011] it was observed that the equilibrium measure (see concluding remark Section 7.1) for the weight V, with $j = 2\nu$ in (2-2), may be reexpressed as

$$\mu_{V_t}(\lambda) = z_0\mu_0(\lambda) + (1 - z_0)\mu_\infty(\lambda), \qquad (7\text{-}92)$$

where μ_0 is the equilibrium measure for $V = 1/2\lambda^2$ (the *semicircle* law) and μ_∞ is the equilibrium measure for $V = \lambda^{2\nu}$; that is, the general measure is a linear combination, over z_0, of two extremal monomial equilibrium measures. For $z_0 \in [0, 1]$ (which corresponds to $t_{2\nu} \in [0, \infty]$), this combination is convex and (7-92) is indeed a measure with a single interval of support in $\lambda \in \mathbb{R}$. This may be analytically continued to a complex z_0 neighborhood of $[0, 1]$ so that (7-92) remains a positive measure along an appropriate connected contour ("single interval") in the complex λ-plane. For $\nu = 2$ this continuation may be made up to a boundary curve in the complex z_0-plane passing through $z_0 = 2$ (with a corresponding image in the complex t_4 plane). (For a related result see [Bertola and Tovbis 2011].) This should be extendable for general ν. The mechanism for carrying out this continuation is to regard (7-92) as a z_0-parametrized family of holomorphic quadratic differentials. The candidate for the measure's support is

then an appropriate bounded real trajectory of the quadratic differential. Outside the boundary curve, the Riemann–Hilbert analysis used in this paper may be analytically deformed and our results extended. The boundary may be regarded as a curve of critical parameters for this deformation. This curve is precisely the locus where the Riemann invariants, that determine the edge of the spectrum (as described in concluding remark 7.1) exhibit a shock.

This scenario is reminiscent of that for the small $\hbar$-limit of the nonlinear Schrödinger equation [Jin et al. 1999; Kamvissis et al. 2003] in which the analogue of our boundary curve is the envelope of *dispersive shocks*. In that setting it is the Zakharov–Shabat inverse scattering problem that shows one how to pass through the dispersive shocks and describe a continuation of measure-valued solutions with so-called *multigap* support. It is our expectation that coupling gravity to an appropriate conformal field theory (to thus arrive at a bona fide string theory) [Mariño 2005] will play a similar role in our setting to determine a unique continuation through the boundary curve of critical parameters to a unique equilibrium measure with multicut support. We also hope that this will help bring powerful methods from the study of dispersive limits of nonlinear PDE into the realm of random matrix theory.

Acknowledgement

References

[Ambjørn et al. 1993] J. Ambjørn, L. Chekhov, C. F. Kristjansen, and Y. Makeenko, "Matrix model calculations beyond the spherical limit", *Nuclear Phys. B* **404**:1-2 (1993), 127–172.

[Bauer and Itzykson 1996] M. Bauer and C. Itzykson, "Triangulations", *Discrete Math.* **156**:1-3 (1996), 29–81.

[Bertola and Tovbis 2011] M. Bertola and A. Tovbis, "Asymptotics of orthogonal polynomials with complex varying quartic weight: global structure, critical point behaviour and the first Painlevé equation", 2011. arXiv 1108.0321

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

[Bleher and Deaño 2012] P. Bleher and A. Deaño, "Topological expansion in the cubic random matrix model", *Int. Math. Res. Not.* (2012), 2699–2755.

[Bleher and Its 2005] P. M. Bleher and A. R. Its, "Asymptotics of the partition function of a random matrix model", *Ann. Inst. Fourier* (*Grenoble*) **55**:6 (2005), 1943–2000.

[Chekhov and Eynard 2006] L. Chekhov and B. Eynard, "Hermitian matrix model free energy: Feynman graph technique for all genera", *J. High Energy Phys.* **2006**:3 (2006), #014.

[Cori and Vauquelin 1981] R. Cori and B. Vauquelin, "Planar maps are well labeled trees", *Canad. J. Math.* **33**:5 (1981), 1023–1042.

[Deaño et al. 2010] A. Deaño, D. Huybrechs, and A. B. J. Kuijlaars, "Asymptotic zero distribution of complex orthogonal polynomials associated with Gaussian quadrature", *J. Approx. Theory* **162**:12 (2010), 2202–2224.

[Deift 1999] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann–Hilbert approach*, Courant Lecture Notes in Mathematics **3**, New York University, Courant Institute of Mathematical Sciences, AMS, New York / Providence, RI, 1999.

[Di Francesco 2006] P. Di Francesco, "2D quantum gravity, matrix models and graph combinatorics", pp. 33–88 in *Applications of random matrices in physics*, edited by E. Brézin et al., NATO Sci. Ser. II Math. Phys. Chem. **221**, Springer, Dordrecht, 2006.

[Di Francesco et al. 1995] P. Di Francesco, P. Ginsparg, and J. Zinn-Justin, "2D gravity and random matrices", *Phys. Rep.* **254**:1-2 (1995), 133.

[Douglas and Shenker 1990] M. R. Douglas and S. H. Shenker, "Strings in less than one dimension", *Nuclear Phys. B* **335**:3 (1990), 635–654.

