Random Matrices and Their Applications
MSRI Publications
Volume **40**, 2001

Matrix Model Combinatorics: Applications to Folding and Coloring

PHILIPPE DI FRANCESCO

Abstract. We present a detailed study of the combinatorial interpretation of matrix integrals, including the examples of tessellations of arbitrary genera, and loop models on random surfaces. After reviewing their methods of solution, we apply these to the study of various folding problems arising from physics, including: the meander (or polymer folding) problem "enumeration of all topologically inequivalent closed nonintersecting plane curves intersecting a line through a given number of points" and a fluid membrane folding problem reformulated as that of "enumerating all vertex-tricolored triangulations of arbitrary genus, with given numbers of vertices of either color".

Contents

1. Introduction 112
2. Matrix Models and Combinatorics 114
    2.1. Gaussian Integrals and Wick's Theorem 114
    2.2. One-Hermitian Matrix Model: Discrete Random Surfaces 116
    2.3. Multi-Hermitian Matrix Case 122
    2.4. A generating Function for Fatgraphs 125
3. Matrix Models: Solutions 127
    3.1. Reduction to Eigenvalues 127
    3.2. Orthogonal Polynomials 127
    3.3. Large N asymptotics I: Orthogonal Polynomials 130
    3.4. Large N asymptotics II: Saddle-Point Approximation 131
    3.5. Critical Behavior and Asymptotic Enumeration 136
    3.6. Gaussian Words 138
4. Folding Polymers: Meanders 139
    4.1. Definitions and Generalities 139
    4.2. A Simple Algorithm. Numerical Results 142
5. Algebraic Formulation: Temperley–Lieb Algebra 145
    5.1. Definition 145
    5.2. Meander Polynomials 146
    5.3. Meander Determinants 147

6. Matrix Model for Meanders 147
  6.1. The Black-And-White Model 148
  6.2. Meander Polynomials and Gaussian Words 150
  6.3. Exact Asymptotics for the Case of Arbitrary Many Rivers 153
  6.4. Exact Meander Asymptotics from Fully-Packed Loop Models Coupled to Two-dimensional Quantum Gravity 154
7. Folding Triangulations 157
  7.1. Folding the Triangular Lattice 158
  7.2. Foldable Triangulations 159
8. Exact Solution 162
  8.1. The Discrete Hirota Equation 162
  8.2. Direct Expansion and Large N Asymptotics 165
9. Conclusion 167
Acknowledgements 167
References 167

1. Introduction

Our first aim of this article is to convince the reader that matrix integrals, exactly calculable or not, can always be interpreted in some sort of combinatorial way as generating functions for decorated graphs of given genus, with possibly specified vertex and/or face valencies. We show this by expressing pictorially the processes involved in computing Gaussian integrals over matrices, what physicists call generically Feynman rules. These matrix diagrammatic techniques have been first developed in the context of quantum chromodynamics in the limit of large number of colors (the size of the matrix) [1; 2], and more recently in the context of two-dimensional quantum gravity, namely the coupling of two-dimensional statistical models (matter theories) to the fluctuations of the two-dimensional space into surfaces of arbitrary topologies (gravity) [3]. These toy models for noncritical string theory are a nice testing ground for physical ideas, and have led to many confirmations of continuum field-theoretical results in quantum gravity. The purely combinatorial aspect of these models has often been treated as side-result, and we believe it deserves more attention, especially in view of some spectacular results. Indeed, the noncritical string machinery allows one to relate critical properties (such as singularities of thermodynamic quantities) of the flat space statistical models to those of the same models defined on random surfaces [4]. This has led for instance to recent progress in the study of random walks, by using the inverse relation to deduce flat space results from gravitational ones [5].

Once the connection is made between a combinatorial graph-related problem and some matrix model, we still have to compute the integral. Several powerful techniques have been developed, mainly in the context of various branches of physics, to compute those integrals. The original one is orthogonal polynomials [6], but it only applies to "simple" models. More general is the saddle-point technique [2], but it only allows for computing these integrals in the limit of large size of the matrices.

In this article, we wish to present applications of the combinatorics of matrix models to some specific questions arising in physics having to do with folding. Folding problems arise in biology and physics in particular when considering polymers or membranes. An ideal polymer is a chain of say n identical constituents represented by segments (chemical bonds) attached to one another by their ends (atoms), around which they can rotate freely. Membranes are two-dimensional generalizations of polymers, that is, reticulated networks made of vertices (atoms) linked by edges (chemical bonds), either in the form of a regular lattice (tethered membranes) or in the form of networks with arbitrary vertex valencies (fluid membranes). Ideally, imposing that all bonds be rigid, the only possibility for a polymer or membrane to change its spatial configuration is through folding, in which atoms (for polymers) or bonds (for membranes) serve as hinges. Quantities of interest for physicists are thermodynamic ones, characteristic of the systems when their size is large. Mathematically, these correspond to asymptotics of say the numbers of distinct configurations of folding of polymers or membranes of given length or area, when the latter tend to infinity. Both in the case of polymer and fluid membrane folding, we will present matrix models allowing for the calculation of such thermodynamic properties.

The polymer folding problem is better known in mathematics as the "meander problem", namely that of enumerating all the topologically inequivalent configurations of a closed nonselfintersecting plane curve intersecting a line through a given number of points. This problem apparently first emerged in some work by Poincaré in the beginning of the century, and reemerged in various areas of mathematics [7; 8; 9], from recreational mathematics to the 16th Hilbert problem to computer science to the theory of knots and links. In physics, the formulation as a folding problem and the relation to matrix integrals [10] have brought very different developments both algebraic [12; 13; 14; 15] and numerical [11; 16; 17]. As a highly nontrivial outcome of our study, we will present the exact meander asymptotics recently derived in [18].

The article is organized as follows. Section 2 gives a detailed presentation of the combinatorial interpretation of Hermitian matrix integrals as generating tools for fatgraphs, and Section 3 reviews their various methods of solution.

Three sections are then devoted to the application of these ideas to the problem of enumerating all distinct compact folding configurations of a closed or open polymer, namely the number of distinct ways to fold a (self-avoiding) chain of identical constituents onto itself: this is known as the meander or semimeander problem. After defining the problem and reviewing a few known results in Section 4, we present in Section 5 an algebraic formulation of the counting problem within the framework of the Temperley–Lieb algebra, omnipresent in the integrable statistical models, as well as the theory of knots and links. This connection produces remarkable results, like the exact expression for the meander determinant, a meander-related quantity. In Section 6, we connect the meander and related problems to a multimatrix integral, that allows for the exact determination of meandric configuration exponents, in particular through the abovementioned connection [4] between flat and curved space models.

The following two sections study a matrix model for generating foldable triangulations, a sort of two-dimensional generalization of the meander problem, but without the self-avoidance constraint. These triangulations form a simple model for fluid membranes encountered in physics and biology. After posing the problem in Section 7, we introduce a two-matrix integral generating these triangulations, and solve it in various ways in Section 8.

The last section gathers a few concluding remarks.

2. Matrix Models and Combinatorics

The aim of this and the next section is to familiarize the reader with the use of matrix integrals as tools for generating and enumerating graphs with various decorations. These correspond in turn to physical models of matter coupled to two-dimensional quantum gravity, in the form of fluctuating surfaces of arbitrary genus. In the following, we first present the combinatorial tools and give recipes to construct ad-hoc matrix integrals for various graph enumeration problems. We then expose various methods of computation of these matrix integrals, concentrating in particular on the planar graph limit.

2.1. Gaussian Integrals and Wick's Theorem. Consider the Gaussian average

$$\langle x^{2n}\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}x^{2n}\,dx = (2n-1)!! = \frac{(2n)!}{2^n n!}. \tag{2–1}$$

Among the many ways to compute this integral, let us choose the so-called source integral method, namely define the Gaussian source integral

$$\Sigma(s) = \langle e^{xs}\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2+sx}\,dx = e^{s^2/2}.$$

Then the average (2–1) is obtained by taking $2n$ derivatives of $\Sigma(s) = e^{s^2/2}$ with respect to s and by setting $s = 0$ in the end. It is then immediate to see that these derivatives must be taken by *pairs*, in which one derivative acts on the exponential and the other one on the prefactor s. Parallelly, we note that $(2n-1)!! = (2n{-}1)(2n{-}3) \times \cdots \times 3 \times 1$ is the total number of distinct associations of $2n$ objects into n pairs. We may therefore formulate pictorially the computation of (2–1) as follows.

We first draw a star-graph (see Figure 1), with one central vertex and $2n$ outcoming half-edges labeled 1 to $2n$ clockwise, one for each x in the integrand (this amounts to labeling the x's in x^{2n} from 1 to $2n$). Now the pairs of derivatives taken on the source integral are in one-to-one correspondence with pairs of edges

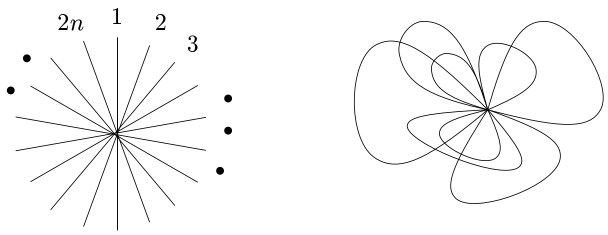


Figure 1. A star diagram with one vertex and $2n$ out-coming half-edges stands for the integrand x^{2n}. In the second diagram, we have represented one nonzero contribution to $\langle x^{2n}\rangle$ obtained by taking derivatives of $\Sigma(s)$ by pairs represented as the corresponding connections of half-edges (the x's and s's are dual to one another, hence derivatives with respect to s are in one to one correspondence with insertions of x in the integrand).

in the pictorial representation. Moreover, to get a nonzero contribution to $\langle x^{2n}\rangle$, we must saturate the set of $2n$ legs by taking n pairs of them. We represent each such saturation by connecting the corresponding edges as in Figure 1. We get exactly $(2n-1)!!$ distinct closed star-graphs with one vertex. We may therefore write the one-dimensional version of Wick's theorem

$$\langle x^{2n}\rangle = \sum_{\text{pairings}} \prod \langle x^2\rangle, \tag{2–2}$$

where the sum extends over all pairings saturating the $2n$ half-edges, and the weight is simply the product over all the edges formed of the corresponding averages $\langle x^2\rangle = (d^2/ds^2)\Sigma(s)|_{s=0} = 1$. Each saturation forms a "Feynman diagram" of the Gaussian average. The edge pairings are called propagators (with value 1 here). The Feynman rules are simply the set of values of these propagators. This may appear like a complicated way of writing a rather trivial result, but it suits our purposes for generalization to matrix models and graphs. Note that the pictorial interpretation we have given for the computation of $\langle x^{2n}\rangle$ is not unique. For instance, we could have arbitrarily split the average into $\langle x^p x^q\rangle$ for some integers p, q with $p + q = 2n$. Then we would have rather represented *two* vertices with respectively p and q out-coming half-edges. Wick's theorem (2–2) states that we must now sum over (possibly disconnected) graphs obtained by saturating the $p + q$ half-edges by pairs. The result of course is still the same.

As a last remark, and to make the contact with graphs, we may consider for instance the formal series expansion

$$z(g_1, g_2, \ldots) = \langle e^{\sum_{i\geq 1} g_i x^i}\rangle \tag{2–3}$$

in powers of $g_1, g_2, \ldots$, whose coefficients are computed using (2–1). Thanks to the previous pictorial interpretation, we may compute a typical term in the series expansion, say the coefficient $\langle \prod (x^i)^{V_i}\rangle$ of $\prod g_i^{V_i}/V_i!$, by drawing V_i star diagrams with i half-edges, $i = 1, 2, \ldots$, and saturating the $2E = \sum iV_i$ half-edges in all possible ways by forming E pairs. Hence computing (2–3) is reinterpreted as a generating function of graphs with specified numbers of vertices of given valencies.

2.2. One-Hermitian Matrix Model: Discrete Random Surfaces. We now repeat the calculations of the previous section with the following Gaussian Hermitian matrix average of an arbitrary function f

$$\langle f(M)\rangle = \frac{1}{Z_0(N)}\int dM\, e^{-N\,\mathrm{Tr}(M^2/2)} f(M),$$

where the integral extends over Hermitian $N\times N$ matrices, with the standard Haar measure $dM = \prod_i dM_{ii}\prod_{i<j} dRe(M_{ij})dIm(M_{ij})$, and the normalization factor $Z_0(N)$ is fixed by requiring that $\langle 1\rangle = 1$ for $f = 1$. Typically, we may take for f a monomial of the form $f(M) = \prod_{(i,j)\in I} M_{ij}$, I a finite set of pairs of indices. Note the presence of the normalization factor N (=the size of the matrices) in the exponential. Note also the slight abuse of notation as we still denote averages with the same bracket sign as in previous section: we may simply include the case of the previous section as the particular case of integration over 1×1 Hermitian matrices (that is, real numbers) here.

As in the one-dimensional case of the previous section, for a given Hermitian $N\times N$ matrix S we introduce the source integral

$$\Sigma(S) = \langle e^{\mathrm{Tr}(SM)}\rangle = e^{\mathrm{Tr}(S^2)/(2N)},$$

easily obtained by completing the square $M^2 - N(SM+MS) = (M-NS)^2 - N^2S^2$ and performing the change of variable $M' = M - NS$. We can use this equation to compute any average of the form

$$\langle M_{ij}M_{kl}\ldots\rangle = \frac{\partial}{\partial S_{ji}}\frac{\partial}{\partial S_{lk}}\cdots\ \Sigma(S)\big|_{S=0}. \tag{2–4}$$

Note the interchange of the indices due to the trace $\mathrm{Tr}(MS) = \sum M_{ij}S_{ji}$. As before, derivatives with respect to elements of S must go by pairs, one of which acts on the exponential and the other one on the S element thus created. In particular, a fact also obvious from the parity of the Gaussian, (2–4) vanishes unless there are an even number of matrix elements of M in the average. In the simplest case of two matrix elements, we have

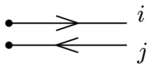
$$\langle M_{ij}M_{kl}\rangle = \frac{\partial}{\partial S_{lk}}\frac{1}{N}S_{ij}e^{\mathrm{Tr}(S^2)/(2N)}\bigg|_{S=0} = \frac{1}{N}\delta_{il}\delta_{jk}. \tag{2–5}$$

Hence the pairs of derivatives must be taken with respect to S_{ij} and S_{ji} for some pair i,j of indices to yield a nonzero result. This leads naturally to the Matrix Wick's theorem:

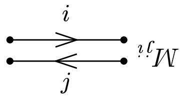
$$\Big\langle \prod_{(i,j)\in I} M_{ij}\Big\rangle = \sum_{\text{pairings}} P \prod_{(ij),(kl)\in P}\langle M_{ij}M_{kl}\rangle, \tag{2–6}$$

where the sum extends over all pairings saturating the (pairs of) indices of M by pairs.

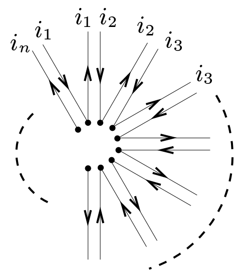
We see that in general, due to the restrictions (2–5) many terms in (2–6) will vanish. We now give a pictorial interpretation for the nonvanishing contributions in (2–6). We represent a matrix element M_{ij} as a half-edge (with a marked end) made of a double-line, each of which is oriented in an opposite direction. We decide that the line pointing from the mark carries the index i, while the other one, pointing to the mark, carries the index j. This reads

$M_{ij} \quad \longleftrightarrow$  .

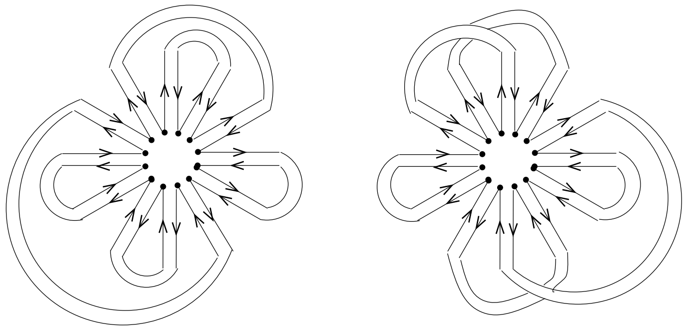
The two-element result (2–5) becomes simply the construction of an edge (with both ends marked) out of two half-edges M_{ij} and M_{kl}, but is nonzero only if the indices i and j are conserved along the oriented lines. This gives pictorially

$\langle M_{ij} M_{ji} \rangle \quad \longleftrightarrow \quad M_{ij}$  . (2–7)

Similarly, an expression of the form $\mathrm{Tr}(M^n)$ will be represented as a star diagram with one vertex connected to n double half-edges in such a way as to respect the identification of the various running indices, namely

$$\mathrm{Tr}(M^n) = \sum_{i_1, i_2, \ldots, i_n} M_{i_1 i_2} M_{i_2 i_3} \ldots M_{i_n i_1} \longleftrightarrow$$  . (2–8)

As a first application of this diagrammatic interpretation of the Wick theorem (2–6), let us compute the large N asymptotics of $\langle \mathrm{Tr}(M^n) \rangle$. To compute $\langle \mathrm{Tr}(M^n) \rangle$, we must first draw a star diagram as in (2–8), then apply (2–6) to express the result as a sum over the saturations of the star with edges connecting its outcoming half-edges by pairs. To get a nonzero result, we must clearly have n even, say $n = 2p$. Note then that there are $(2p-1)!!$ such pairings, allowing us to recover the result of the previous section by setting $N = 1$, and simply replacing all oriented double lines by unoriented single ones. But if instead we take N to be large, we see that only a fraction of these $(2p-1)!!$ pairings will contribute at leading order. Indeed, assume first we restrict the set of pairings to "planar" ones (see Figure 2(a)), namely such that the saturated star diagrams have a petal structure in which edges only connect pairs of half-edges of the form (ij), $i < j$ and (kl), $k < l$ with either $j < k$ or $k < i < j < l$ or $i < k < l < j$ (in



(a) (b)

Figure 2. An example of planar (petal) diagram (a) and a nonplanar one (b). Both diagrams have $n = 2p = 12$ half-edges, connected with $p = 6$ edges. The diagram (a) has $p + 1 = 7$ faces bordered by oriented loops, whereas (b) only has 3 of them. The Euler characteristic reads $2 - 2h = F - E + 1$ ($V = 1$ in both cases), and gives the genus $h = 0$ for (a), and $h = 2$ for (b).

the labeling of half-edges, we have taken cyclic boundary conditions, namely the labels 1 and $n+1$ are identified). In other words, the petals are either juxtaposed or included into one-another. We may compute the genus of the petal diagrams by noting that they form a tessellation of the sphere (=plane plus point at infinity). This tessellation has $V = 1$ vertex (the star), $E = p$ edges, and F faces, including the "external" face containing the point at infinity. The planarity of the diagram simply expresses that its genus h vanishes, namely

$$2 - 2h = 2 = F - E + V = F + 1 - p \quad \Rightarrow \quad F = p + 1.$$

Such diagrams receive a total contribution $1/N^p$ from the propagators (weight $1/N$ per connecting edge), but we still have to sum over the remaining matrix indices $j_1, j_2, \dots, j_{p+1}$ running over the $p + 1$ oriented loops we have created, which form the boundaries of the $F = p + 1$ faces. This gives an extra weight N per loop (or face) of the diagram, hence a total contribution of N^{p+1}. So all the petal diagrams contribute the same total factor N to $\langle \mathrm{Tr}(M^n) \rangle$. Now any nonpetal (that is, nonplanar; see Figure 2(b)) diagram must have at least *two less* oriented loops, hence contributes at most $1/N$ to $\langle \mathrm{Tr}(M^n) \rangle$. Indeed, its Euler characteristic is negative or zero, hence it has $F \le E - V = p - 1$ and it contributes at most for $N^{F-p} \le 1/N$. So, to leading order in N, only the genus zero (petal) diagrams contribute. We simply have to count them. This is a standard problem in combinatorics: one may for instance derive a recursion relation for the number c_p of petal diagrams with $2p$ half-edges, by fixing the left end of an edge (say at position 1), and summing over the positions of its right end (at positions $2j$, $j = 1, 2, \dots, p$), and noting that the petal thus formed may contain c_{j-1} distinct petal diagrams and be next to c_{p-j} distinct ones. This gives the recursion relation

$$c_p = \sum_{j=1}^{p} c_{j-1} c_{p-j} \qquad c_0 = 1,$$

solved by the Catalan numbers

$$c_p = \frac{(2p)!}{(p+1)!p!}. \tag{2–9}$$

Finally, we get the one-matrix planar Gaussian average

$$\lim_{N\to\infty} \frac{1}{N} \langle \mathrm{Tr}(M^n) \rangle = \begin{cases} c_p & \text{if } n = 2p, \\ 0 & \text{otherwise.} \end{cases} \tag{2–10}$$

This exercise shows us what we have gained by considering $N \times N$ matrices rather than numbers: we have now a way of discriminating between the various genera of the graphs contributing to Gaussian averages. This fact will be fully exploited in the next example.