[Ercolani 2011] N. M. Ercolani, "Caustics, counting maps and semi-classical asymptotics", *Nonlinearity* **24**:2 (2011), 481–526.

[Ercolani and McLaughlin 2001] N. M. Ercolani and K. T.-R. McLaughlin, "Asymptotics and integrable structures for biorthogonal polynomials associated to a random two-matrix model: advances in nonlinear mathematics and science", *Phys. D* **152/153** (2001), 232–268.

[Ercolani and McLaughlin 2003] N. M. Ercolani and K. D. T.-R. McLaughlin, "Asymptotics of the partition function for random matrices via Riemann–Hilbert techniques and applications to graphical enumeration", *Int. Math. Res. Not.* **2003**:14 (2003), 755–820.

[Ercolani and Pierce 2012] N. M. Ercolani and V. U. Pierce, "The continuum limit of Toda lattices for random matrices with odd weights", *Commun. Math. Sci.* **10**:1 (2012), 267–305.

[Ercolani et al. 2008] N. M. Ercolani, K. D. T.-R. McLaughlin, and V. U. Pierce, "Random matrices, graphical enumeration and the continuum limit of Toda lattices", *Comm. Math. Phys.* **278**:1 (2008), 31–81.

[Eynard 2004] B. Eynard, "Topological expansion for the 1-Hermitian matrix model correlation functions", *J. High Energy Phys.* 11 (2004), 031.

[Eynard and Orantin 2009] B. Eynard and N. Orantin, "Topological recursion in enumerative geometry and random matrices", *J. Phys. A* **42**:29 (2009), 293001, 117.

[Fokas et al. 1992] A. S. Fokas, A. R. Its, and A. V. Kitaev, "The isomonodromy approach to matrix models in 2D quantum gravity", *Comm. Math. Phys.* **147**:2 (1992), 395–430.

[Goulden et al. 2009] I. P. Goulden, D. M. Jackson, and R. Vakil, "A short proof of the λ_g-conjecture without Gromov-Witten theory: Hurwitz theory and the moduli of curves", *J. Reine Angew. Math.* **637** (2009), 175–191.

[Gross and Migdal 1990] D. J. Gross and A. A. Migdal, "A nonperturbative treatment of two-dimensional quantum gravity", *Nuclear Phys. B* **340**:2-3 (1990), 333–365.

[Janson et al. 1993] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel, "The birth of the giant component", *Random Structures Algorithms* **4**:3 (1993), 231–358.

[Jin et al. 1999] S. Jin, C. D. Levermore, and D. W. McLaughlin, "The semiclassical limit of the defocusing NLS hierarchy", *Comm. Pure Appl. Math.* **52**:5 (1999), 613–654.

[Kamvissis et al. 2003] S. Kamvissis, K. D. T.-R. McLaughlin, and P. D. Miller, *Semiclassical soliton ensembles for the focusing nonlinear Schrödinger equation*, Annals of Mathematics Studies **154**, Princeton Univ. Press, Princeton, NJ, 2003.

[Kazarian and Lando 2007] M. E. Kazarian and S. K. Lando, "An algebro-geometric proof of Witten's conjecture", *J. Amer. Math. Soc.* **20**:4 (2007), 1079–1089.

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

[Le Gall 2010] J.-F. Le Gall, "Geodesics in large planar maps and in the Brownian map", *Acta Math.* **205**:2 (2010), 287–360.

[Macdonald 1995] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Oxford Univ. Press, New York, 1995.

[Mariño 2005] M. Mariño, "Les Houches lectures on matrix models and topological strings", 2005. arXiv hep-th/0410165

[Mirzakhani 2007] M. Mirzakhani, "Weil–Petersson volumes and intersection theory on the moduli space of curves", *J. Amer. Math. Soc.* **20**:1 (2007), 1–23.

[Okounkov and Pandharipande 2009] A. Okounkov and R. Pandharipande, "Gromov–Witten theory, Hurwitz numbers, and matrix models", pp. 325–414 in *Algebraic geometry, I* (Seattle, 2005), edited by D. Abramovich et al., Proc. Sympos. Pure Math. **80**, Amer. Math. Soc., Providence, RI, 2009.

[Pierce 2007] V. U. Pierce, "Combinatoric results for graphical enumeration and the higher Catalan numbers", 2007. arXiv 0703160v1

[Schaeffer 1997] G. Schaeffer, "Bijective census and random generation of Eulerian planar maps with prescribed vertex degrees", *Electron. J. Combin.* **4**:1 (1997), R20.

[Stanley 1988] R. P. Stanley, "Differential posets", *J. Amer. Math. Soc.* **1**:4 (1988), 919–961.

[Szegő 1939] G. Szegő, *Orthogonal polynomials*, AMS Colloquium Publications **23**, Amer. Math. Soc., 1939.

[Tutte 1968] W. T. Tutte, "On the enumeration of planar maps", *Bull. Amer. Math. Soc.* **74** (1968), 64–74.

[Whitham 1974] G. B. Whitham, *Linear and nonlinear waves*, Pure and Applied Mathematics, Wiley, New York, 1974.

ercolani@math.arizona.edu   *Department of Mathematics, The University of Arizona, 617 N. Santa Rita Avenue, P.O. Box 210089, Tucxon, 85721-0089, United States*