We would now like to present an application of (2–6) with important physical and mathematical consequences. We apply the matrix Wick theorem (2–6) to the following generating function $f(M) = \exp(N \sum_{i\geq 1} g_i \,\mathrm{Tr}(M^i)/i)$, to be understood as a formal power series of g_i, for $i = 1, 2, 3, 4, \ldots$:

$$\begin{aligned} Z_N(g_1, g_2, \ldots) &= \langle e^{N \sum_{i\geq 1} g_i \,\mathrm{Tr}(M^i/i)} \rangle \\ &= \sum_{n_1, n_2, \ldots \geq 0} \prod_{i\geq 1} \frac{(Ng_i)^{n_i}}{i^{n_i} n_i!} \left\langle \prod_{i\geq 1} \mathrm{Tr}(M^i)^{n_i} \right\rangle \\ &= \sum_{n_1, n_2, \ldots \geq 0} \prod_{i\geq 1} \frac{(Ng_i)^{n_i}}{i^{n_i} n_i!} \sum N^{-E(\Gamma)} N^{F(\Gamma)}, \end{aligned} \tag{2–11}$$

where the last sum is over all labeled fatgraphs Γ with n_i i-valent vertices. This comes from a direct application of (2–6).

gravi

In (2–11), we have first represented pictorially the integrand $\prod_i (\mathrm{Tr}(M^i))^{n_i}$ as a succession of n_i i-valent star diagrams like that of (2–8), $i = 1, 2, \ldots$. Then we have summed over all possible saturations of all the marked half-edges of all these stars, thus forming (not necessarily connected) ribbon or fatgraphs Γ with some labeling of their half-edges (see Figure 6–15 for an example of connected fatgraph). In (2–11), we have denoted by $E(\Gamma)$ the total number of edges of Γ, connecting half-edges by pairs, that is, the number of propagators needed (yielding a factor $1/N$ each, from (2–5)). The number $F(\Gamma)$ is the total number of faces of Γ. The faces of Γ are indeed well-defined because Γ is a fatgraph, that is, a graph with edges made of doubly oriented parallel lines carrying the corresponding matrix indices $i = 1, 2, \ldots, N$: the oriented loops we have created by the pairing process are interpreted as face boundaries, in one-to-one correspondence with faces of Γ. But the traces of the various powers

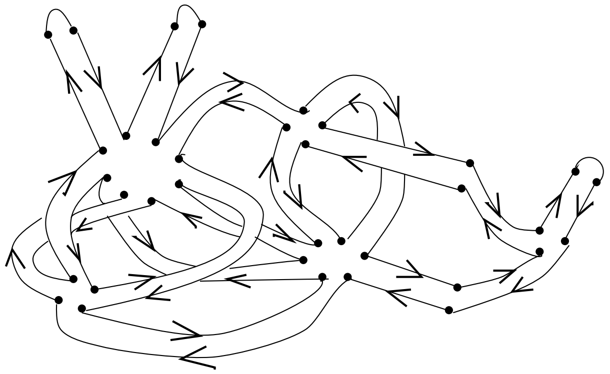


Figure 3. A typical connected fatgraph Γ, corresponding to the average $\langle \mathrm{Tr}(M)^3\,\mathrm{Tr}(M^2)^2\,\mathrm{Tr}(M^3)\,\mathrm{Tr}(M^4)^2\,\mathrm{Tr}(M^6)\,\mathrm{Tr}(M^8)\rangle$. The graph was obtained by saturating the ten star diagrams corresponding to the ten trace terms, namely with $n_1 = 3$ univalent vertices, $n_2 = 2$ bi-valent ones, $n_3 = 1$ tri-valent one, $n_4 = 2$ four-valent ones, $n_6 = 1$ six-valent one and $n_8 = 1$ eight-valent one, hence a total of $V = 10$ vertices. This graph corresponds to one particular Wick pairing for which we have drawn the $E = 16$ connecting edges, giving rise to $F = 2$ oriented loops bordering the faces of Γ.

of M still have to be taken, which means all the indices running from 1 to N have to be summed over all these loops. This results in the factor N per face of Γ in (2–11). Finally, the sum extends over all (possibly disconnected) fatgraphs Γ with labeled half-edges. Each such labeled graph corresponds to exactly one Wick pairing of (2–6). Summing over all the possible labelings of a given unlabeled fatgraph Γ results in some partial cancellation of the symmetry prefactors $\prod_i 1/(i^{n_i} n_i!)$, which actually leaves us with the inverse of the order of the symmetry group of the unlabeled fatgraph Γ, denoted by $1/|\mathrm{Aut}\,\Gamma|$. This gives the final form

$$Z_N(g_1, g_2, \ldots) = \sum_{\substack{\text{fatgraphs}\\ \Gamma}} \frac{N^{V(\Gamma)-E(\Gamma)+F(\Gamma)}}{|\mathrm{Aut}\,\Gamma|} \prod_{i\geq 1} g_i^{n_i(\Gamma)}, \qquad (2\text{–}12)$$

where $n_i(\Gamma)$ denotes the total number of i-valent vertices of Γ and $V(\Gamma) = \sum_i n_i(\Gamma)$ is the total number of vertices of Γ. To restrict the sum in (2–12) to only connected graphs, we simply have to formally expand the logarithm of Z_N, resulting in the final identity

$$F_N(g_1, g_2, \ldots) = \mathrm{Log}\, Z_N(g_1, g_2, \ldots) = \sum_{\substack{\text{connected}\\ \text{fatgraphs }\Gamma}} \frac{N^{2-2h(\Gamma)}}{|\mathrm{Aut}\,\Gamma|} \prod_i g_i^{n_i(\Gamma)}, \qquad (2\text{–}13)$$

where we have identified the Euler characteristic $\chi(\Gamma) = F - E + V = 2 - 2h(\Gamma)$, where $h(\Gamma)$ is the genus of Γ (number of handles). Equation (2–13) gives a clear geometrical meaning to the Gaussian average of our choice of $f(M)$: it amounts to computing the generating function for fatgraphs of given genus and given vertex valencies. Such a fatgraph Γ is in turn dual to a tessellation Γ^* of a Riemann surface of same genus, by means of n_i i-valent polygonal tiles, $i = 1, 2, \ldots$.

The result (2–13) is therefore a statistical sum over discretized random surfaces (the tessellations), that can be interpreted in physical terms as the free energy of two-dimensional quantum gravity. The name free energy stands generically for the logarithm of the partition function Z_N. It simply identifies the Gaussian matrix integral with integrand $f(M)$ as a discrete sum over configurations of tessellated surfaces of arbitrary genera, weighted by some exponential factor. More precisely, imagine only $g_3 = g \neq 0$ while all other g_i's vanish. Then (2–13) becomes a sum over fatgraphs with cubic vertices (also called ϕ^3 fatgraphs), dual to triangulations T of Riemann surfaces of arbitrary genera. Assuming these triangles have all unit area, then $n_3(\Gamma) = A(T)$ is simply the total area of the triangulation T. Hence (2–13) becomes

$$F_N(g) = \sum_{\text{connected triangulations } T} \frac{g^{A(T)} N^{2-2h(T)}}{|\mathrm{Aut}\, T|} \tag{2–14}$$

and the summand $g^A N^{2-2h} = e^{-S_E}$ is nothing but the exponential of the discrete version of Einstein's action for General Relativity in 2 dimensions, which reads, for any surface S

$$\begin{aligned} S_E(\Lambda, \mathcal{N}|S) &= \Lambda \int_S d^2x \sqrt{|g|} + \frac{\mathcal{N}}{4\pi} \int_S d^2x \sqrt{|g|} R \\ &= \Lambda A(S) + \mathcal{N}(2 - 2h(S)), \end{aligned}$$

where g is the metric of S, R its scalar curvature, and the two multiplicative constants are respectively the cosmological constant Λ and the Newton constant $\mathcal{N}$. In the preceding equation, we have identified the two invariants of S: its area $A(S)$ and its Euler characteristic $\chi(S) = 2 - 2h(S)$ (using the Gauss–Bonnet formula). The contact with (2–14) is made by setting $g = e^{-\Lambda}$ and $N = e^{-\mathcal{N}}$. If we now include all g_i's in (2–13) we simply get a more elaborate discretized model, in which we can keep track of the valencies of vertices of Γ (or tiles of the dual Γ^*).

Going back to the purely mathematical interpretation of (2–13), we start to feel how simple matrix integrals can be used as tools for generating all sorts of graphs whose duals tessellate surfaces of arbitrary given topology. The size N of the matrix relates to the genus, whereas the details of the integrand relate to the structure of vertices. An important remark is also that the large N limit of (2–13) extracts the genus zero contribution, namely that coming from planar graphs. So as a by-product, it will be possible to extract from asymptotics of matrix integrals for large size N some results on planar graphs.

2.3. Multi-Hermitian Matrix Case. The results of previous section can be easily generalized to multiple Gaussian integrals over several Hermitian matrices. More precisely, let $M_1,\ M_2,\ \dots M_p$ denote p Hermitian matrices of same size $N\times N$, and $Q_{a,b},\ a,b=1,2,\dots,p$ the elements of a positive definite form Q. We consider the multiple Gaussian integrals of the form

$$\langle f(M_1,\dots,M_p)\rangle = \frac{\int dM_1\dots dM_p\, e^{-(N/2)\sum_{a,b=1}^{p}\mathrm{Tr}(M_aQ_{ab}M_b)}f(M_1,\dots,M_p)}{\int dM_1\dots dM_p\, e^{-(N/2)\sum_{a,b=1}^{p}\mathrm{Tr}(M_aQ_{ab}M_b)}}, \quad (2\text{–}15)$$

where by a slight abuse of notation we still denote averages by the same bracket sign as before: the one-Hermitian matrix case of the previous section corresponds simply to $p=1$ and $Q_{1,1}=1$. The averages (2–15) are computed by extending the source integral method of previous section: for some Hermitian matrices $S_1,\dots,S_p$ of size $N\times N$, we define and compute the multisource integral

$$\Sigma(S_1,\dots,S_p)=\langle e^{\sum_{a=1}^{p}\mathrm{Tr}(S_aM_a)}\rangle = e^{(1/2N)\sum_{a,b=1}^{p}\mathrm{Tr}(S_a(Q^{-1})_{a,b}S_b)}$$

and use multiple derivatives thereof to compute any expression of the form (2–15), before taking $S_a\to 0$. As before, derivatives with respect to elements of the S's must go by pairs to yield a nonzero result. For instance, in the case of two matrix elements of M's we find the propagators

$$\big\langle (M_a)_{ij}(M_b)_{kl}\big\rangle = \frac{1}{N}\delta_{il}\delta_{jk}(Q^{-1})_{a,b}. \quad (2\text{–}16)$$

In general we will apply the multimatrix Wick theorem

$$\Big\langle \prod_{(a,i,j)\in J}(M_a)_{ij}\Big\rangle = \sum_{\substack{\text{pairings}\\ P}}\ \prod_{\substack{\text{pairs}\\ (aij),(bkl)\in P}}\big\langle (M_a)_{ij}(M_b)_{kl}\big\rangle \quad (2\text{–}17)$$

expressing the multimatrix Gaussian average of any product of matrix elements of the M's as a sum over all pairings saturating the matrix half-edges, weighted by the corresponding value of the propagator (2–16). Note that half-edges must still be connected according to the rule (2–7), but that in addition, depending on the form of Q, some matrices may not be allowed to connect to one another (e.g. if $(Q^{-1})_{ab}=0$ for some a and b, then $\langle M_aM_b\rangle=0$, and there cannot be any edge connecting a matrix with index a to one with index b).

This gives us much freedom in "cooking up" multimatrix models to evaluate generating functions of graphs with specific decorations such as colorings, spin models, etc... This is expected to describe the coupling of matter systems (e.g. a spin model usually defined on a regular lattice) to two-dimensional quantum gravity (by letting the lattice fluctuate into tessellations of arbitrary genera).

An important example is the so-called gravitational $O(n)$ model [19], say on cubic fatgraphs. Its regular lattice version [24] is as spin model defined on the honeycomb (hexagonal) lattice, with only cubic (trivalent) vertices. A spin configuration of the model is simply a map from the set of vertices of the lattice to the n-dimensional unit sphere, hence the name inherited from the obvious symmetry of the target space. For $n = 1$, this is the Ising model, in which the spin variable may only take values ± 1. For $n \geq 2$ this gives an infinite continuum set of configurations, over which we have to perform a statistical sum, henceforth an integration over all the spin values. With a suitable choice of nearest neighbor interaction between adjacent spins, this statistical sum was shown to reduce to the following very simple loop model on the honeycomb lattice. The partition function $Z_{O(n)}$ of the regular lattice model is expressed as a sum over configurations of closed nonintersecting loops drawn on the honeycomb lattice, each weighed by the factor n, whereas each edge of each loop receives a Boltzmann factor K corresponding to the inverse temperature of the statistical mechanical model. The gravitational version thereof is very simple to figure out: we must replace the honeycomb lattice by its spatial fluctuations into arbitrary cubic fatgraphs. Then, on each fatgraph, we must consider arbitrary closed nonintersecting loop configurations, with the same weights as in the regular lattice case. The free energy of this model reads

$$F_N(g|K,n) = \sum_{\substack{\text{connected}\\ \text{cubic fatgraphs } \Gamma}} \frac{g^{A(\Gamma)} N^{2-2h(\Gamma)}}{|\mathrm{Aut}\,\Gamma|} \sum_{\substack{\text{loop}\\ \text{configurations } \ell}} n^{C(\ell)} K^{E(\ell)}, \qquad (2\text{–}18)$$

where as before we have inserted a weight g per cubic vertex ($A(\Gamma)$ is the total area of the triangulation dual to Γ, also the total number of vertices of Γ), and for each loop configuration ℓ on Γ, we have denoted by $C(\ell)$ and $E(\ell)$ respectively the total number of connected components of ℓ (number of loops) and the total number of edges occupied by the loops.

We now present a multimatrix model for (2–18). Loop models can be easily represented by using the so-called *replica trick*: as we wish to attach a weight n per loop, we may introduce a color variable $c = 1, 2, \ldots, n$ by which we may paint each connected component of loop independently. Summing over all possible colorings will yield the desired weight. Hence we are allowing the loops to be *replicated*, namely we may now choose among n different types of loops to build our configurations. To represent these configurations, we need $n+1$ Hermitian matrices of same size $N \times N$, say $A_1, A_2, \ldots, A_n$ for the n colored loop half-edges, and B for the unoccupied half-edges. As loops are nonintersecting the only allowed vertices are of the form B^3 (unoccupied vertex) and $A_c^2 B$ (loop of color c going through a vertex, with exactly one unoccupied edge). The propagators must force loops to conserve the same color on each connected component, while unoccupied half-edges can only be connected to unoccupied ones, namely

$$\big\langle (A_c)_{ij} (A_d)_{kl} \big\rangle = \frac{K}{N} \delta_{cd} \delta_{il} \delta_{jk}, \quad \langle B_{ij} B_{kl} \rangle = \frac{1}{N} \delta_{il} \delta_{jk}, \quad \big\langle (A_c)_{ij} B_{kl} \big\rangle = 0,$$

where we also have added the weight K per edge of loop. This defines the inverse of the quadratic form Q we need for the corresponding matrix model, according to (2–16). This leads to the multimatrix integral

$$\begin{aligned} Z_N(g|K,n) &= \frac{\int dB\, dA_1 \ldots dA_n\, e^{-N\,\mathrm{Tr}(V(B,A_1,\ldots,A_n))}}{\int dB\, dA_1 \ldots dA_n\, e^{-N\,\mathrm{Tr}(V_0(B,A_1,\ldots,A_n))}} \\ V(B,A_1,\ldots,A_n) &= \frac{1}{2}B^2 + \frac{1}{2K}\sum_{c=1}^{n} A_c^2 - \frac{g}{3}B^3 - gB\sum_{c=1}^{n} A_c^2 \\ V_0(B,A_1,\ldots,A_n) &= \frac{1}{2}B^2 + \frac{1}{2K}\sum_{c=1}^{n} A_c^2. \end{aligned} \tag{2–19}$$

To get the free energy of the gravitational $O(n)$ model (2–18), we simply have to take the logarithm of (2–19), to extract the contribution from connected fatgraphs only. As before, we may decorate the model with higher order terms of the form $\sum_i g_i B^i/i$ to allow for arbitrary unoccupied vertex configurations and keep track of the fine structure of the fatgraphs, without altering the loop configurations. Another possibility is to remove the B^3 term from (2–19), to restrict the loop configurations to compact ones, that is, covering all the vertices of Γ.

Another standard example, with a more complicated quadratic form, is the so-called gravitational q-states Potts model [20], defined say on arbitrary cubic diagrams Γ. Imagine that the diagram Γ is decorated by maps σ from its set of vertices $v(\Gamma)$ to $\mathbb{Z}_q$. We may alternatively view this as "spin" (or color) variables $\sigma \in \mathbb{Z}_q$ living on the vertices of Γ. The model is then defined by attaching Boltzmann factors corresponding to different energies of configurations of neighboring spins according to whether they are identical or distinct. More precisely, if v, v' are two vertices connected by an edge e in Γ, we define the weight

$$w_e(\sigma) = e^{K\delta_{\sigma(v),\sigma(v')}}$$

for some positive real parameter K (inverse temperature). The gravitational Potts model free energy is then defined as

$$F_N(g|K,q) = \sum_{\substack{\text{connected} \\ \text{cubic fatgraphs } \Gamma}} \frac{g^{A(\Gamma)} N^{2-2h(\Gamma)}}{|\mathrm{Aut}\,\Gamma|} \sum_{\substack{\text{maps} \\ \sigma: v(\Gamma)\to\mathbb{Z}_q}} \prod_{\substack{\text{edges} \\ e \text{ of } \Gamma}} w_e(\sigma). \tag{2–20}$$

We now "cook up" a multimatrix model integral for (2–20). We wish to generate all connected cubic fatgraphs Γ: this would be easily done by considering the example of the previous section with all g's equal to zero except $g_3 = g$. To represent the spin configurations however, we need to be able to distinguish between q different types of vertices, according to the value of $\sigma(v) \in \mathbb{Z}_q$. This is done by introducing q Hermitian matrices $M_1, \ldots, M_q$ with the same size $N \times N$, with a Gaussian potential of the form $N\,\mathrm{Tr}(\sum M_a Q_{ab} M_b)/2$, such that

$$(Q^{-1})_{ab} = 1 + (e^K - 1)\delta_{ab} \tag{2–21}$$

in order for the propagators (2–16) to receive the extra weight e^K when the "spins" a and b are identical, and 1 when they are distinct. Equation (2–21) is easily inverted into

$$Q_{ab} = \frac{1}{e^K - 1}\Big(\delta_{ab} - \frac{1}{e^K - 1 + q}\Big). \tag{2–22}$$

The free energy of the gravitational Potts model on cubic diagrams is therefore given by

$$\begin{aligned} F_N(g|K,q) &= \operatorname{Log} Z_N(g|K,q), \\ Z_N(g|K,q) &= \frac{\int dM_1 \dots dM_q \, e^{-N \operatorname{Tr} V(M_1,\dots,M_q)}}{\int dM_1 \dots dM_q \, e^{-N \operatorname{Tr} V_0(M_1,\dots,M_q)}}, \\ V(M_1,\dots,M_q) &= \frac{1}{2} \sum_{a,b=1}^{q} Q_{ab} M_a M_b - \frac{g}{3} \sum_{a=1}^{q} M_a^3, \\ V_0(M_1,\dots,M_q) &= \frac{1}{2} \sum_{a,b=1}^{q} Q_{ab} M_a M_b, \end{aligned} \tag{2–23}$$

with Q as in (2–22). When $q = 2$ the model is nothing but the Ising model (also equivalent to the $O(n = 1)$ model, as mentioned above). To recover (2–20) from (2–23), we simply apply the multimatrix Wick theorem, and note that the overall contribution for each graph is $N^{V-E+F} = N^{2-2h}$ as before, as we get a weight N per vertex, $1/N$ per edge from propagators, and N per oriented loop of indices, irrespectively of the spin values.

2.4. A generating Function for Fatgraphs. In this section, we use the philosophy developed in the two previous ones to present a model of *dually weighted* fatgraphs, namely in which we wish to specify not only the structure of vertices but that of faces as well [21]. This uses the slightly more general notion of index-dependent propagators, obtained by considering one-Hermitian matrix models with Gaussian potentials of the form

$$V(M) = \frac{1}{2} \sum_{ijkl} M_{ij} Q_{ij;kl} M_{kl}, \tag{2–24}$$

where Q is some arbitrary tensor depending on the four matrix indices, but such that $\operatorname{Tr} V(M)$ remains positive. Repeating the calculations of Section 2.3, we find propagators of the form

$$\langle M_{ij} M_{kl} \rangle = \frac{1}{2N} \Big(\frac{1}{Q_{ij;kl}} + \frac{1}{Q_{kl;ij}}\Big). \tag{2–25}$$

This gives even more freedom in tailoring specific models for combinatorial purposes. We now examine the choice

$$Q_{ij;kl} = \lambda_i \lambda_j \delta_{il} \delta_{jk}$$

for some $\lambda_1,\dots,\lambda_N>0$. We will also refer to the $N\times N$ diagonal matrix Λ with entries $\Lambda_{ij}=\lambda_i\delta_{ij}$. Then the propagator (2–25) becomes

$$\langle M_{ij}M_{kl}\rangle=\frac{1}{N\lambda_i\lambda_j}\delta_{il}\delta_{jk}. \tag{2–26}$$

If we consider the Gaussian average with respect to (2–24) of the function $f(M)=\exp(-N\sum_{i\ge1}g_i\operatorname{Tr}(M^i)/i)$, we find the following interpretation after taking the logarithm:

$$\begin{aligned}F_N(\Lambda;g_1,g_2,\dots)&=\operatorname{Log}\langle e^{-N\sum_{i\ge1}(g_i/i)\operatorname{Tr}(M^i)}\rangle\\&=\sum_{\substack{\text{connected}\\\text{fatgraphs }\Gamma}}\frac{N^{2-2h(\Gamma)}}{|\operatorname{Aut}\Gamma|}\prod_{i\ge1}g_i^{n_i(\Gamma)}\prod_{\substack{\text{faces}\\ f\text{ of }\Gamma}}\sum_{i=1}^{N}\frac{1}{\lambda_i^{m(f)}},\end{aligned}\tag{2–27}$$

where, for each face f of Γ, we have denoted by $m(f)$ the perimeter of its boundary (number of adjacent edges, or valency). The latter term comes from the summation over all the indices running over the oriented loops of the graph. But this time, we do not use $\sum_{i=1}^N 1=N$, but rather notice that each edge of a loop carrying the running index i comes with a factor $1/\lambda_i$ from the propagators (2–26). Hence each face comes with a total factor $1/\lambda_i^{m(f)}$, $m(f)$ the face valency, which has to be summed over from $i=1$ to N, to yield (2–27). Introducing the variables

$$G_k=\operatorname{Tr}(\Lambda^{-k})=\sum_{i=1}^{N}\frac{1}{\lambda_i^k}, \tag{2–28}$$

we finally get the expansion

$$F_N(\Lambda;g_1,g_2,\dots)=\sum_{\text{conn. fatgraphs }\Gamma}\frac{N^{2-2h(\Gamma)}}{|\operatorname{Aut}\Gamma|}\prod_{i\ge1}g_i^{n_i(\Gamma)}G_i^{m_i(\Gamma)}, \tag{2–29}$$

where $m_i(\Gamma)$ denotes the number of i-valent faces of Γ. The expression (2–29) looks almost symmetric in g_i and G_i: it would actually be the case if the G_i were an infinite set of independent variables like the g_i's, but they are not. Due to the definition (2–28), only the N first G's are generically independent variables, all the others are dependent. Here is how we should read (2–29). For any given N, let us expand $F_N(\Lambda;g_1,g_2,\dots)$ as a sum over fatgraphs with $\le N$ edges. Then the right-hand side only involves the numbers $G_1,\dots,G_N$, as well as $g_1,\dots,g_N$, and is manifestly symmetric in the g's and G's, as the two sets of variables are exchanged by going to the dual graphs. The coefficient of each monomial corresponds to the graphs with specified genus, numbers of i-valent vertices, and numbers of i-valent faces. To get information on larger graphs, we simply have to take N larger and larger.

3. Matrix Models: Solutions

We hope that this series of examples has convinced the reader of the relevance of matrix integrals to graph combinatorics. We will now see how to extract useful information from these matrix integrals. In this section, we will mainly cover the one-matrix integrals defined in Section 2.2. Multi-matrix techniques are very similar, and we will present them later, when we need them. More precisely, we will study the one-matrix integral

$$Z_N(V) = \frac{\int dM\, e^{-N \operatorname{Tr} V(M)}}{\int dM\, e^{-N \operatorname{Tr} V_0(M)}}, \tag{3–1}$$

with an arbitrary polynomial potential $V(x) = \frac{x^2}{2} + \sum_{3 \le i \le k} \frac{g_i}{i} x^i$ and $V_0(x) = \frac{x^2}{2}$. This contains as a limiting case the partition function (2–11). We are not worrying at this point about convergence issues for these integrals, as they must be understood as formal tools allowing for computing well-defined coefficients in formal series expansions in the g's. (The reader may be more comfortable assuming $\mathrm{Re}(g_k) > 0$ to actually ensure convergence throughout this section.)

3.1. Reduction to Eigenvalues. The step zero in computing the integral (3–1) is the reduction to N one-dimensional integrals, namely over the real eigenvalues $m_1, \dots, m_N$ of the Hermitian matrix M. This is done by performing the change of variables $M \to (m, U)$, where $m = \mathrm{diag}(m_1, \dots, m_N)$, and U is a unitary diagonalization matrix such that $M = UmU^\dagger$, hence $U \in U(N)/U(1)^N$ as U may be multiplied by an arbitrary matrix of phases. The Jacobian of the transformation is readily found to be the squared Vandermonde determinant

$$J = \Delta(m)^2 = \prod_{1 \le i < j \le N} (m_i - m_j)^2 .$$

Performing the change of variables in both the numerator and denominator of (3–1) we obtain

$$Z = \frac{\int_{\mathbb{R}^N} dm_1 \dots dm_N\, \Delta(m)^2 e^{-N \sum_{i=1}^N V(m_i)}}{\int_{\mathbb{R}^N} dm_1 \dots dm_N\, \Delta(m)^2 e^{-N \sum_{i=1}^N \frac{m_i^2}{2}}}. \tag{3–2}$$

3.2. Orthogonal Polynomials. The standard technique of computation of (3–2) uses orthogonal polynomials. The idea is to disentangle the Vandermonde determinant squared interaction between the eigenvalues. The solution is based on the following simple lemma: If $p_m(x) = x^m + \sum_{j=0}^{m-1} p_{m,j} x^j$ are monic polynomials of degree m, for $m = 0, 1, \dots, N-1$, then

$$\Delta(m) = \det(m_i^{j-1})_{1 \le i,j \le N} = \det(p_{j-1}(m_i))_{1 \le i,j \le N}. \tag{3–3}$$

This is easily proved by performing suitable linear combinations of columns. We now introduce the unique set of monic polynomials p_m, of degree $m =$

$0, 1, \ldots, N-1$, that are orthogonal with respect to the one-dimensional measure $d\mu(x) = \exp(-NV(x))dx$, namely such that

$$(p_m, p_n) = \int_{\mathbb{R}} p_m(x) p_n(x)\, d\mu(x) = h_m \delta_{m,n}.$$

They allow us to rewrite the numerator of (3–2), using (3–3), as

$$\sum_{\sigma,\tau \in S_N} \varepsilon(\sigma\tau) \prod_{i=1}^{N} \int_{\mathbb{R}} d\mu(m_i)\, p_{\sigma(i)-1}(m_i) p_{\tau(i)-1}(m_i) e^{-NV(m_i)} = N! \prod_{j=0}^{N-1} h_j. \quad (3\text{–}4)$$

We may apply the same recipe to compute the denominator, with the result $N! \prod_{j=0}^{N-1} h_j^{(0)}$, where the $h_j^{(0)}$ are the squared norms of the orthogonal polynomials with respect to to the Gaussian measure $d\mu_0(x) = \exp(-Nx^2/2)\, dx$. Hence the h's determine $Z_N(V)$ entirely through

$$Z_N(V) = \prod_{i=0}^{N-1} \frac{h_i}{h_i^{(0)}}. \quad (3\text{–}5)$$

We may repeat this calculation in the presence of a "spectator" term of the form $\sum_i f(m_i)/N$ for some arbitrary function f. The term corresponding to $i = 1$ in this sum simply reads, by obvious modification of (3–4),

$$\begin{aligned}\sum_{\sigma,\tau \in S_N} \varepsilon(\sigma\tau) \int_{\mathbb{R}} d\mu(m_1)\, \frac{f(m_1)}{N} p_{\sigma(1)-1}(m_1) p_{\tau(1)-1}(m_1) e^{-NV(m_1)} \\ \times \prod_{i=2}^{N} \int_{\mathbb{R}} d\mu(m_i)\, p_{\sigma(i)-1}(m_i) p_{\tau(i)-1}(m_i) e^{-NV(m_i)} \\ = N! \prod_{j=0}^{N-1} h_j \times \frac{1}{N} \int_{\mathbb{R}} dm\, f(m) \sum_{i=0}^{N-1} \frac{p_i(m)^2}{N h_i} e^{-NV(m)},\end{aligned}$$

which is independent of the choice of index i (which here equals 1). Therefore, the complete average with respect to the full matrix measure reads

$$\begin{aligned}\left\langle \frac{1}{N} \sum_{i=1}^{N} f(m_i) \right\rangle_V &= \frac{\int_{\mathbb{R}^N} dm_1 \ldots dm_N\, \frac{1}{N} \sum_{i=1}^{N} f(m_i) \Delta(m)^2 e^{-N \sum V(m_i)}}{\int_{\mathbb{R}^N} dm_1 \ldots dm_N\, \Delta(m)^2 e^{-N \sum V(m_i)}} \\ &= \int_{\mathbb{R}} dm\, f(m) \sum_{i=0}^{N-1} \frac{p_i(m)^2}{N h_i} e^{-NV(m)}, \quad (3\text{–}6)\end{aligned}$$

where we have added the subscript V to recall that the average is normalized with respect to the full measure (using V instead of V_0 in the denominator term). In the case where $f(m)$ is a polynomial, this can be immediately rewritten in terms of the r's and therefore the h's only, as we'll explain now.

To further compute the h's, we introduce operators Q and P, acting on the polynomials p_m:

$$Qp_m(x) = xp_m(x), \quad Pp_m(x) = \frac{d}{dx} p_m(x),$$

with the obvious commutation relation

$$[P, Q] = 1.$$

Using the self-adjointness of Q with respect to the scalar product $(f, g) = \int f(x)g(x)d\mu(x)$, it is easy to prove that

$$Qp_m(x) = xp_m(x) = p_{m+1}(x) + s_m p_m(x) + r_m p_{m-1}(x) \tag{3–7}$$

for some constants r_m and s_m, and that $s_m = 0$ if the potential $V(x)$ is even. The same reasoning yields

$$r_m = \frac{h_m}{h_{m-1}} \quad \text{for } m = 1, 2, \ldots,$$

and we also set $r_0 = h_0$ for convenience.

Moreover, expressing both (Pp_m, p_m) and (Pp_m, p_{m-1}) in two ways, using integration by parts, we easily get the master equations

$$\frac{m}{N} = \frac{(V'(Q)p_m, p_{m-1})}{(p_{m-1}, p_{m-1})}, \quad 0 = (V'(Q)p_m, p_m), \tag{3–8}$$

which amount to a recursive system for s_m and r_m. Note that the second line of (3–8) is automatically satisfied if V is even: it vanishes as the integral over $\mathbb{R}$ of an odd function. Indeed, this latter equation allows for computing the s's out of the r's, therefore becomes a tautology when all the s's vanish. Assuming for simplicity that V is even, the first equation of (3–8) gives a nonlinear recursion relation for the r's. The degree k of V actually determines the number of terms in the recursion, namely $k - 1$. So, we need to feed the $k - 2$ initial values of $r_0, r_1, r_2, \ldots, r_{k-3}$ into the recursion relation, and we obtain the exact value of $Z_N(V)$ by substituting $h_i = r_0 r_1 \ldots r_i$ in both the numerator and the denominator of (3–5). Note that for $V_0(x) = x^2/2$ the recursion (3–8) reduces simply to

$$\frac{m}{N} = \frac{(Qp_m^{(0)}, p_{m-1}^{(0)})}{(p_{m-1}^{(0)}, p_{m-1}^{(0)})} = r_m^{(0)},$$

and therefore $h_m^{(0)} = h_0^{(0)} m!/N^m = \sqrt{2\pi}\, m!/N^{m+1/2}$. The $p_m^{(0)}$ are simply the Hermite polynomials.

Finally, the free energy of the model (3–1) reads

$$F_N(V) = \operatorname{Log} Z_N(V) = N \operatorname{Log} r_0 \sqrt{\frac{N}{2\pi}} + \sum_{i=1}^{N-1} (N - i) \operatorname{Log} \frac{N r_i}{i} \tag{3–9}$$

in terms of the r's. In the case when V is even and $f(m)$ a polynomial, the average (3–6) is clearly expressible in terms of the r's only, by use of the recursion relation (3–7) (with $s_m = 0$). Indeed, when $f(m) = m^n$, we get

$$\left\langle \frac{1}{N} \operatorname{Tr}(M^n) \right\rangle_V = \frac{1}{N} \sum_{i=0}^{N-1} \frac{(Q^n p_i, p_i)}{h_i}. \tag{3–10}$$

When $n = 2$ for instance, this simply reads

$$\left\langle \frac{1}{N} \operatorname{Tr}(M^2) \right\rangle_V = \frac{r_1}{N} + \sum_{i=1}^{N-1} \frac{r_{i+1} + r_i}{N}. \tag{3–11}$$

3.3. Large N asymptotics I: Orthogonal Polynomials. As mentioned before, the large N limit of matrix integrals always has an interpretation as sum over planar (genus zero) fatgraphs. In particular, the free energy $F_N(V)$ of (3–9) can be expressed in an analogous way as (2–13) as a sum over connected fatgraphs, except that all but a finite number of g's vanish, namely $g_m = 0$ for $m \geq k+1$, and also for $m \leq 2$. The large N contribution is

$$F_N(V) \sim N^2 f_0(V) + O(1), \tag{3–12}$$

where $f_0(V)$ is the planar free energy, namely that obtained by restricting oneself to genus zero fatgraphs.

In view of the expression (3–9), it is straightforward to get asymptotics like (3–12), by first noting that as $h_0 \sim \sqrt{2\pi/N}$, the first term in (3–9) doesn't contribute to the leading order N^2 and then by approximating the sum by an integral of the form

$$f_0(V) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N-1} \left(1 - \frac{i}{N}\right) \operatorname{Log} \frac{r_i}{i/N} = \int_0^1 dz\,(1-z) \operatorname{Log} \frac{r(z)}{z}, \tag{3–13}$$

where we have assumed that the sequence r_i tends to a function $r_i \equiv r(i/N)$ of the variable $z = i/N$ when N becomes large. This can actually be proved rigorously. The limiting function $r(z)$ in (3–13) is determined by the equations (3–8), that become polynomial in this limit. In the case V even for instance, where $V(x) = x^2/2 + \sum_{i=2}^{p} g_{2i} x^{2i}/(2i)$ ($k = 2p$), we simply get

$$z = r(z) + \sum_{i=2}^{p} \binom{2i-1}{i} g_{2i} r(z)^i.$$

The function $r(z)$ is the unique root of this polynomial equation that tends to z for small z (it can be expressed using the Lagrange inversion method for instance, as a formal power series of the g's), and the free energy follows from (3–13). This allows also for computing the large N limit of averages of the form (3–6) or (3–10). For instance, the large N limit of (3–11) reads

$$\lim_{N\to\infty} \left\langle \frac{1}{N} \operatorname{Tr}(M^2) \right\rangle_V = 2 \int_0^1 dz\, r(z),$$

and more generally we have

$$\lim_{N\to\infty}\langle \frac{1}{N}\operatorname{Tr}(M^n)\rangle_V = \begin{cases} \binom{2p}{p}\int_0^1 dz\, r(z)^p & \text{if } n=2p,\\ 0 & \text{otherwise.}\end{cases} \tag{3–14}$$

In the case $V=V_0$ (and $r(z)=z$) of Gaussian averages, (3–14) reduces to the result 2–10.

3.4. Large N asymptotics II: Saddle-Point Approximation. We now present another solution, which does not rely on the orthogonal polynomial technique nor requires its applicability. We start from the N-dimensional integrals (3–2), that we rewrite

$$Z_N(V) = \frac{\int dm_1 \ldots dm_N e^{-N^2 S(m_1,\ldots,m_N)}}{\int dm_1 \ldots dm_N e^{-N^2 S_0(m_1,\ldots,m_N)}}, \tag{3–15}$$

where we have introduced the "actions"

$$S(m_1,\ldots,m_N) = \frac{1}{N}\sum_{i=1}^N V(m_i) - \frac{1}{N^2}\sum_{1\le i\ne j\le N}\operatorname{Log}|m_i-m_j|,$$

$$S_0(m_1,\ldots,m_N) = \frac{1}{N}\sum_{i=1}^N V_0(m_i) - \frac{1}{N^2}\sum_{1\le i\ne j\le N}\operatorname{Log}|m_i-m_j|.$$

For large N the numerator and denominator of (3–15) are dominated by the semiclassical (or saddle-point) minimum of S and S_0 respectively. For S, the saddle-point equations read

$$\frac{\partial S}{\partial m_j} = 0 \implies V'(m_j) = \frac{2}{N}\sum_{\substack{1\le i\le N\\ i\ne j}}\frac{1}{m_j-m_i} \tag{3–16}$$

for $j=1,2,\ldots,N$. Introducing the discrete resolvent

$$\omega_N(z) = \frac{1}{N}\sum_{i=1}^N \frac{1}{z-m_i}$$

and multiplying (3–16) by $1/(N(z-m_j))$ and summing over j, we easily get the equation

$$\begin{aligned} V'(z)\omega_N(z) &+ \frac{1}{N}\sum_{j=1}^N \frac{V'(m_j)-V'(z)}{z-m_j}\\ &= \frac{1}{N^2}\sum_{1\le i\ne j\le N}\frac{1}{m_j-m_i}\left(\frac{1}{z-m_j}-\frac{1}{z-m_i}\right)\\ &= \frac{1}{N^2}\sum_{1\le i\ne j\le N}\frac{1}{(z-m_i)(z-m_j)}\\ &= \omega_N(z)^2 + \frac{1}{N}\omega_N'(z). \end{aligned}$$

Assuming ω_N tends to a differentiable function $\omega(z)$ when $N \to \infty$ we may neglect the last derivative term, and we are left with the quadratic equation

$$\begin{aligned} &\omega(z)^2 - V'(z)\omega(z) + P(z) = 0, \\ &P(z) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \frac{V'(z) - V'(m_j)}{z - m_j}, \end{aligned} \tag{3–17}$$

where $P(z)$ is a polynomial of degree $k-2$. The existence of the limiting resolvent $\omega(z)$ boils down to that of the limiting density of distribution of eigenvalues

$$\rho(z) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \delta(z - m_j),$$

normalized by the condition

$$\int_{\mathbb{R}} \rho(z)\, dz = 1, \tag{3–18}$$

since there are exactly N eigenvalues on the real axis. This density is related to the resolvent through

$$\omega(z) = \int \frac{\rho(x)}{z - x}\, dx = \sum_{m=1}^{\infty} \frac{1}{z^m} \int_{\mathbb{R}} x^{m-1} \rho(x)\, dx \tag{3–19}$$

where the expansion holds in the large z limit, and the integral extends over the support of ρ, included in the real line. Conversely, the density is obtained from the resolvent by use of the discontinuity equation across its real support:

$$\rho(z) = \frac{1}{2i\pi} \lim_{\varepsilon\to 0} \omega(z + i\varepsilon) - \omega(z - i\varepsilon), \quad \text{for } z \in \operatorname{supp} \rho. \tag{3–20}$$

Solving the quadratic equation (3–17) as

$$\omega(z) = \frac{V'(z) - \sqrt{(V'(z))^2 - 4P(z)}}{2},$$

we must impose the large z behavior inherited from (3–18) and (3–19), namely that $\omega(z) \sim 1/z$ for large z. For $k \geq 2$, the polynomial in the square root has degree $2(k-1)$: expanding the square root for large z up to order $1/z$, all the terms cancel up to order 0 with $V'(z)$, and moreover the coefficient in front of $1/z$ must be 1 (this fixes the leading coefficient of P). The other coefficients of P are fixed by the higher moments of the measure $\rho(x)dx$. For instance, when $k = 2$ and $V = V_0$, we get

$$\omega_0(z) = \tfrac{1}{2}\big(z - \sqrt{z^2 - 4}\big).$$

It then follows from (3–20) that the density has the compact support $[-2, 2]$ and has the celebrated "Wigner's semicircle law" form

$$\rho_0(z) = \frac{1}{2\pi} \sqrt{4 - z^2}.$$

Viewing the resolvent ω_0 as the generating function for the moments of the measure whose density is ρ_0 (through the expansion (3–19)), we immediately get the values of the moments

$$\int_{\mathbb{R}} x^n \rho_0(x)dx = \begin{cases} c_p & \text{if } n = 2p, \\ 0 & \text{otherwise,} \end{cases} \tag{3–21}$$

with c_p as in (2–9). These are indeed immediately identified with the planar limit of the Gaussian Hermitian matrix averages (with potential $V_0(x) = x^2/2$) by using the following identity, valid for any V:

$$\lim_{N\to\infty} \langle \frac{1}{N} \operatorname{Tr} M^n \rangle_V = \int_{\mathbb{R}} x^n \rho(x)\, dx$$

The result (3–21) agrees with (2–10) and (3–14). Actually, comparing the preceding limit with (3–10), we deduce the following expression for the limiting density $\rho(x)$ in terms of the orthogonal polynomials of the previous solution:

$$\rho(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \frac{p_i(x)^2}{h_i} e^{-NV(x)}.$$

When $V = V_0$, we recover from this equation the Wigner's semicircle from standard asymptotics of the Hermite polynomials.

In the general case, the density reads

$$\rho(z) = \frac{1}{2\pi} \sqrt{4P(z) - (V'(z))^2}$$

and may have a disconnected support, made of a union of intervals. It is however interesting to restrict oneself to the case when the support of ρ is made of a single real interval $[a, b]$. It means that the polynomial $V'(z)^2 - 4P(z)$ has single roots at $z = a$ and $z = b$ and that all other roots have even multiplicities. In other words, we may write

$$\omega(z) = \frac{1}{2}(V'(z) - Q(z)\sqrt{(z-a)(z-b)}),$$

where $Q(z)$ is a polynomial of degree $k - 2$, entirely fixed in terms of V by the asymptotics $\omega(z) \sim 1/z$ for large $|z|$. For instance, for an even quartic potential, we have

$$\begin{aligned} V(z) &= \frac{z^2}{2} - g\frac{z^4}{4} \\ \omega(z) &= \frac{1}{2}\Big(z - gz^3 - \Big(1 - g\frac{a^2}{2} - gz^2\Big)\sqrt{z^2 - a^2}\Big), \\ a^2 &= \frac{2}{3g}\big(1 - \sqrt{1 - 12g}\,\big), \\ \rho(z) &= \frac{1}{2\pi}(1 - \tfrac{1}{2}ga^2 - gz^2)\sqrt{a^2 - z^2}. \end{aligned} \tag{3–22}$$

The planar free energy (3–12) is finally obtained by substituting the limiting densities ρ, ρ_0 in the saddle point actions, namely

$$f_0(V) = S(\rho_0, V_0) - S(\rho, V),$$
$$S(\rho, V) = \int dx\, \rho(x) V(x) - \int dx\, dy\, \rho(x)\rho(y)\, \mathrm{Log}\,|x-y|.$$

This expression seems more involved than our previous result (3–13), but is equivalent to it. In the case of the quartic potential of (3–22), we find the genus zero free energy $f_0(g) \equiv f_0(V)$

$$f_0(g) = \frac{1}{2}\,\mathrm{Log}\,\frac{a^2}{4} + \frac{1}{384}(a^2-4)(a^2-36), \tag{3–23}$$

with a^2 as in (3–22).

In cases where the orthogonal polynomial technique does not apply (like in complicated multimatrix integrals), however, the saddle-point technique always gives access to the planar limit.

For completeness we quickly mention the corresponding results for the gravitational $O(n)$ model introduced in Section 2.3 (see [22] for details and a more general solution). We have to evaluate the large N asymptotics of the partition function

$$Z = \frac{\int dA_1 \dots dA_n dB e^{-N\,\mathrm{Tr}(W(B,A_1,A_2,\dots,A_n))}}{\int dA_1 \dots dA_n dB e^{-N\,\mathrm{Tr}(W_0(B,A_1,A_2,\dots,A_n))}},$$
$$W(B,A_1,A_2,\dots,A_n) = V(B) + \frac{1}{2}\sum_{c=1}^{n} A_c^2 - gB\sum_{c=1}^{n} A_c^2, \tag{3–24}$$
$$W_0(B,A_1,A_2,\dots,A_n) = \frac{1}{2}B^2 + \frac{1}{2}\sum_{c=1}^{n} A_c^2,$$

a simple generalization of (2–19) with some arbitrary potential $V(B)$. Note that W is simply quadratic in the A_c, so we can perform the Gaussian integrals overs the A's first. More precisely, the potential takes the form

$$\mathrm{Tr}\, W(B,A_1,A_2,\dots,A_n) = \mathrm{Tr}\, V(B) + \frac{1}{2}\sum_{c=1}^{n}\sum_{ijkl=1}^{N} (A_c)_{ik}\mathbf{Q}_{ik;jl}(A_c)_{lj},$$
$$\mathbf{Q}_{ik;jl} = \delta_{ij}\delta_{kl} - g(\delta_{ij}B_{kl} + \delta_{kl}B_{ij}).$$

More compactly, the quadratic form reads

$$\mathbf{Q} = I \otimes I - g(I \otimes B + B \otimes I).$$

The Gaussian integral over the A's, normalized as in (3–24), gives $(\det \mathbf{Q})^{-1/2}$ for each integral; hence

$$Z = \frac{1}{\int dB\, e^{-N\,\mathrm{Tr}\,\frac{B^2}{2}}} \int dB\, e^{-N\,\mathrm{Tr}\,V(B)} \det(I\otimes I - g(I\otimes B + B\otimes I))^{-n/2}.$$

Thus the $O(n)$ model reduces to a one-matrix model, but with a complicated potential. We may now apply both the reduction to an eigenvalue integral and the large N technique sketched above. We get the eigenvalue integral

$$Z = \frac{1}{(2\pi)^{N/2}} \int db_1 \dots db_N \, e^{-N^2 S(b_1,\dots,b_N)},$$

$$S(b_1,\dots,b_N) = \frac{1}{N}\sum_{i=1}^{N} V(b_i) - \frac{1}{N^2}\sum_{1\le i\neq j\le N} \operatorname{Log}|b_i - b_j| + \frac{n}{2N^2}\sum_{1\le i,j\le N} \operatorname{Log}(1-g(b_i+b_j)). \tag{3–25}$$

In the last term we may impose the constraint $i \neq j$, as the terms $i = j$ only contribute for $O(\frac{1}{N})$ to S. The corresponding saddle-point equations $\partial_{b_i} S = 0$ read

$$V'(b_i) = \frac{2}{N}\sum_{j\neq i}\frac{1}{b_i - b_j} + \frac{gn}{N}\sum_{j\neq i}\frac{1}{1-g(b_i+b_j)}$$

for $i = 1, 2, \dots, N$. Setting $x_i = 1 - 2gb_i$, and $v'(x) = \frac{1}{2g}V'(b)$, we get

$$v'(x_i) = \frac{2}{N}\sum_{j\neq i}\frac{1}{x_i - x_j} + \frac{n}{N}\sum_{j\neq i}\frac{1}{x_i + x_j}. \tag{3–26}$$

Multiplying this by $\frac{1}{N}(\frac{1}{z-x_i} - \frac{1}{z+x_i})$ and summing over i, we arrive at the large N quadratic equation:

$$\begin{aligned} -Q(z) + v'(z)\omega(z) + v'(-z)\omega(-z) &= \omega(z)^2 + \omega(-z)^2 + n\omega(z)\omega(-z), \\ \omega(z) &= \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{z-x_i} \\ Q(z) &= \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\frac{v'(z)-v'(x_i)}{z-x_i} - \frac{v'(-z)-v'(x_i)}{z+x_i}, \end{aligned} \tag{3–27}$$

while the saddle-point equation turns into a discontinuity equation across the support of the limiting eigenvalue distribution ρ:

$$\omega(z+i0) + \omega(z-i0) + n\omega(-z) = v'(z) \quad \text{for } z \in \operatorname{supp}\rho. \tag{3–28}$$

For generic n, $\omega(z)$ is well defined only in an infinite covering of the complex plane, the transition from one sheet to another being governed by (3–28). However, setting

$$n = 2\cos(\pi\nu),$$

we easily see that if $\nu = r/s$ is rational, this number becomes finite. Indeed, when expressed in terms of the shifted resolvent

$$w(z) = \frac{nv'(-z) - 2v'(z)}{n^2-4} + \omega(z), \tag{3–29}$$

the discontinuity and quadratic equations (3–28) and (3–27) become respectively

$$\begin{aligned} w(z+i0)+w(z-i0)+nw(-z)&=0,\\ w(z)^2+w(-z)^2+nw(z)w(-z)&=P(z), \end{aligned} \tag{3–30}$$

where

$$P(z)=\frac{1}{n^2-4}(nv'(z)v'(-z)-v'(z)^2-v'(-z)^2)-Q(z).$$

Introducing the two functions

$$w_\pm(z)=e^{\pm i\pi\frac{\nu}{2}}w(z)+e^{\mp i\pi\frac{\nu}{2}}w(-z) \tag{3–31}$$

namely with $w_-(z)=w_+(-z)$, we immediately get from (3–30) that

$$\begin{aligned} w_+(z+i0)&=-e^{i\pi\nu}w_-(z-i0),\\ w_-(z+i0)&=-e^{-i\pi\nu}w_+(z-i0),\\ w_+(z)w_-(z)&=P(z). \end{aligned} \tag{3–32}$$

For rational $\nu=r/s$, we see that the combination

$$w_+(z)^s+(-1)^{r+s}w_-(z)^s=2S(z) \tag{3–33}$$

is regular across the support of ρ. For instance, if $v(x)$ is polynomial or meromorphic, so are $P(z)$ and $S(z)$. We may now solve (3–32) and (3–33) for $w_\pm$ as

$$w_\pm(z)=(S(z)\pm\sqrt{S(z)^2-P(z)^s})^{\frac{1}{s}},$$

and finally get the resolvent using (3–31) and (3–29):

$$\begin{aligned} \omega(z)=\frac{2v'(z)-nv'(-z)}{n^2-4}-\frac{1}{2i\sin(\pi r/s)}\Big(&e^{i\pi r/(2s)}\big(S(z)+\sqrt{S(z)^2-P(z)^s}\,\big)^{1/s}\\ &-e^{-i\pi r/(2s)}\big(S(z)-\sqrt{S(z)^2-P(z)^s}\,\big)^{1/s}\Big). \end{aligned} \tag{3–34}$$

Assuming that v is polynomial, we must further fix the coefficients of P and S by imposing the asymptotic behavior $\omega(z)\sim 1/z$ for large $|z|$. If $\deg(v)=k$, then $\deg(P)=2(k-1)$ and $\deg(S)=s(k-1)$, and P and S are entirely fixed if we impose moreover that ρ has a support made of a single interval $[a,b]$.

3.5. Critical Behavior and Asymptotic Enumeration. We saw in Section 2 how to write the generating functions for various families of (decorated) fatgraphs in terms of Hermitian matrix integrals. It is a standard fact that the critical properties of these generating functions can be translated into asymptotics for these numbers of fatgraphs, say for large numbers of vertices. Indeed, writing such a generating function as $F(g)=\sum_{n\ge 0}F_ng^n$, where F_n denotes a number of connected (decorated) fatgraphs with n vertices, assume F has a critical singular part of the form

$$F(g)_{\text{sing}}\sim(g_c-g)^{2-\gamma} \tag{3–35}$$

when g approaches some critical value g_c, then for large n the numbers F_n behave as

$$F_n \sim g_c^{-n} n^{-3+\gamma}.$$

We apply this to the example of (3–1) with the quartic potential $V(x) = \frac{1}{2}x^2 - \frac{1}{4}gx^4$. The generating function

$$F_N(g) = \operatorname{Log} Z_N(V) = \sum_{h \geq 0} N^{2-2h} f_h(g)$$

decomposes into the generating functions $h_h(g)$:

$$f_h(g) = \sum_{n \geq 0} g^n f_{h,n},$$

where $f_{h,n}$ denotes the number of fatgraphs of genus h with n vertices. Using the genus zero answer for $f_0(g)$ (3–23), we find that the critical singularity is attained when $g = g_c = \frac{1}{12}$, in which case $f_0(g)_{\text{sing}} \sim (g_c - g)^{5/2}$ and $\gamma = -1/2$ in (3–35). This implies the following asymptotics for the numbers of connected genus zero fatgraphs

$$f_{0,n} \sim \frac{12^n}{n^{7/2}}.$$

More refined applications of the orthogonal polynomial techniques lead to the genus h result

$$f_{h,n} \sim \frac{12^n}{n^{1+(5/2)(1-h)}}$$

and to a score of interesting behaviors for the numbers of decorated fatgraphs for specific potentials V, as well as for the $O(n)$ model with also specific potentials, for which any positive rational value of $-\gamma$ in (3–35) may be reached. The exponent γ is called the string susceptibility exponent. More generally, it has been shown that the coupling of a critical matter theory to gravity results in a behavior (3–35) of the genus zero free energy with γ given by the formula [4]

$$\gamma = \frac{c - 1 - \sqrt{(25-c)(1-c)}}{12}, \tag{3–36}$$

where c is the central charge of the conformal field theory underlying the flat space critical matter model [23]. The exponent $\gamma = -\frac{1}{2}$ is characteristic of the "pure gravity" models, in which there is no matter theory, in other words $c = 0$. In the case of the so-called dense critical phase of the $O(n)$ model, one has [24]

$$c(n) = 1 - 6\frac{e^2}{1-e}, \quad n = 2\cos(\pi e), \tag{3–37}$$

and therefore

$$\gamma(n) = -\frac{e}{1-e}$$

from (3–36). This is confirmed by the saddle-point results [22] mentioned above.

3.6. Gaussian Words. In the case of multimatrix integrals, one may wonder how the beautifully simple result represented by (2–10) and (3–21) is generalized. More precisely, we are interested in the large N limit of the multi-Gaussian average of the trace of any word:

$$\begin{aligned}\eta_{n_1,n_2,\ldots,n_{mk}} &= \big\langle\langle \mathrm{Tr}(M_1^{n_1}\ldots M_k^{n_k} M_1^{n_{k+1}}\ldots M_k^{n_{mk}})\rangle\big\rangle \\ &\equiv \lim_{N\to\infty} \frac{\int dM_1\ldots dM_k e^{-N\,\mathrm{Tr}\sum_{i=1}^k M_i^2}\,\mathrm{Tr}(M_1^{n_1}\ldots M_k^{n_k} M_1^{n_{k+1}}\ldots M_k^{n_{mk}})}{N\int dM_1\ldots dM_k e^{-N\,\mathrm{Tr}\sum_{i=1}^k M_i^2}}. \qquad (3\text{–}38)\end{aligned}$$

In this setting, the result (2–10) corresponds to $k = m = 1$ and reads

$$\eta_n = \begin{cases} c_p & \text{if } n = 2p, \\ 0 & \text{otherwise,} \end{cases}$$

for all $n \geq 0$. The simplest way of computing (3–38) in a given situation is the blind application of Wick's theorem (2–17), by further restricting the Wick pairings to the *planar* ones only, as $N \to \infty$. For instance, for $k = 2$ and $m = 2$, we have

$$\begin{aligned}\eta_{n_1,n_2,n_3,n_4} &= \big\langle\langle \mathrm{Tr}(M_1^{n_1} M_2^{n_2} M_1^{n_3} M_2^{n_4})\rangle\big\rangle \\ &= \eta_{n_1+n_3}\eta_{n_2}\eta_{n_4} + \eta_{n_2+n_4}\eta_{n_1}\eta_{n_3} - \eta_{n_1}\eta_{n_2}\eta_{n_3}\eta_{n_4},\end{aligned}$$

where we have isolated respectively the pairings in which elements of the blocks $M_1^{n_1}$ and $M_1^{n_3}$ may be connected, then those in which elements of the blocks $M_2^{n_2}$ and $M_2^{n_4}$ may be connected, and we have subtracted the terms in which all pairings take place within the same powers of the M's to avoid counting them twice (once in each of the previous terms).

To compute the most general trace of a word in say k Gaussian Hermitian matrices, we simply have to apply the following quadratic recursion relation, for $\omega = e^{2i\pi/k}$

$$\eta_{n_1,\ldots,n_{mk}} = -\sum_{i=1}^{mk-1} \omega^i \eta_{n_1,\ldots,n_i}\eta_{n_{i+1},\ldots,n_{mk}}. \qquad (3\text{–}39)$$

This relation is a compact rephrasing of the Wick theorem in the case of planar pairings, as the reader will check easily (the only relation to be used is

$$\sum_{0\leq i\leq k-1} \omega^i = 0;$$

see [10] for details). This gives a priori access to the average of any trace of word. The numbers η are natural multidimensional generalizations of the Catalan numbers (2–9).

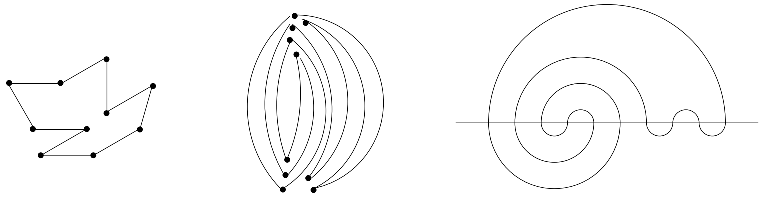


Figure 4. A typical polymer with $2n = 10$ segments is depicted, together with one of its compact folding configurations, in which the segments have been deformed and slightly pulled apart for clarity. The corresponding meander is obtained by drawing an horizontal line (river) intersecting the folding configuration (road) through $2n = 10$ points (bridges). The connected half segments of polymer have been replaced by semicircular portions of road for simplicity.

4. Folding Polymers: Meanders

We now present a first important application of matrix integrals to the fundamental combinatorial problem of meander enumeration, also equivalent to the compact folding problem of polymers. After defining meanders and reviewing various approaches to their enumeration, we will define in Section 6 a matrix model for meanders and discuss its solution in some particular cases. We finally present a general argument leading to the exact values of the meandric asymptotic configuration exponents.

4.1. Definitions and Generalities. A polymer is a chain of say m identical constituents, modelled by rigid segments attached by their ends, that serve as hinges in the folding process. A useful representation to bear in mind is that of a strip of stamps of size $1 \times m$, attached by their ends. Given a closed polymer chain with $m = 2n$ segments (one in which the chain is closed so as to form a loop; see Figure 4), we address the question of enumeration of all the ways of compactly folding the chain, in such a way that in the final (folded) state all the segments are piled up onto one another.

As illustrated in Figure 4, we may keep track of the folding by intersecting the final state with an oriented line (say east to west), and slightly pulling apart the segments of the folded polymer. Redrawing the result with a horizontal east-west oriented line, and representing any two connected half-segments by semicircles, we arrive at a meander, i.e., the configuration of a nonintersecting loop (road, made of semicircular bits) crossing the horizontal line (river, flowing east to west) through a given number $m = 2n$ of points (bridges), as depicted in Figure 4. The number of bridges will also be called the order of the meander. The number of distinct meanders of order $2n$ is denoted by M_{2n} [8].

We extend this definition to a set of k roads namely meanders with k possibly interlocking connected components and a total number of bridges $2n$ (corresponding to the simultaneous folding of k possibly interlocking polymers with a total number of $2n$ segments). The number of meanders of order $2n$ with k connected components is denoted by $M_{2n}^{(k)}$. Note that necessarily $1 \le k \le n$. These numbers are summarized in the meander polynomial [10]

$$m_{2n}(q) = \sum_{k=1}^{n} M_{2n}^{(k)}\ q^k. \tag{4–1}$$

Given a meander of order $2n$, the river cuts the plane into two parts. The upper part is called an arch configuration a of order $2n$: it is a set of n nonintersecting semicircles drawn in the upper half plane delimited by the river and joining the $2n$ equally spaced bridges by pairs. The lower part is clearly the reflection b^t of another arch configuration b of order $2n$. Let A_{2n} denote the set of arch configurations of order $2n$, and for any $a, b \in A_{2n}$, let $c(a,b)$ denote the number of connected components of road obtained by forming a (multicomponent) meander with a as upper half and b^t as lower half. Then

$$m_{2n}(q) = \sum_{a,b \in A_{2n}} q^{c(a,b)}.$$

The number of distinct arch configurations is nothing but the Catalan number

$$|A_{2n}| = c_n = \frac{(2n)!}{(n+1)!n!},$$

since A_{2n} can be easily mapped onto the planar Wick pairings of a single star with $2n$ branches leading to (2–10): indeed, we just have to pull all the bridges together so as to form a $2n$-legged vertex; this results in a planar petal diagram like that of Figure 2(a). As an immediate consequence we get $m_{2n}(1) = |A_{2n}|^2 = c_n^2 \sim 4^{2n}/(\pi n^3)$ by use of Stirling's formula for large n. In general, the meander polynomial is expected to behave for large n as

$$m_{2n}(q) \sim c(q) \frac{R(q)^{2n}}{n^{\alpha(q)}} \tag{4–2}$$

and we just proved that $R(1) = 4$, $\alpha(1) = 3$ and $c(1) = 1/\pi$. In Section 6.4 below, we will present an argument leading to the general determination of the exact value of the meander configuration exponent $\alpha(q)$ as a function of q. Note that

$$s(q) = \lim_{n \to \infty} \frac{1}{2n} \operatorname{Log} m_{2n}(q) = \operatorname{Log} R(q)$$

is nothing but the thermodynamic entropy of folding per segment of a multicomponent closed polymer.

We may also address the problem of compactly folding an open polymer chain of $n-1$ segments [10], attached to a wall by one of its ends (think of a strip of stamps, attached in a book by its left end). As illustrated in Figure 5, we may intersect the folded polymer with a circle that also crosses the book's end. Extending the polymer itself into a half-line, and deforming the circle accordingly

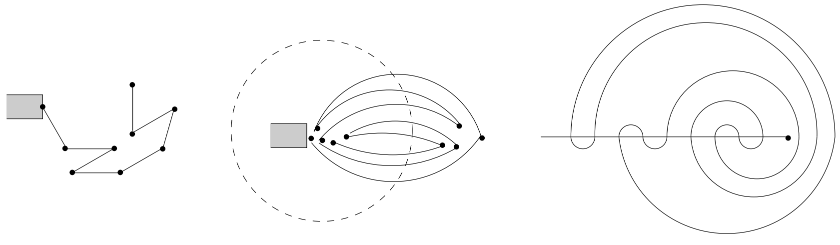


Figure 5. An open polymer with 8 segments, together with one typical compactly folded configuration, and its semimeander version of order 9, obtained by intersecting the 8 segments of the configuration with a circle (represented in dashed line), that also intersects the support (wall) to which the open polymer is attached. The free end of the folded polymer is then stretched into a straight half-line, that becomes the semimeander's river, the free end becoming its source. The dashed circle is deformed into a winding road, crossing the river through 9 bridges. For simplicity, we have represented all the pieces of the road joining the various segments by means of semicircles. Note that here the road winds three times around the source of the river.

without creating new intersections, we arrive at a semimeander of order n, i.e., the configuration of a closed non-self-intersecting loop (road, the former circle) crossing a semi-infinite line (river with a source, the former polymer+book) through n points (bridges).

Note that, in a semimeander, the road may wind around the source of the river, as illustrated in Figure 5. We denote by $\bar{M}_n$ the number of topologically inequivalent semimeanders of order n, and by $\bar{M}_n^{(k)}$ the number of semimeanders with k connected components, $1 \leq k \leq n$. We also have the semimeander polynomial

$$\bar{m}_n(q) = \sum_{k=1}^{n} \bar{M}_n^{(k)}\ q^k.$$

A semimeander may be viewed as a particular kind of meander by opening the river as sketched in Figure 6 so as to double the number of bridges $n \to 2n$, and by connecting them as they were in the semimeander in the upper-half of the meander, and through the reflection of a rainbow arch configuration r_{2n}^t in the lower one, made of n concentric semicircles. The semimeander polynomial is easily rewritten as

$$\bar{m}_n(q) = \sum_{a \in A_{2n}} q^{c(a, r_{2n})} \tag{4–3}$$

and we have the value at $q = 1$: $\bar{m}_{2n}(1) = c_n \sim 4^n/(\sqrt{\pi} n^{3/2})$. We also expect the asymptotics

$$\bar{m}_{2n}(q) \sim \bar{c}(q) \frac{\bar{R}(q)^n}{n^{\bar{\alpha}(q)}} \tag{4–4}$$

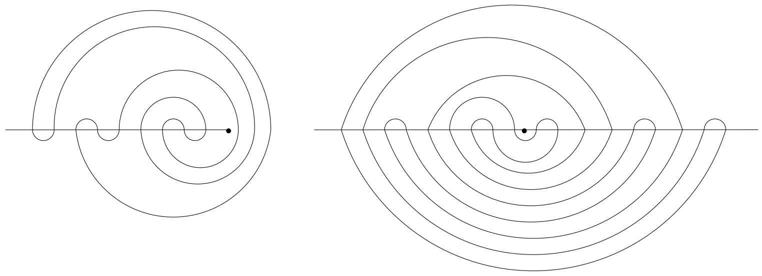


Figure 6. A semimeander of order n is opened into a meander of order $2n$. Think of the lower part of the river as pivoting around its source to form a straight line, while all bridge connections are deformed without any new intersections. The lower part of the meander is a rainbow, made of n concentric semicircles. The winding number is just the number of semicircular arches passing at the vertical of the center (former source).

and we have $\bar{R}(1) = 4$, $\bar{\alpha}(1) = \frac{3}{2}$ and $\bar{c}(q) = 1/\sqrt{\pi}$. Note again that the semimeander configuration exponent $\bar{\alpha}(q)$ will be determined as a function of q in Section 6.4 below.

Conversely, a meander may be viewed as a semimeander with no winding. It is therefore sufficient to solve the semimeander enumeration problem, provided we keep track of the winding numbers.

4.2. A Simple Algorithm. Numerical Results. Multi-component semimeanders of order n are in one-to-one correspondence with arch configurations of order $2n$ (4–3). Our algorithm [11] is based on a recursive construction of arch configurations that allows to keep track of both the numbers of connected components and windings. To build an arch configuration of order $2n+2$ from one of order $2n$, we may do either of the following transformations, both involving the addition of two bridges along the river, respectively to the left and right of the previous ones:

(I) For each external arch (semicircle contained in no other) connecting say the bridges i and j, replace it by two semicircles, one connecting the new leftmost bridge to i, and one connecting j to the new rightmost bridge.

(II) Add a large external arch connecting the two added bridges: it circles the whole previous arch configuration.

The corresponding semimeanders are obtained by completing this arch with the reflection of the rainbow r_{2n+2} as lower part. It is easy to show that applying (I)–(II) to all of A_{2n} yields exactly A_{2n+2}. Moreover, (I) preserves the number of connected components, while (II) obviously increases it by 1 (the net result is to add a circle around the semimeander). Applying successions of $(I) - (II)$ on the "root" (semimeander of order 1), we may build the tree of semimeanders of Figure 7.

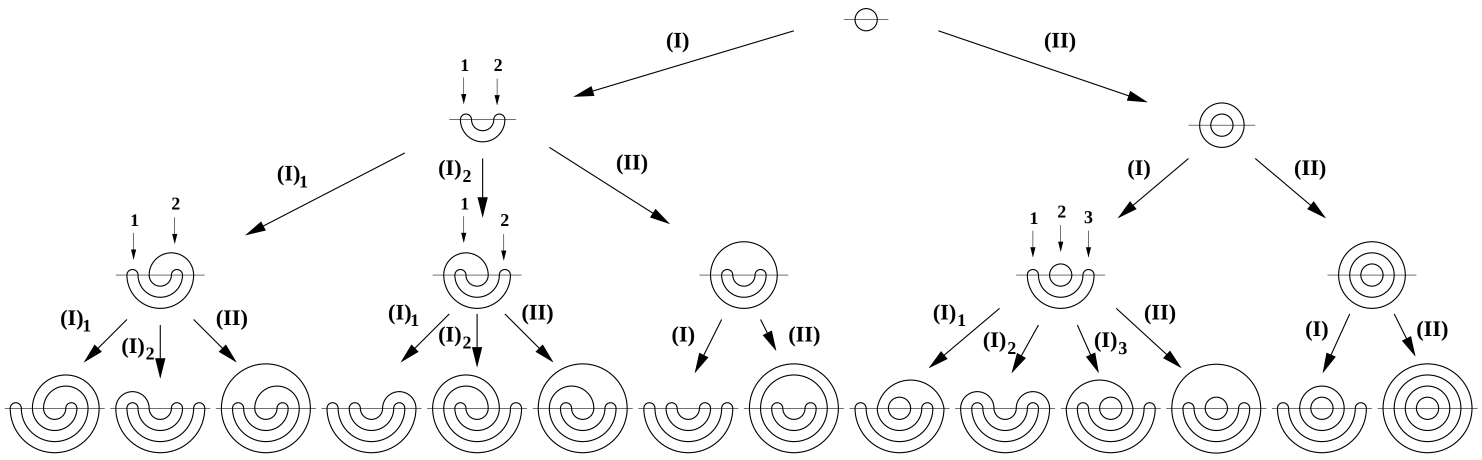


Figure 7. The tree of semimeanders down to order $n = 4$. This tree is constructed by repeated applications of the transformations (I) and (II) on the semimeander of order 1 (root). We have indicated by small vertical arrows the multiple choices for the process (I), each of which is indexed by its number. The number of connected components of a given semimeander is equal to the number of processes (II) in the path going from the root to it, plus one (that of the root).

n	$\bar{M}_n$	n	$\bar{M}_n$	k	$\bar{M}_{27}^{(k)}$	k	$\bar{M}_{27}^{(k)}$
1	1	16	1053874	1	369192702554	16	2376167414
2	1	17	3328188	2	2266436498400	17	628492938
3	2	18	10274466	3	6454265995454	18	153966062
4	4	19	32786630	4	11409453277272	19	34735627
5	10	20	102511418	5	14161346139866	20	7159268
6	24	21	329903058	6	13266154255196	21	1333214
7	66	22	1042277722	7	9870806627980	22	220892
8	174	23	3377919260	8	6074897248976	23	31851
9	504	24	10765024432	9	3199508682588	24	3866
10	1406	25	35095839848	10	1483533803900	25	374
11	4210	26	112670468128	11	619231827340	26	26
12	12198	27	369192702554	12	236416286832	27	1
13	37378	28	1192724674590	13	83407238044		
14	111278	29	3925446804750	14	27346198448		
15	346846			15	8352021621		

Table 1. The number $\bar{M}_n^{(k)}$ of semimeanders of order n with k connected components, obtained by exact enumeration on the computer. On the left, the number of one-component semimeander ($k = 1$) is given for $n \le 29$; on the right, n is fixed to 27 and $1 \le k \le n$.

The preceding algorithm is easily implemented on a computer and yields all semimeander numbers with fixed winding and connected components up to some quite large orders. We give examples in Table 1.

These numbers allow in turn for probing the asymptotics (4–2) and (4–4), as functions of q. In particular, we have these large q asymptotics of the radii:

$$
\begin{aligned}
R(q) &= 2\sqrt{q}\Big(1 + \frac{1}{q} + \frac{3}{2q^2} - \frac{3}{2q^3} - \frac{29}{8q^4} - \frac{81}{8q^5} - \frac{89}{16q^6} + O\Big(\frac{1}{q^7}\Big)\Big),\\
\bar{R}(q) &= q + 1 + \frac{2}{q} + \frac{2}{q^2} + \frac{2}{q^3} - \frac{4}{q^5} - \frac{8}{q^6} - \frac{12}{q^7} - \frac{10}{q^8} - \frac{4}{q^9} + \frac{12}{q^{10}} + \frac{46}{q^{11}}\\
&\qquad + \frac{98}{q^{12}} + \frac{154}{q^{13}} + \frac{124}{q^{14}} + \frac{10}{q^{15}} - \frac{102}{q^{16}} + \frac{20}{q^{17}} - \frac{64}{q^{18}} + O\Big(\frac{1}{q^{19}}\Big).
\end{aligned}
$$

At finite values of q, the numerical results displayed in Figure 8 reveal an interesting phase transition between a phase for $q < q_c$ of irrelevant winding: the numbers of meanders and semimeanders are then asymptotically equivalent $R(q) = \bar{R}(q))$, and a strong winding phase $q > q_c$, where the winding number is proportional to n, and therefore $\bar{R}(q) \gg R(q)$. The transition point is estimated as $q_c \simeq 2$ with poor precision. We will propose an exact value for q_c in Section 6.4 below.

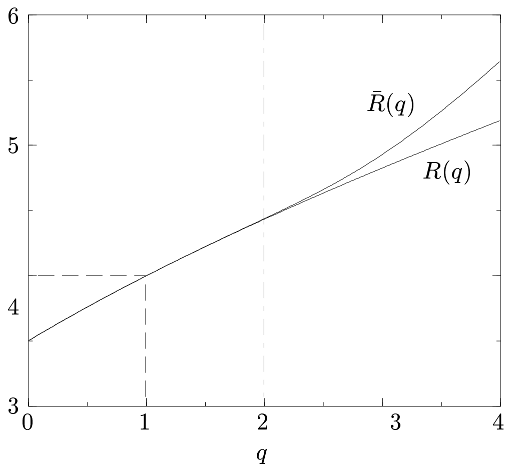


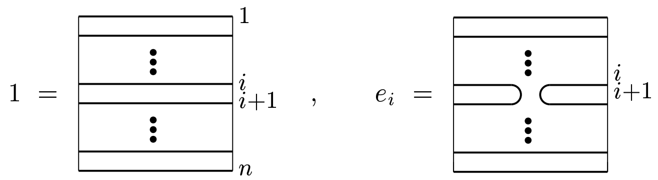
Figure 8. The functions $\bar{R}(q)$ and $R(q)$ for $0 \le q \le 4$ as results of large n extrapolations. The two curves coincide for $0 \le q \le 2$ and split for $q > 2$ with $\bar{R}(q) > R(q)$. Apart from the exact value $\bar{R}(1) = R(1) = 4$, we find the estimates $\bar{R}(0) = 3.50(1)$, $\bar{R}(2) = 4.44(1)$, $\bar{R}(3) = 4.93(1)$ and $\bar{R}(4) = 5.65(1)$.

5. Algebraic Formulation: Temperley–Lieb Algebra

5.1. Definition. The Temperley–Lieb algebra of order n and parameter q, denoted by $TL_n(q)$, is defined through its n generators $1, e_1, e_2, \ldots, e_{n-1}$ subject to the relations

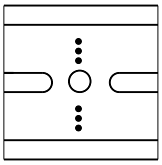
$$\begin{array}{lll} \text{(i)} & e_i^2 = q e_i & \text{for } i = 1, 2, \ldots, n-1; \\ \text{(ii)} & [e_i, e_j] = 0 & \text{if } |i-j| > 1; \\ \text{(iii)} & e_i e_{i\pm 1} e_i = e_i & \text{for } i = 1, 2, \ldots, n-1. \end{array} \tag{5–1}$$

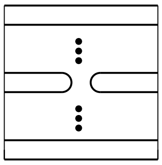
This definition becomes clear in the "domino" pictorial representation, where the generators are represented as dominos as follows:



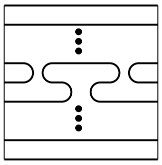
and a product of elements is represented by the concatenation of the corresponding dominos. Note that we have numbered the left and right "string ends" of the dominos from top to bottom, 1 to n. Relation (ii) expresses the locality of the e's, namely that the e's commute whenever they involve distant strings.

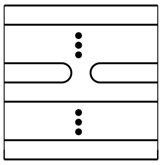
Relations (i) and (iii) read respectively

(i) $e_i^2 =$ 

$= qe_i = q$ 

(the loop has been erased, but affected the weight q) and

(iii) $e_i e_{i+1} e_i =$ 

$= e_i =$ 

(obtained by stretching the $(i{+}2)$-nd string).

Moreover the algebra is equipped with a trace, defined as follows. Given a domino d, we put it on a cylinder by identifying its left and right string ends. We then count the number $n(d)$ of connected strings in the resulting picture. The trace is then simply $\mathrm{Tr}(d) = q^{n(d)}$. This definition is extended to any element of the algebra by linearity. This trace has the important Markov property which allows to compute it by induction:

$$\mathrm{Tr}(d(e_1, e_2, \dots, e_j)e_{j+1}) = \frac{1}{q}\,\mathrm{Tr}(d(e_1, e_2, \dots, e_j)) \tag{5–2}$$

for any algebra element $d(e_1, \dots, e_j)$ involving only the generators $e_1, e_2, \dots e_j$.

The crucial remark actually concerns the left ideal $I_{2n}(q)$ of $TL_{2n}(q)$ generated by the element $e_1 e_3 \dots e_{2n-1}$. The corresponding dominos have their right string ends paired by single arches linking the ends $2i-1$ and $2i$, $i = 1, 2, \dots, n$. Therefore their $2n$ left string ends are also connected among themselves, thus forming an arch configuration of order $2n$ after eliminating the internal loops and stretching the strings (by use of (i)–(iii)) and giving them a semicircular shape. We therefore have a bijection between the set A_{2n} of arch configurations of order $2n$ and a basis of the ideal $I_{2n}(q) = TL_{2n}(q)e_1e_3 \dots e_{2n-1}$, made only of "reduced" dominos, namely with all their loops removed and their strings stretched. The set of reduced dominos of $I_{2n}(q)$ is denoted by D_{2n}. As a consequence we have $\dim(I_{2n}(q)) = c_n$. We denote by the same letter $a \in A_{2n}$ or D_{2n} the arch configuration or the corresponding reduced domino.

5.2. Meander Polynomials. Given a pair a, b of reduced dominos of $I_{2n}(q)$, we may form the scalar product

$$(a, b) = \frac{1}{q^n}\,\mathrm{Tr}(b^t a),$$

where by b^t we mean the reflected domino with respect to one of its vertical edges. The middle of the concatenated domino $b^t a$ is nothing but the meander with a as upper-half and b^t as lower one (tilted by 90°), while it also contains n small circles formed in the cylinder identification; hence

$$(a, b) = q^{c(a,b)}.$$

We deduce the following expressions for the meander and semimeander polynomials in purely algebraic terms [12]:

$$m_{2n}(q) = \sum_{a,b \in D_{2n}} (a, b), \quad \bar{m}_n(q) = \sum_{a \in D_{2n}} (a, r_{2n}).$$

Since the trace can be computed by induction, using (5–2), we now get an inductive way of computing the meander and semimeander polynomials. This is however far from explicit, and does not allow a priori for a good asymptotic study of these polynomials.

5.3. Meander Determinants. On the other hand, we may consider other meander-related quantities that can be explicitly calculated using Temperley–Lieb algebra representation theory. the most interesting of them is the "meander determinant" constructed as follows. We first define the Gram matrix for the reduced basis of $I_{2n}(q)$ as the $c_n \times c_n$ matrix $G_{2n}(q)$ with entries

$$G_{2n}(q)_{a,b} = (a, b) \quad \text{for all } a, b \in D_{2n}.$$

The meander determinant is then defined as the Gram determinant

$$\Delta_{2n}(q) = \det(G_{2n}(q)).$$

This determinant is computed by performing the explicit Gram–Schmidt orthonormalization of the scalar product $(\cdot\,, \cdot)$, and using along the way the representation theory of $TL_n(q)$. The result is remarkably simple [12; 13]:

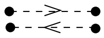
$$\Delta_{2n}(q) = \prod_{m=1}^{n} U_m(q)^{a_{m,n}}, \quad a_{m,n} = \binom{2n}{n-m} - 2\binom{2n}{n-m-1} + \binom{2n}{n-m-2},$$

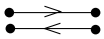
where the $U_m(q)$ are the Chebyshev polynomials of the second kind defined by $U_m(q) = \sin((m+1)\theta)/\sin(\theta)$ with $2\cos\theta = q$.

6. Matrix Model for Meanders

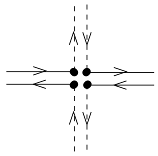
We have seen in Section 2 how to tailor matrix integrals to suit our combinatorial needs. The meanders are basically obtained by intersection of two types of curves: the roads and the rivers. In the following, we will define a matrix model that generates arbitrary configurations of any number of rivers (now viewed as closed curves, so in the case of just one river we should close it into a circle), crossed by any number of non-(self)-intersecting roads [10].

6.1. The Black-And-White Model. We now construct a Hermitian matrix integral that generates the meander polynomials. The computation of such an integral must involve fatgraphs with double-line edges, which we will eventually interpret as the river(s) and the road(s). Let's paint in white the river edges, and in black the road edges. We therefore have a "black and white" graph made of black and white loops which intersect each other through simple intersections. To assign a weight say q per black loop (component of road) and p per white loop (component of river), the simplest way is to use the replica trick of Section 2.3: for positive integer values of p and q, introduce q "black" Hermitian matrices $B_1, B_2, \dots, B_q$ and p "white" Hermitian matrices $W_1, W_2, \dots, W_p$, all of size $N \times N$, with the only nonvanishing propagators

white edges: $\langle (W_a)_{ij}(W_b)_{kl}\rangle = \frac{1}{N}\delta_{a,b}\delta_{il}\delta_{jk} =$ 

black edges: $\langle (B_a)_{ij}(B_b)_{kl}\rangle = \frac{1}{N}\delta_{a,b}\delta_{il}\delta_{jk} =$ 

and only simple intersection vertices

$\mathrm{Tr}(W_a B_b W_a B_b) =$ 

for all $1 \le x \le p$ and $1 \le y \le q$.

The case of a unique river will then be recovered by taking the limit $p \to 0$. This suggests to introduce the "black and white" matrix integral

$$\begin{aligned} Z_{q,p}(N;x) &= \frac{1}{Z_0}\int \prod_{a=1}^{q} dB_a \prod_{b=1}^{p} dW_b e^{-N\,\mathrm{Tr}\,V(\{B_a\},\{W_b\})}, \\ V(\{B_a\},\{W_b\}) &= \frac{1}{2}\bigg(\sum_{a=1}^{q} B_a^2 + \sum_{b=1}^{p} W_b^2 - x\sum_{a=1}^{q}\sum_{b=1}^{p} B_a W_b B_a W_b\bigg), \end{aligned} \tag{6–1}$$

where Z_0 is a normalization factor ensuring that $Z_{q,p}(N;0) = 1$. As before, the corresponding free energy may be formally expanded as a sum over all possible connected black and white graphs as

$$\begin{aligned} F_{q,p}(N;x) &= \frac{1}{N^2}\,\mathrm{Log}\, Z_{q,p}(N;x) \\ &= \sum_{\text{black \& white }\Gamma} \frac{1}{|\mathrm{Aut}\,\Gamma|} N^{-2g(\Gamma)} x^{v(\Gamma)} q^{L_b(\Gamma)} p^{L_w(\Gamma)}, \end{aligned} \tag{6–2}$$

where, as in Section 2.2, $g(\Gamma)$, $v(\Gamma)$ and $\mathrm{Aut}(\Gamma)$ denote respectively the genus, number of vertices, and symmetry group of Γ, whereas $L_b(\Gamma)$ and $L_w(\Gamma)$ denote respectively the numbers of black and white loops and black and white edges in Γ. To get a generating function for meander polynomials from (6–2), we simply have to take $N \to \infty$ to only retain the planar graphs (genus zero), and then only compute the coefficient of w in the resulting expression as a power series of w, which leads to

$$F_q(x) \equiv \lim_{N\to\infty} \frac{\partial F_{q,p}(N;x)}{\partial p}\bigg|_{p=0} = \sum_{n=1}^{\infty} \frac{x^{2n}}{4n} m_{2n}(q), \tag{6–3}$$

where we simply have identified planar black and white fatgraphs with meanders (with the river closed into a loop), whose symmetry group is $\mathbb{Z}_{2n} \times \mathbb{Z}_2$ for the cyclic symmetry along the looplike river, and the symmetry between inside and outside of that loop. The meander polynomial is as in (4–1). The large n asymptotic behavior of the meander polynomial $m_{2n}(q)$ can be directly linked to the critical behavior of the generating function $F_q(x)$ (6–3). Indeed, (4–2) translates into a singular part

$$F_q(x)_{\text{sing}} \sim (x(q) - x)^{\mu(q)},$$

where

$$x(q) = \frac{1}{R(q)}, \qquad \mu(q) = \alpha(q).$$

So, if we can investigate the critical properties of $F_q(x)$, the meander asymptotics will follow.

The black-and-white model in the large N limit also allows for a simple representation of semimeanders as a correlation function. Indeed, considering the operator

$$\phi_1 = \lim_{N\to\infty} \frac{1}{N} \operatorname{Tr}(W_1)$$

and computing the correlation function

$$\langle \phi_1 \phi_1 \rangle_{q,p} = \frac{1}{Z_0} \int \prod dW_b dB_a \phi_1 \phi_1 e^{-N \operatorname{Tr} V(\{B_a\},\{W_b\})}$$

by use of the Wick theorem, we see that ϕ_1 creates a white line at a point, hence the fatgraphs contributing to the $p \to 0$ limit will simply have a white segment joining the two endpoints created by the two ϕ_1's, intersected by arbitrary configurations of road. Given such a planar graph, we may always send one of the two endpoints to infinity thus yielding a semi-infinite river, and the fatgraphs with n intersections are just the semimeanders of order n. Hence

$$\sum_{n\geq 1} \bar{m}_n(q) x^n = \langle \phi_1 \phi_1 \rangle_{q,0}.$$

Again, we will get the semimeander asymptotics from the singular behavior of this correlation function when $x \to x_c$.

Before taking the limit as $p \to 0$, we could have written

$$F_{q,p}(x) \equiv \lim_{N\to\infty} F_{q,p}(N;x) = \sum_{n=1}^{\infty} \frac{x^{2n}}{4n} m_{2n}(q,p), \tag{6–4}$$

where we have defined a meander polynomial $m_{2n}(q,p)$ for multicomponent meanders with also multiple rivers (and a weight q per road and p per river). In particular $m_{2n}(q,p) \sim p m_{2n}(q)$ when $p \to 0$. Similarly,

$$\langle \phi_1 \phi_1 \rangle_{q,p} = \sum_{n \geq 1} \bar{m}_n(q,p) x^n \tag{6–5}$$

defines a semimeander polynomial with also multiple rivers, one of which is semi-infinite. In Section 6.3 below, we will derive the exact asymptotics of the polynomial $m_{2n}(q,p=1)$ from the black-and-white matrix model, while in Section 6.4, we will present an argument yielding all the critical configuration exponents for large n.

Note also that if we keep N finite, we get an all genus expansion

$$\frac{\partial F_{q,p}(N;x)}{\partial p}\bigg|_{p=0} = \sum_{g \geq 0} N^{-2g} \sum_{n=1}^{\infty} \frac{x^n}{2n} m_n^{(g)}(q), \tag{6–6}$$

where we have defined the genus g meander polynomials $m_n^{(g)}(q)$ that counts the natural higher genus generalization of meanders of given genus and number of connected components of road, and with one river.

6.2. Meander Polynomials and Gaussian Words. The integral (6–1) can be considerably simplified by noting that it is just multi-Gaussian say in the white matrix vector $\vec{W} = (W_1, \ldots, W_p)$. Using the recipes of Section 3.4, we first identify the quadratic form involving each W_b in (6–1), namely

$$\begin{aligned} W_b^t \mathbf{Q} W_b &= \sum_{i,j,k,l=1}^{N} (W_b)_{ji} Q_{ij,kl} (W_b)_{kl} \\ &= \mathrm{Tr}\left(W_b^2 - x \sum_{a=1}^{q} W_b B_a W_b B_a \right); \end{aligned}$$

hence

$$Q_{ij,kl} = \delta_{ik}\delta_{jl} - x \sum_{a=1}^{q} (B_a)_{ik}(B_a)_{lj},$$

or, more compactly,

$$\mathbf{Q} = I \otimes I - x \sum_{a=1}^{q} B_a \otimes B_a^t. \tag{6–7}$$

After integration over the W's in (6–1), we are left with

$$Z_{q,p}(N;x) = \frac{\int \prod_{a=1}^{q} dB_a e^{-\frac{N}{2}\mathrm{Tr}\, B_a^2} \det(\mathbf{Q})^{-p/2}}{\int \prod_{a=1}^{q} dB_a e^{-\frac{N}{2}\mathrm{Tr}\, B_a^2}}.$$

Extracting the $p \to 0$ limit is easy, with the result

$$\frac{\partial F_{q,p}(N;x)}{\partial p}\bigg|_{p=0} = \frac{1}{N^2} \langle -\frac{1}{2}\,\mathrm{Tr}\,\mathrm{Log}\,\mathbf{Q} \rangle,$$

where the bracket stands as in Section 2.3 for the multi-Gaussian average with respect to the B's. Expanding Log $\mathbf{Q}$ of (6–7) as a formal power series of x, we get finally

$$\frac{\partial F_{q,p}(N;x)}{\partial p}\bigg|_{p=0} = \sum_{m\geq 1}\frac{(-x)^m}{2m}\sum_{a_1,a_2\ldots,a_m=1}^{q}\langle|\frac{1}{N}\operatorname{Tr}(B_{a_1}B_{a_2}\ldots B_{a_m})|^2\rangle.$$

Comparing this with (6–6), we identify the all-genus meander polynomial generating function

$$\sum_{g\geq 0}N^{-2g}m_n^{(g)}(q) = \sum_{a_1,a_2\ldots,a_n=1}^{q}\langle|\frac{1}{N}\operatorname{Tr}(B_{a_1}B_{a_2}\ldots B_{a_n})|^2\rangle. \tag{6–8}$$

Concentrating on the genus zero case, we may express the meander polynomial (4–1) as a multi-Gaussian integral

$$m_{2n}(q) = \sum_{a_1,a_2\ldots,a_{2n}=1}^{q}\lim_{N\to\infty}\langle|\frac{1}{N}\operatorname{Tr}(B_{a_1}B_{a_2}\ldots B_{a_n})|^2\rangle, \tag{6–9}$$

where we have noted that the right-hand side of (6–8) vanishes for large N if n is odd, by a simple parity argument. Moreover, the right-hand-side of (6–9) must be computed using the planar Wick theorem. It is clear however that no pairing can be made between B matrix elements pertaining to the two trace terms, as that would violate planarity. In other words, in the large N limit, we have $\langle\operatorname{Tr} f(B_j)\operatorname{Tr} g(B_j)\rangle = \langle\operatorname{Tr} f(B_j)\rangle\langle\operatorname{Tr} g(B_j)\rangle$ for any functions f and g of the B's. Hence, using the definition (3–38), we may finally express the meander polynomial in terms of planar multi-Gaussian averages of words

$$m_{2n}(q) = \sum_{a_1,a_2\ldots,a_{2n}=1}^{q}|\gamma_{a_1,a_2,\ldots,a_{2n}}|^2, \tag{6–10}$$

where, for $a_1,\ldots a_p\in\{1,2,\ldots,q\}$, we have set

$$\gamma_{a_1,a_2,\ldots,a_p} = \eta_{m_1,\ldots,m_{qr-1}} \quad \text{if } B_{a_1}B_{a_2}\ldots B_{a_p} = B_1^{m_1}\ldots B_q^{m_q}B_1^{m_{q+1}}\ldots B_q^{m_{qr-1}}.$$

Having proved (6–10) through some quite lengthy process, it is instructive to interpret (6–10) directly. Using the planar Wick theorem to compute the right-hand side of (6–10), we will represent for $\gamma_{a_1,\ldots,a_{2n}}$ the chain of $2n$ matrix elements as equally spaced points along a line, each with the corresponding color $a_i\in\{1,2,\ldots,q\}$ (see Figure 9 for an illustration). A given planar Wick pairing is nothing but an arch configuration linking these points by pairs, *provided they have the same color*. As we have to multiply two γ's, we get a pair of such colored arch configurations. The fact that the two γ's are conjugate of one another simply means that the two arch configurations may be superposed, and that their bridge colors match: indeed, $\gamma^*_{a_1,\ldots,a_{2n}} = \gamma_{a_{2n},\ldots,a_1}$, so we just have to represent its pairings head down and use the same line as for γ. So the net result is the production, for each (pair of) Wick pairing pertaining to the

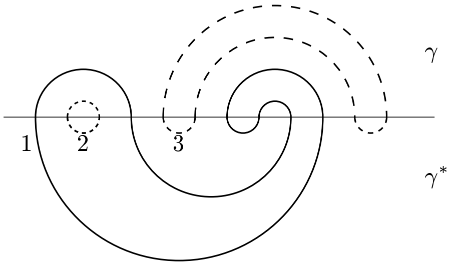


Figure 9. A typical pair of Wick pairings used to compute the quantity $|\gamma_{1,2,2,1,3,3,1,1,1,1,3,3}|^2$. Each pair is represented by a semicircle with the corresponding color (1 = solid, 2 = dashed, 3 = dotted). A pairing for γ is represented, as a colored arch configuration of order 12 here. A pairing for γ^* is also represented, but head-down, to match the bridge colors. The net result is a colored meander with 3 components of road.

two γ's and for each choice of colors $a_1, a_2, \ldots, a_{2n}$ of the bridges, of a *colored* multicomponent meander, namely with colored roads. As the colors are summed over, we get a final weight of q per connected component of road, which matches the definition of the meander polynomial.

This interpretation of the result (6–10) allows to immediately generalize it to the semimeander case:

$$\bar{m}_n(q) = \sum_{a_1,a_2,\ldots,a_n=1}^{q} \gamma_{a_1,a_2,\ldots,a_n,a_n,a_{n-1},\ldots,a_2,a_1}. \tag{6–11}$$

Indeed, the planar Wick pairings for the γ form again colored arch configurations of order $2n$ with the constraint that they have symmetrically identical colors. This latter constraint can be represented by a lower rainbow diagram r_{2n} with the corresponding colors. The net result is a colored semimeander, and the result (6–11) follows exactly as before.

By analogy with (6–8), we may also define the genus g semimeander polynomials as

$$\sum_{g\geq 0} N^{-2g}\bar{m}_n^{(g)}(q) = \sum_{a_1,a_2\ldots,a_n=1}^{q} \left\langle \left| \frac{1}{N}\operatorname{Tr}(B_{a_1}B_{a_2}\ldots B_{a_n}B_{a_n}\ldots B_{a_1})\right| \right\rangle.$$

Having expressed all our quantities of interest in terms of γ's, we now have an alternative route for evaluating them. We have to use the recursion relations (3–39) to first compute the γ's and then to substitute them in the above expressions for (semi-)meander polynomials. This is still however a tedious work, and it doesn't really allow for any asymptotic study at large n. One advantage though of these expressions is that as the γ's are positive numbers, they allow for exact lower bounds on the meander polynomials.

A last remark is in order: in this section, the meander and semimeander polynomials have been computed only for positive integer values of q. Analytic continuation to any positive real q is however immediate as these are polynomials.

6.3. Exact Asymptotics for the Case of Arbitrary Many Rivers. In this section, following [25], we will use the black and white matrix model for meanders to evaluate the exact asymptotics of slightly different quantities, namely the "multimeander" polynomials $m_{2n}(q,1)$ counting the meanders with fixed numbers of connected components of roads and arbitrary numbers of rivers, closed into nonintersecting loops. However this side result already announces the flavor of what true meander asymptotics should look like.

We wish to compute the function $Z_{q,1}$ (6–1) by first integrating over all the B matrices, rather than the single W one. Using the form (6–7) with $B \to W$, we are then left with

$$Z_{q,1}(N;x) = \frac{\int dW e^{-N\,\mathrm{Tr}(W^2/2)} \det(1\otimes 1 - xW\otimes W^t)^{-q/2}}{\int dW e^{-N\,\mathrm{Tr}(W^2/2)}}.$$

This is just a Gaussian average over one Hermitian matrix W. Applying the reduction method of Section 3.1, we arrive at the eigenvalue integral

$$Z_{q,1}(N;x) = \frac{1}{(2\pi N)^{N/2}} \int dw_1 \ldots dw_N \frac{\prod_{i<j}(w_i-w_j)^2}{\prod_{i,j}(1-xw_iw_j)^{q/2}} e^{-N\sum_i w_i^2/2}.$$

This is very similar to the eigenvalue integral (3–25) we obtained in the case of the $O(n)$ model of Section 3.4. Actually, upon the change of variables

$$w = \frac{1}{\sqrt{x}}\frac{1-z}{1+z}$$

we have

$$\begin{aligned} Z_{q,1}(N;x) = {} & \frac{2^{\frac{N^2}{2}(2-q)}}{(2\pi N x)^{N/2}} \\ & \times \int dz_1 \ldots dz_N \frac{\prod_{i<j}(z_i-z_j)^2}{\prod_{i,j}(z_i+z_j)^{q/2}} \prod_i (1+z_i)^{N(q-2)} e^{-N\sum_i \frac{1}{2x}\left(\frac{1-z_i}{1+z_i}\right)^2}. \end{aligned}$$

We now evaluate the large N behavior of the integral using the saddle-point technique of Section 3.4. The action to be minimized reads now

$$\begin{aligned} S(z_1,\ldots,z_N) &= \frac{1}{N}\sum_{i=1}^N v(z_i) - \frac{1}{N^2}\sum_{1\le i\ne j\le N} \mathrm{Log}\,|z_i-z_j| - \frac{q}{2N^2}\sum_{1\le i,j\le N}\mathrm{Log}(z_i+z_j) \\ v(z) &= \frac{1}{2x}\Big(\frac{1-z}{1+z}\Big)^2 + (q-2)\,\mathrm{Log}(1+z) \\ & . \end{aligned} \tag{6–12}$$

Note that the sum in the last term of S may be restricted to $i \neq j$ up to a term of order $O(1/N)$ that does not contribute to the large N leading asymptotics. Expressing that $\partial_{z_i} S = 0$, we finally get

$$v'(z_i) = \frac{2}{N} \sum_{j \neq i} \frac{1}{z_i - z_j} + \frac{q}{N} \sum_{j \neq i} \frac{1}{z_i + z_j}.$$

This is exactly the saddle-point equation (3–26) for $n = q$ and the particular choice (6–12) of the potential $v(z)$. Note that v' is a meromorphic function of z with a third order pole at $z = -1$. As before, we assume that the limiting eigenvalue distribution has a support made of a single interval $[a, b]$, $0 < a < b$. This requirement turns out to fix entirely the meromorphic functions $S(z)$ and $P(z)$ in (3–34). The critical singularity of the genus zero free energy $f \sim (x(q) - x)^{2-\gamma}$ is found to lie at [25]

$$x = x(q) = \frac{f^2}{2 \sin(\pi \frac{f}{2})},$$

where we have set $q = 2\cos(\pi f)$, and the corresponding critical exponent γ reads

$$\gamma = -\frac{f}{1-f}.$$

As mentioned before, these translate into multiriver meander asymptotics

$$m_{2n}(q,1) \sim \frac{R(q,1)^{2n}}{n^{\alpha(q,1)}}$$

as

$$R(q,1) = \frac{1}{x(q)} = 2\,\frac{\sin^2(\pi f/2)}{f^2}, \quad \alpha(q,1) = \frac{2-f}{1-f}. \tag{6–13}$$

In particular, we recover from these values the case of meanders with one river and arbitrary many connected components of road, with $q = 0$, $f = \frac{1}{2}$, and $R(1,0) = R(0,1) = 4$, $\alpha(1,0) = \alpha(0,1) = 3$. We list a few of these values for various fractions f in Table 2.

6.4. Exact Meander Asymptotics from Fully-Packed Loop Models coupled to Two-dimensional Quantum Gravity. In a recent work [18] it was noticed that the black-and-white matrix model is the natural random surface version of some Fully Packed loop model on the square lattice. The latter is defined by assigning a color (B or W) to each edge of the square lattice, in such a way that two edges of each color meet at each vertex. These edges then form (Fully Packed) loops each of which is assigned a weight p or q for W and B loops respectively. The model is called the $FPL^2(p,q)$ model [26]. When defined on a random surface of genus zero, the model assigns colors to the edges of a random fatgraph with only vertices of the form $BWBW$ (crossing) or $BBWW$ (avoiding). The black-and-white model of Section 6.1 does not have the second kind of vertices. Therefore the original Fully Packed loop model has been further restricted.

q	f	$R(q,1)$	$\alpha(q,1)$	$R(q)$
0	$\frac{1}{2}$	4	3	3.50
1	$\frac{1}{3}$	$\frac{9}{2}$	$\frac{5}{2}$	4
$\sqrt{2}$	$\frac{1}{4}$	$16-8\sqrt{2}=4.68\ldots$	$\frac{7}{3}$	$4.13\ldots$
$\sqrt{3}$	$\frac{1}{6}$	$36-18\sqrt{3}=4.82\ldots$	$\frac{11}{5}$	$4.27\ldots$
2	0	$\pi^2/2=4.93\ldots$	2	$4.42\ldots$

Table 2. Multi-river meander asymptotics. We have listed a few values of $q=2\cos\pi f$, together with the corresponding values of $R(q,1)$ and $\alpha(q,1)$, and the numerical values of $R(q)$ obtained from the direct enumeration results of Section 4.2 for comparison. Note that $R(q,1)\simeq R(q)+\frac{1}{2}$ with good precision.

The detailed study of the $FPL^2(p,q)$ model shows two remarkable facts: (i) it is critical for all values of $0\le p,q\le 2$ (ii) it is represented in the continuum limit by a Conformal Theory with central charge

$$c_{FPL}(q,p)=3-6\Big(\frac{e^2}{1-e}+\frac{f^2}{1-f}\Big),$$

where $p=2\cos\pi e$ and $q=2\cos\pi f$. This was proved by mapping the $FPL^2(p,q)$ model onto a three-dimensional height model, where the heights are defined in the center of each face, with an Ampère-like rule prescribing the transitions from one face to its neighbors. In the continuum limit, the height variable becomes a three-dimensional free field (conformal theory with central charge $c=3$), and the corrective weights assigning the factors p and q per loop of each color account for the correction of c by electric charges e,f.

The restriction we impose here on the $FPL^2(p,q)$ model on a random surface amounts to restricting the height variable to only two dimensions instead of three. The correct formula for the flat space Fully Packed Loop theory is therefore

$$c(q,p)=2-6\Big(\frac{e^2}{1-e}+\frac{f^2}{1-f}\Big),\quad p=2\cos\pi e,\quad q=2\cos\pi f, \tag{6–14}$$

with $e,f\in\big[0,\frac{1}{2}\big]$ (so $0\le p,q\le 2$). We therefore state that the black-and-white model of Section 6.1 is described in the planar (large N) limit by the gravitational version of a conformal theory with central charge (6–14), namely the same theory defined on fluctuating surfaces, that have to be summed over statistically. Note that the $O(n)$ model, whose gravitational version has been introduced in Section 2.3 is also described in the dense phase by a conformal theory of central charge $c=1-6g^2/(1-g)$ where $n=2\cos\pi g$. So we are now dealing with a sort of double $O(n)$ model coupled to gravity.

The coupling to gravity of a conformal theory with central charge $c\le 1$ has been extensively studied within the context of noncritical string theory. The gravitational theory has a new parameter x, called the cosmological constant, coupled to the area of the surfaces we have to sum over. More precisely, the free energy for a conformal theory coupled to gravity in genus zero reads

$$F = \operatorname{Log} Z = \sum_{A\geq 0} x^A \sum_{\substack{\text{connected surfaces } \Gamma \\ \text{of area } A}} Z_{CFT}(\Gamma), \tag{6–15}$$

where $Z_{CFT}(\Gamma)$ denotes the partition function of the conformal theory on the genus zero connected surface Γ. Comparing Z with the black-and-white model partition function (6–1), we see that x plays the role of cosmological constant, as $n = A$ are the areas of the tessellations dual to the fatgraphs of the model. When the conformal theory has central charge c, the free energy (6–15) has been shown to have a singularity of the form (3–35), (3–36):

$$F \sim (x_c - x)^{2-\gamma} \qquad \gamma = \tfrac{1}{12}\big(c - 1 - \sqrt{(1-c)(25-c)}\,\big)$$

when x approaches some critical value x_c. This is easily translated into the large area asymptotics of the partition function of the model on surfaces of fixed area

$$F_A \sim \frac{x_c^{-A}}{A^{3-\gamma}}. \tag{6–16}$$

Moreover, the operators of the conformal theory get "dressed" by gravity, and their correlation functions have singularities of the form

$$\langle \phi_{m_1} \dots \phi_{m_k} \rangle \sim (x_c - x)^{\sum \Delta_{m_i} - \gamma + 2 - k},$$

where the "dressed dimensions" Δ_m are related to the conformal dimensions h_m of their undressed versions in the conformal theory through [4]

$$\Delta_m = \frac{\sqrt{1-c+24h_m} - \sqrt{1-c}}{\sqrt{25-c} - \sqrt{1-c}}. \tag{6–17}$$

We now have all the necessary material to compute the configuration exponents of all the meandric numbers of interest. Applying the result (6–16) to the central charge (6–14), we find the configuration exponent of the multiriver meander polynomial (6–4)

$$\begin{aligned} m_{2n}(q,p) &\sim \frac{R(q,p)^{2n}}{n^{\alpha(q,p)}}, \\ \alpha(q,p) &= 2 + \tfrac{1}{12}\sqrt{1-c(q,p)}\big(\sqrt{25-c(q,p)} + \sqrt{1-c(q,p)}\,\big), \end{aligned}$$

as well as that of the multiriver semimeander polynomial (6–5)

$$\bar{m}_n(q,p) \sim \frac{R(q,p)^n}{n^{\bar{\alpha}(q,p)}}, \qquad \bar{\alpha}(q,p) = \alpha(q,p) - 1 + 2\Delta_1, \tag{6–18}$$

where Δ_1 is the dressed dimension of the operator creating a white endpoint. In the conformal theory, this operator is known to have the dimension [27]

$$h_1 = \frac{1-e}{16} - \frac{e^2}{4(1-e)},$$

where $p = 2\cos \pi e$. We simply have to apply (6–17) to the preceding equation and substitute the value of Δ_1 back into (6–18).

For $p = 1$ ($e = \frac{1}{3}$) and q arbitrary, we find

$$\alpha(q,1) = \frac{2-f}{1-f}, \qquad \bar{\alpha}(q,1) = \frac{1}{1-f},$$

for all $q = 2\cos \pi f$. The first of these equations agrees with the saddle point result (6–13). The second is readily obtained by noticing that $\Delta_1 = 0$ when $e = \frac{1}{3}$, and therefore $\bar{\alpha}(q,1) = \alpha(q,1) - 1$.

For $p = q = 0$ ($e = f = \frac{1}{2}$), we get the exact values of the meander and semimeander configuration exponents:

$$\alpha(0,0) = 2 + \frac{1}{12}\sqrt{5}\big(\sqrt{5} + \sqrt{29}\big),$$
$$\bar{\alpha}(0,0) = 1 + \frac{1}{24}\sqrt{11}\big(\sqrt{5} + \sqrt{29}\big).$$

Note that the arguments of this section do not give any prediction for nonuniversal quantities such as $R(q,p)$ (which is expected to depend on q and p explicitly, not just on $c(q,p)$).

Finally, for $p = 0$ and q arbitrary, we find:

$$\alpha(q,0) = 2 + \frac{1}{12}\sqrt{1-c(q)}\big(\sqrt{25-c(q)} + \sqrt{1-c(q)}\,\big),$$
$$\bar{\alpha}(q,0) = 1 + \frac{1}{24}\sqrt{3-4c(q)}\big(\sqrt{25-c(q)} + \sqrt{1-c(q)}\,\big),$$

with $c(q)$ given by (3–37). Note that the second of these equations breaks down when $q = q_c$ corresponding to $c(q_c) = \frac{3}{4}$, namely

$$q_c = 2\cos\left(\pi\frac{\sqrt{97}-1}{48}\right).$$

We identify this as the critical value of q beyond which the winding becomes relevant in semimeanders, namely when circles (roads intersecting the river only once) dominate the semimeander configurations.

7. Folding Triangulations

We now turn to our second application of matrix integrals, having to do with the generation of foldable two-dimensional triangulations. These triangulations are a simple model for so-called tethered or fluid membranes, objects of physical and biological interest.

Although we now deal with the folding of two-dimensional objects (as opposed to the one-dimensional polymers of part B), we will rephrase the problem as that of enumerating vertex-tricolored triangulations. A suitable matrix model will be presented and solved using various techniques.

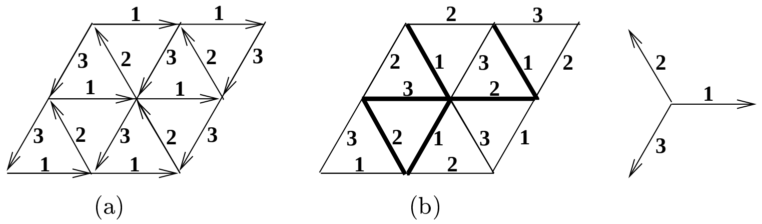


Figure 10. A choice (a) for the tangent vectors of the triangular lattice, together with the corresponding coloring of the edges by $1, 2, 3$. This is the flat configuration of the membrane. A folding configuration (b) with the corresponding folded bonds (thick black lines) and edge coloring. The three colors correspond to the three unit vectors with vanishing sum represented above.

As mentioned above, one usually distinguishes between two distinct models for membranes:

Tethered Membranes are represented by reticulated networks, typically some domain of a regular lattice with physical vertices and rigid bonds, allowed only to change their spatial configuration through folding.

Fluid Membranes are represented also by now irregular reticulated networks in which the valency of vertices is no longer fixed, this disorder accounting for the fluidity. But the bonds are still rigid, and the membrane is still allowed to change its spatial configuration through folding.

In the following, we first briefly describe some results [28] on the triangular lattice folding (tethered membrane folding), before turning to the problem of folding triangulations [29] (fluid membranes). Here we only consider the so-called "phantom folding" of membranes; that is, we allow the network to interpenetrate itself, and are only interested in the statistics of the final folded states, independently of the actual feasibility of the folding process. Moreover, the folding of the triangular lattice will be two-dimensional, in that the folded configurations will be subsets of the original lattice.

7.1. Folding the Triangular Lattice. We consider the triangular lattice, or rather a rhombus-shaped portion of it with size $N \times P$. We view this as a reticulated network, with solid edges of unit length. Each such edge carries a unit tangent vector $\vec{t}$, with the constraint that

$$\sum_{\text{around faces}} \vec{t} = \vec{0}. \tag{7–1}$$

With the choice of tangent vectors indicated in Figure 10(a), a folding of the triangular lattice is a continuous map $\rho : S \to \mathbb{R}^2$, preserving the length of the tangent vectors and satisfying the condition (7–1) around each triangular face of S. Let $\vec{t}_1, \vec{t}_2, \vec{t}_3$ denote the unit tangent vectors to a given face of S. Their images $\rho(\vec{t}_i)$ are three unit vectors with vanishing sum, according to the face rule (7–1).

Fixing the image of one tangent vector of S to be a given unit vector $\vec{e_1}$, we see that the images of the tangent vectors of S may only take three values $\vec{e_1}, \vec{e_2}, \vec{e_3}$, where $\vec{e_1}, \vec{e_2}, \vec{e_3}$ are three unit vectors with vanishing sum, hence forming angles of $2\pi/3$. We associate colors numbered $1, 2, 3$ to these three possible images.

A folding map ρ of the triangular lattice is therefore a coloring of its edges, with the three colors $1, 2, 3$, such that the three colors of edges around each face are all distinct. An example of such a coloring is given in Figure 10(b) together with the corresponding folding configuration. The dual of this coloring model is the problem of tricoloring the edges of the hexagonal (honeycomb) lattice in such a way that the three edges adjacent to each vertex are painted with distinct colors $1, 2, 3$. It has been solved by Baxter [30], by use of the Bethe Ansatz. Baxter's results yield in particular the exact value for the thermodynamic entropy of folding per face of the triangular lattice $s = \lim_{N,P\to\infty} \frac{1}{NP} \operatorname{Log} Z_{N,P}$, where $Z_{N,P}$ denotes the total number of distinct folded configurations of our portion of lattice. It reads

$$s = \operatorname{Log}\left(\frac{\sqrt{3}}{2\pi} \Gamma\left(\tfrac{1}{3}\right)^{3/2} \right). \tag{7–2}$$

This was originally proved by explicitly diagonalizing a large (transfer) matrix, indexed by the coloring configurations of rows of N edges in the honeycomb lattice, and describing the "row-to-row transfer", that is, the allowed coloring configurations for two such neighboring rows. The thermodynamic entropy (7–2) is then the logarithm of the largest (Perron–Frobenius) eigenvalue of this matrix. The diagonalization is performed using a particular ansatz for the form of the eigenvectors, the Bethe Ansatz. The proof of (7–2) being highly technical, we will not reproduce it here, but refer the interested reader to the original papers [30]. We simply mention that this model is part of the class of Two-dimensional Integrable Lattice Models, for which a Bethe Ansatz solution exists.

A score of other lattice folding problems have been studied [31]. Actually, one can classify all the (compactly) foldable lattices [32], and define higher-dimensional generalizations of their folding problems. Remarkably, all of them lead to some equivalent coloring problems, also rephrased into Fully Packed loop enumeration problems. For instance, in the case of folding the triangular lattice, we have seen that the model is equivalent to that of tricoloring the edges of the lattice. Concentrating say on the colors 1 and 2, we see that edges of alternating colors $1212\ldots$ form loops on the dual (honeycomb) lattice, and moreover each vertex of the dual is visited by exactly one of these loops: the loops $1212\ldots$ are therefore fully packed, as well as the $2323\ldots$ and $1313\ldots$ Remarkably, the same type of Fully Packed loop models have emerged in our study of meanders (see Section 6.4).

7.2. Foldable Triangulations. Fluid membranes are modelled by irregular networks of vertices linked by edges, in which the valencies of the vertices are arbitrary, as well as the genus of the underlying surface, which might have an arbitrary topology. As advocated in Section 2, matrix models provide us with a means of generating such graphs.

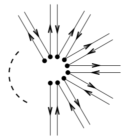
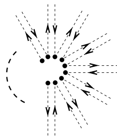
We restrict ourselves to networks with only triangular faces, namely triangulations. The initial question one can ask is: are all triangulations foldable? It is indeed desirable to consider only triangulations with a large number of folded configurations, otherwise the effect of folding might be wiped out in the limit of large size. We therefore demand that our triangulations be *compactly* foldable, namely that one can fold them completely onto just one of their (equilateral triangular) faces. With this constraint, it is clear that not all triangulations turn out to be "foldable". Indeed, let us paint by three distinct colors $1, 2, 3$ the three vertices of the image triangle, and paint accordingly the vertices of the preimages under the folding map. This results in the tricoloring of the vertices of the initial triangulation, in such a way that the three colors around each triangular face are distinct. So only the vertex-tricolorable triangulations will be compactly foldable. Another way of viewing this restriction is to recall that we first need to attach tangent vectors to the edges of the triangulation, in such a way that (7–1) is satisfied. It is straightforward to see that this is possible only if the vertices of the triangulation are all even, as around such a vertex, we must have an alternance of tangent vectors pointing to and from it. This condition turns out to be sufficient in genus zero to grant the tricolorability of the triangulation. The situation in higher genus is unclear [33].

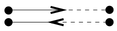
We now introduce a generating function $Z(x_1, x_2, x_3; t; N)$ for possibly disconnected vertex-tricolored triangulations of arbitrary genus, such that

$$\begin{aligned} F(x_1, x_2, x_3; t; N) &= \operatorname{Log} Z(x_1, x_2, x_3; t; N) \\ &= \sum_{\substack{\text{vertex}-\text{tricolored} \\ \text{connected triangulations } T}} x_1^{n_1(T)} x_2^{n_2(T)} x_3^{n_3(T)} \frac{t^{A(T)/2} N^{2-2h(T)}}{|\operatorname{Aut} T|}, \end{aligned}$$

where $n_i(T)$ denote the total numbers of vertices of color i, $A(T)$ the total number of faces, $h(T)$ the genus and $|\operatorname{Aut} T|$ the order of the symmetry group of the tricolored triangulation T.

The construction of a matrix model to represent $Z(x_1, x_2, x_3; t; N)$ is based on the following simple remark: in a given tricolored triangulation T if we remove say all the vertices of color 3 and all edges connected to them, we end up with a bicolored graph, with unconstrained vertex valencies. Such bicolored graphs are easily built out of the Feynman graphs of a two Hermitian matrix model, say M_1 and M_2, the index 1 and 2 standing for the color, with colored vertices

$\operatorname{Tr}(M_1^n) =$  $\longleftrightarrow N x_1$, $\operatorname{Tr}(M_2^n) =$  $\longleftrightarrow N x_2$, connected through propagators imposing the alternance of colors, namely

$$\langle (M_a)_{ij}(M_b)_{kl}\rangle \;=\; (1-\delta_{ab})\delta_{jk}\delta_{il}\frac{t}{N} =$$  ,

where $a,b=1,2$. We introduce the corresponding matrix integral, but keep N fixed while the matrices are taken of size $n\times n$, n possibly different from N. This gives the partition function

$$\begin{aligned} Z_n(x_1,x_2;t;N) &= \frac{1}{\varphi_n(t,N)}\int dM_1\, dM_2\, e^{-N\ \mathrm{Tr}\, V(M_1,M_2;x_1,x_2,t)}, \\ V(M_1,M_2;x_1,x_2,t) &= x_1 \operatorname{Log}(1-M_1) + x_2 \operatorname{Log}(1-M_2) + \frac{1}{t}M_1M_2, \end{aligned} \tag{7–3}$$

where the normalization factor $\varphi_n(t,N)$ ensures that $Z_n(0,0;t;N)=1$. Clearly we must take the integral (7–3) only at the level of formal series, by expanding it in power series of x_1 and x_2, and computing the Gaussian integrals with measure $dM_1dM_2e^{-(N/t)M_1M_2}$ by the Feynman procedure. In particular, this gives the following rule for two-dimensional integrals

$$\begin{aligned} \langle x^\alpha y^\beta\rangle &= \frac{N}{t}\int dxdye^{-(N/t)xy}x^\alpha y^\beta \\ &= \delta_{\alpha\beta}\Gamma(\alpha+1)\left(\frac{t}{N}\right)^{\alpha}. \end{aligned} \tag{7–4}$$

The integral could be made rigorous and convergent by considering instead a normal matrix $M_1\to M$, $M_2\to M^*$, but we will content ourselves with formal power series anyway.

We now compute the Feynman graph expansion of the free energy

$$\begin{aligned} F_n(x_1,x_2;t;N) &= \operatorname{Log} Z_n(x_1,x_2;t;N) \\ &= \sum_{\substack{\text{bicolored } \textit{connected} \\ \text{graphs } \Gamma}} \frac{1}{|\mathrm{Aut}\,\Gamma|} x_1^{n_1(\Gamma)} x_2^{n_2(\Gamma)} t^{E(\Gamma)} N^{(V(\Gamma)-E(\Gamma))} n^{F(\Gamma)}, \end{aligned}$$

where we have denoted by $n_i(\Gamma)$ the number of vertices of color i, $V(\Gamma)=n_1(\Gamma)+n_2(\Gamma)$, $E(\Gamma)$ the number of edges, $F(\Gamma)$ the number of faces (the boundary of which are loops of running indices, accounting for a factor n each), and $|\mathrm{Aut}\,\Gamma|$ the order of the symmetry group of the bicolored fatgraph Γ. Adding a central vertex of color 3 in the middle of each face of Γ, and connecting it to all the vertices around the face with edges will result in a vertex-tricolored triangulation T. The number of such added vertices is nothing but $n_3(T)=F(\Gamma)$. Introducing

$$x_3 = \frac{n}{N},$$

we may rewrite

$$F_n(x_1,x_2;t;N) = F(x_1,x_2,x_3;t;N) \tag{7–5}$$

by using the Euler relation $2-2h(\gamma) = 2-2h(T) = V(\Gamma)-E(\Gamma)+F(\Gamma)$ and the fact that $A(T)=2E(\Gamma)$, as each edge of Γ gives rise to two triangles of T, one in each of the two faces adjacent to the edge. It is also a simple exercise to show that $|\mathrm{Aut}\, T| = |\mathrm{Aut}\, \Gamma|$. Considering (7–5) as a formal power series of t with polynomial coefficients in x_1, x_2, x_3, we see that the knowledge of $F_n(x_1, x_2; t; N)$ for integer values of n determines completely the polynomial dependence on $x_3 = n/N$ at each order in t. Hence computing $F_n(x_1, x_2; t; N)$ through the integral formulation (7–3) will yield the generating function for compactly foldable triangulations.

8. Exact Solution

We now present the exact computation of the generating function for compactly foldable triangulations, $F(x_1, x_2, x_3; t; N)$. Although it can be obtained through orthogonal polynomial techniques generalizing that of Section 3.2, we choose to present an alternative and powerful approach, using the Discrete Hirota equation, in the form of a recursion relation for the quantities $F_n(x_1, x_2; t; N)$, in terms of n, Nx_1 and Nx_2. We will then write some direct formal series expansion for the solution, and use saddle-point techniques to extract the large N behavior of the free energy. This will give a nice formula for the generating function of vertex-tricolored triangulations of genus zero.

8.1. The Discrete Hirota Equation. The step zero in computing the integral (7–3) is like in Section 3.1 the reduction to eigenvalue integrals. Diagonalizing both matrices as $M_i = U_i m_i U_i^\dagger$, with m_i diagonal and U_i unitary, the change of variables $(M_1, M_2) \to (U_1, m_1; U_2, m_2)$ has the Jacobian $J = \Delta(m_1)^2 \Delta(m_2)^2$. We encounter a problem as the potential $V(M_1, M_2; x_1, x_2, t)$ is not invariant under unitary conjugation, namely the term $\mathrm{Tr}(M_1 M_2) = \mathrm{Tr}(\Omega m_1 \Omega^\dagger m_2)$, with the unitary matrix $\Omega = U_2^\dagger U_1$. So the integral over U_i is no longer trivial, and yields a factor

$$\int d\Omega e^{-(N/t)\,\mathrm{Tr}(\Omega m_1 \Omega^\dagger m_2))} \propto \frac{\det\left[e^{-(N/t) m_{1,i} m_{2,j}}\right]_{1 \le i,j \le n}}{\Delta(m_1)\Delta(m_2)} \tag{8–1}$$

up to some normalization factor depending on n and N only. This is the celebrated Itzykson–Zuber formula of integration over the unitary group [35], itself a particular case of the Duistermaat–Heckmann localization formula [36], since the determinant in (8–1) may be viewed as a sum over the critical (saddle-) points of the integrand, at $\Omega = P_\sigma$ the permutations of the eigenvalues.

The integral (7–3) then reduces to

$$Z_n(x_1, x_2; t; N) = \frac{1}{\psi_n(t, N)} \int dm_1 dm_2 \Delta(m_1)\Delta(m_2) e^{-N\,\mathrm{Tr}(V(m_1, m_2; x_1, x_2, t))},$$

where the normalization factor $\psi_n(t, n)$ ensures that $Z_n(0, 0; t; N) = 1$. It is easily derived by expanding the two determinants as sums over permutations, with the result

$$\begin{aligned}\psi_n(t,n) &= \int dm_1 dm_2 \Delta(m_1)\Delta(m_2) e^{-Nu\,\mathrm{Tr}(m_1 m_2)} \\ &= \sum_{\sigma,\tau\in S_n} \mathrm{sgn}(\sigma\tau) \prod_{i=1}^{n} \int dm_{1,i} dm_{2,i} m_{1,i}^{\sigma(i)-1} m_{2,i}^{\tau(i)-1} e^{-(N/t) m_{1,i} m_{2,i}} \\ &= n! \prod_{i=1}^{n} \frac{(i-1)!}{(N/t)^i}, \end{aligned} \tag{8–2}$$

by use of (7–4).

Writing

$$Z_n(x_1,x_2;t;N) = \frac{1}{\psi_n(t,N)} \int \Delta(m_1)\Delta(m_2) \times \prod_{i=1}^{n} e^{-(N/t) m_{1,i} m_{2,i}} (1-m_{1,i})^{-a} (1-m_{2,i})^{-b} dm_{1,i} dm_{2,i}$$

with $a = Nx_1$, $b = Nx_2$, and using the basic definition of determinants

$$\begin{aligned}\prod_{i=1}^{n} (1-m_{k,i})^{-a_k} \Delta(1-m_k) &= \det\Big[(1-m_{k,i})^{j-a_k-1}\Big]_{1\le i,j\le n} \\ &= \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} (1-m_{k,i})^{\sigma(i)-a_k-1}\end{aligned}$$

for $k = 1, 2$, and the shorthand notation

$$Z_n(a,b) = Z_n(x_1,x_2;t;N) \quad \text{for } a = Nx_1,\ b = Nx_2,$$

we finally get

$$\begin{aligned} Z_n(a,b) &= \frac{1}{\psi_n(t,N)} \sum_{\sigma,\tau\in S_n} \mathrm{sgn}(\sigma\tau) \prod_{i=1}^{n} \int dm_{1,i}\, dm_{2,i} \\ &\qquad \times (1-m_{1,i})^{\sigma(i)-a-1} (1-m_{2,i})^{\tau(i)-b-1} e^{-(N/t) m_{1,i} m_{2,i}} \\ &= \frac{n!}{\psi_n(t,N)} \sum_{\nu\in S_n} \mathrm{sgn}(\nu) \prod_{i=1}^{n} \int dx\, dy\, (1-x)^{i-a-1} (1-y)^{\nu(i)-b-1} e^{-(N/t)xy}, \end{aligned}$$

where we have set $\nu = \tau\sigma^{-1}$, with the same signature as $\sigma\tau$, and explicitly factored out the sum over σ. Moreover, the dummy integration variables have been rebaptized x and y, and the integral can be computed by expanding the integrand as a power series of x, y and then using term by term the prescription (7–4). The partition function takes therefore the form

$$Z_n(a,b) = \frac{n!}{\psi_n(t,N)} D_n(a,b), \tag{8–3}$$

where $D_n(a,b)$ is the $n\times n$ determinant

$$D_n(a,b)=\det\left[\int dx\,dy\,(1-x)^{i-a-1}(1-y)^{j-b-1}e^{-Nuxy}\right]_{1\le i,j\le n} \tag{8–4}$$

and $\psi_n(t,N)/n!=D_n(0,0)=\prod_{1\le i\le n}(i-1)!/(N/t)^i$.

The Hirota equation [34] is simply the rephrasing in terms of $D_n(a,b)$ of the following identity satisfied by any $(n+1)\times(n+1)$ determinant D and the minors $D_{i,j}$ obtained by erasing the i-th row and j-th column, as well as the minors $D_{i_1,i_2;j_1,j_2}$ obtained by erasing the rows i_1,i_2 and columns j_1,j_2 in D:

$$D\,D_{1,n+1;1,n+1}=D_{n+1,n+1}\,D_{1,1}-D_{1,n+1}\,D_{n+1,1}.$$

When expressed in terms of $D=D_{n+1}(a+1,b+1)$, this implies the quadratic equation

$$\begin{aligned}D_{n+1}&(a{+}1,b{+}1)D_{n-1}(a,b)\\&=D_n(a{+}1,b{+}1)D_n(a,b)-D_n(a,b{+}1)D_n(a{+}1,b).\end{aligned} \tag{8–5}$$

Finally, using $\psi_{n+1}(t,N)\psi_{n-1}(t,N)/\psi_n(t,N)^2=n/(Nu)=nt/N$ from (8–2), we get the Hirota Bilinear equation for $Z_n(a,b)$ by substituting (8–3) into (8–5):

$$\begin{aligned}n\frac{t}{N}&Z_{n+1}(a{+}1,b{+}1)\,Z_{n-1}(a,b)\\&=Z_n(a{+}1,b{+}1)\,Z_n(a,b)-Z_n(a,b{+}1)\,Z_n(a{+}1,b).\end{aligned} \tag{8–6}$$

This recursion relation determines the partition function $Z_n(a,b)$ entirely, once we apply the following initial conditions. Let

$$F_n(a,b)=\operatorname{Log}Z_n(a,b)=\sum_{m\ge1}(t/N)^m\omega_m(a,b,n)$$

be the generating function for connected tricolored triangulations, with $a=Nx_1,b=Nx_2$ and $n=Nx_3$. Then, introducing the shorthand notation $\delta_x f(x)=f(x+1)-f(x)$ for finite differences, the Hirota equation (8–6) turns into a finite difference equation for $F_n(a,b)$:

$$\delta_a\delta_bF_n(a,b)=-\operatorname{Log}\left(1-n\frac{t}{N}e^{\delta_nF_n(a+1,b+1)-\delta_nF_{n-1}(a,b)}\right). \tag{8–7}$$

This turns into a nonlinear recursion relation for the coefficients $\omega_m(a,b,n)$. But thanks to their interpretation in terms of tricolored triangulation counting, namely $\sum\omega_m(Nx_1,Nx_2,Nx_3)(t/N)^m=\sum N^{2-2h}x_1^{n_1}x_2^{n_2}x_3^{n_3}\#(\text{tricol.triang.})$, all of these are polynomials of a,b,n, and all have at least a term abn in factor, as there is always at least one vertex of each color $1,2,3$ in such a triangulation. This allows for performing the discrete integration step involved in the recursion, and yields the complete solution for $F_n(a,b)$ as a formal power series of t. The first few terms read:

$$\begin{aligned}
\omega_1 &= nab,\\
\omega_2 &= \frac{nab}{2}(n+a+b),\\
\omega_3 &= \frac{nab}{3}(n^2+3(a+b)n+a^2+3ab+b^2+1),\\
\omega_4 &= \frac{nab}{4}\big(n^3+6(a+b)n^2+(6a^2+17ab+6b^2+5)n+(a+b)(a^2+5ab+b^2+5)\big).
\end{aligned}$$

Note the symmetry in a, b, n, now manifest. In the limit of large N, the genus zero free energy

$$f_0(x_1,x_2,x_3;t) = \lim_{N\to\infty} \frac{1}{N^2} F_{Nx_3}(Nx_1, Nx_2) \tag{8–8}$$

satisfies the differential equation

$$\partial_{x_1}\partial_{x_2} f_0 = -\operatorname{Log}\big(1 - t x_3 e^{\partial_{x_3}(\partial_{x_1}+\partial_{x_2}+\partial_{x_3})f_0}\big). \tag{8–9}$$

as a consequence of (8–7).

8.2. Direct Expansion and Large N Asymptotics. We will explicitly compute the partition function $Z_n(a,b)$ (8–3) by expressing the determinant (8–4) as a formal series expansion in t. In the integrand of (8–4) we expand all terms of the form

$$(1-u)^{-c} = \sum_{k\geq 0} u^k \frac{\Gamma(k+c)}{\Gamma(c)k!},$$

for $(u,c) = (x, a+1-i)$ and $(y, b+1-j)$ respectively, and then compute the defining integral by use of the prescription (7–4). After a little algebra, we arrive at

$$\begin{aligned}
D_n(a,b) &= \det\left[\sum_{k\geq 0} \frac{1}{k!}\frac{\Gamma(k+a-i+1)}{\Gamma(1+a-i)}\frac{\Gamma(k+b-j+1)}{\Gamma(1+b-j)}\Big(\frac{t}{N}\Big)^{k+1}\right]_{1\leq i,j\leq n}\\
&= \sum_{k_1,\ldots,k_n\geq 0}\prod_{i=1}^{n} \frac{(t/N)^{k_i+1}\Gamma(k_i+a-i+1)}{k_i!\,\Gamma(1+a-i)\Gamma(1+b-i)} \det\big[\Gamma(k_i+b-j+1)\big]_{1\leq i,j\leq n}.
\end{aligned}$$

Factoring $\Gamma(k_i+b-n+1)$ out of each line of the remaining determinant, we are left with the computation of the determinant

$$\det\big[(k_i+b-j)(k_i+b-j-1)\ldots(k_i+b-n+1)\big] = \det\big[q_{n-j}(k_i)\big],$$

where the polynomials q_m are monic of degree m, and as such satisfy

$$\det\big[q_{n-j}(k_i)\big] = \det\big[k_i^{j-1}\big] = \Delta(k)$$

from (3–3), which gives the Vandermonde determinant of the k's. Repeating the same trick with the columns of the determinant, we finally get the symmetric expression

$$Z_n(a,b)=\sum_{k_1,\ldots,k_n\geq 0}\Delta(k)^2\prod_{i=1}^{n}\frac{(t/N)^{k_i+1-i}}{i!\;k_i!}\frac{\Gamma(k_i+a-n+1)}{\Gamma(1+a-i)}\frac{\Gamma(k_i+b-n+1)}{\Gamma(1+b-i)}. \tag{8–10}$$

By construction, this formal series of t is a solution to the Hirota equation (8–6). Note the remarkable similarity between the expansion (8–10) and the reduction to eigenvalues of the one-matrix integral (3–2), except that the integration over eigenvalues is replaced by a sum over nonnegative integers. We indeed have the same two types of terms: (i) the squared Vandermonde determinant, that is, the repulsion term, and (ii) the product of ratios of gamma functions, that can be exponentiated into a potential term. The expansion (8–10) allows us for a computation of the large N limit of the free energy, by applying the saddle-point techniques. Indeed, writing $k_i = N\alpha_i$, we may estimate the expansion by a real integral over the α's when N is large, of the form $\int d\alpha_1 \ldots d\alpha_n \exp(-N^2 S(\{\alpha_i\}))$, where the index i itself ranging from 1 to $n = Nx_3$ may be written as $i = Ns$, $s \in [0,1]$. Expressing that this integral is dominated by the minimum of S, it is a straightforward, though tedious calculation to get the large N asymptotics of Z. We refer the reader to [29] for details, and simply give the beautifully simple result here. The genus zero free energy of the model (8–8) satisfies the following

$$\begin{aligned} t(t\partial_t)^2 f_0(x_1,x_2,x_3;t) &= F_1F_2F_3,\\ F_1(1-F_2-F_3) &= tx_1,\\ F_2(1-F_3-F_1) &= tx_2,\\ F_3(1-F_1-F_2) &= tx_3, \end{aligned} \tag{8–11}$$

where $F_i \equiv F_i(x_1,x_2,x_3;t) = tx_i + O(t^2)$ are formal series of t with polynomial coefficients of the x's. Actually it is easy to see that F_i is the generating function for tricolored rooted trees, whose root has color i. This result still awaits a good combinatorial interpretation. One way of proving (8–11) *a posteriori* is to check that it satisfies the differential equation (8–9). As an independent check, in the particular case $x_1 = x_2 = x_3 = z$ we find $F_1 = F_2 = F_3 = F$ and $F(1-2F) = tz$, in agreement with a former result of Tutte [37].

Analyzing the critical properties of f_0 as a function of t, we find a singular behavior of the form $f_0(t) \sim (t_c - t)^{5/2}$ for generic values of the x's ($x_i > 0$ and $x_i \neq x_j$ for $i \neq j$). This leads to a number of tricolored triangulations behaving as $T(x_1,x_2,x_3)^A/A^{7/2}$ in terms of their fixed area A. We recover therefore the value $\gamma = -1/2$ (3–36) of the pure gravity, meaning that the tricoloring constraint does not affect the configuration exponent of triangulations, but does affect the leading behavior through the function $T(x_1,x_2,x_3) \neq 12$.

9. Conclusion

The combinatorial applications of matrix integrals are various and many, and it would be impossible to even enumerate them all here. In this article we have chosen to concentrate on folding problems with simple physical interpretations, but this is mainly a matter of taste.

In the study of meanders, we have shown that even an almost one-dimensional problem had to be first viewed as a random graph (or surface) problem, and then matrix and quantum gravity techniques have allowed for obtaining many interesting results.

In the case of foldable triangulation enumeration, the matrix model has allowed for the derivation of a very simple formula for the genus zero counting function, that still has to be interpreted combinatorially.

The main lesson we would like to draw is that the matrix models give a maybe less intuitive but definitely different angle on graph-related combinatorial problems.

Acknowledgements

I would like to thank A. Its and P. Bleher, organizers of the semester "Random Matrices and Applications" at MSRI (Spring 1999) for their hospitality and their suggestion to write this article. This work was partially supported by the NSF grant PHY-9722060.

References

[1] G. 't Hooft, *Nucl. Phys.* **B72** (1974), 461.

[2] E. Brézin, C. Itzykson, G. Parisi and J.-B. Zuber, "Planar diagrams", *Comm. Math. Phys.* **59** (1978), 35–51.

[3] See for instance the review by P. Di Francesco, P. Ginsparg and J. Zinn-Justin, "2D gravity and random matrices", *Physics Reports* **254** (1995), 1–131, and references therein.

[4] V. G. Knizhnik, A. M. Polyakov and A. B. Zamolodchikov, "Fractal structure of 2D-quantum gravity", *Mod. Phys. Lett.* **A3** (1988), 819–826; F. David, "Conformal field theories coupled to 2-D gravity in the conformal gauge", *Mod. Phys. Lett.* **A3** (1988), 1651–1656; J. Distler and H. Kawai, "Conformal field theory and 2D quantum gravity", *Nucl. Phys.* **B321** (1989), 509–527.

[5] B. Duplantier, "Two-dimensional copolymers and exact conformal multifractality", Phys. Rev. Lett. **82** (1999), 880–883; "Exact multifractal exponents for two-dimensional percolation", Phys. Rev. Lett. **82** (1999), 3940–3943; "Conformally invariant fractals and potential theory", Phys. Rev. Lett. **84** (2000), 1363–1367.

[6] D. Bessis, "A new method in the combinatorics of the topological expansion", *Comm. Math. Phys.* **69** (1979), 147–163; D. Bessis, C. Itzykson and J.-B. Zuber, "Quantum field theory techniques in graphical enumeration", *Adv. Appl. Math.* **1** (1980), 109–157.

[7] A. Sainte-Laguë, *Avec des nombres et des lignes* (*Récréations Mathématiques*), Vuibert, Paris, 1937; J. Touchard, "Contributions à l'étude du problème des timbres poste", Canad. J. Math. **2** (1950), 385–398; W. Lunnon, "A map-folding problem", Math. of Computation **22** (1968), 193–199; DiK. Hoffman, K. Mehlhorn, P. Rosenstiehl and R. Tarjan, "Sorting Jordan sequences in linear time using level-linked search trees", Information and Control **68** (1986), 170–184; V. Arnold, "The branched covering of $CP_2 \to S_4$, hyperbolicity and projective topology", Siberian Math. Jour. **29** (1988), 717–726; K. H. Ko and L. Smolinsky, "A combinatorial matrix in 3-manifold theory", Pacific. J. Math **149** (1991), 319–336.

[8] S. Lando and A. Zvonkin, "Plane and projective meanders", *Theor. Comp. Science* **117** (1993), 227–241; "Meanders", *Selecta Math. Sov.* **11** (1992), 117–144.

[9] R. Bacher, "Meander algebras", prépublication de l'Institut Fourier n^o 478, 1999.

[10] P. Di Francesco, O. Golinelli and E. Guitter, "Meander, folding and arch statistics", *Math. Comput. Modelling* **26** (1997), 97–147.

[11] P. Francesco, O. Golinelli and E. Guitter, "Meanders: a direct enumeration approach", *Nuc. Phys.* **B482** [FS] (1996), 497–535.

[12] P. Di Francesco, O. Golinelli and E. Guitter, "Meanders and the Temperley-Lieb algebra", *Commun. Math. Phys.* **186** (1997), 1–59.

[13] P. Di Francesco, "SU(N) Meander determinants", *J. Math. Phys.* **38** (1997), 5905–5943; "Meander determinants", *Commun. Math. Phys.* **191** (1998), 543–583.

[14] P. Di Francesco, "Truncated Meanders", preprint UNC-CH-MATH-98/3, to appear in *Proc. Amer. Math. Soc.*

[15] Y. Makeenko, "Strings, matrix models and meanders", *Nuclear Phys. B Proc. Suppl.* **49** (1995), 226–237; Y. Makeenko and H. Win Pe, "Supersymmetric matrix models and the meander problem", preprint ITEP-TH-13/95 (1996); G. Semenoff and R. Szabo, "Fermionic Matrix Models", *Internat. J. Modern Phys.* **A12**, (1997), 2135–2291.

[16] O. Golinelli, "A Monte-Carlo study of meanders", Eur. Phys. J. **B14** (2000), 145–155.

[17] I. Jensen, "Enumerations of plane meanders", preprint cond-mat/9910313. InstDi

[18] P. Di Francesco, O. Golinelli and E. Guitter, "Meanders: exact asymptotics", Nucl. Phys. **B570** (2000), 699–712.

[19] I. Kostov, Mod. Phys. Lett. **A4** (1989) 217; M. Gaudin and I. Kostov, "$O(n)$ model on a fluctuating planar lattice: Some exact results", *Phys. Lett.* **B220** (1989), 200–206; I. Kostov and M. Staudacher, "Multicritical phases of the $O(n)$ model on a random lattice", *Nucl. Phys.* **B384** (1992), 459–483.

[20] V. Kazakov, Nucl. Phys. **B4** (Proc. Suppl.) (1998), 93; J.-M. Daul, "Q-states Potts model on a random planar lattice", preprint hep-th/9502014; P. Zinn-Justin, "The dilute Potts model on random surfaces", preprint cond-mat/9903385; B. Eynard and G. Bonnet, "The Potts-q random matrix model: loop equations, critical exponents, and rational case", Phys. Lett. **B463** (1999), 273–279.

[21] P. Francesco and C. Itzykson, "A generating function for fatgraphs", *Ann. . Henri Poincaré*, **59**:2 (1993), Nienhuis117–139; V. Kazakov, M. Staudacher and T. Wynter, "Character expansion methods and for matrix models of dually weighted graphs",

Comm. Math. Phys. **177**, (1996) 451–468, "Almost flat planar diagrams", *Comm. Math. Phys.* **179** (1996), 235–256; Exact solution of discrete two-dimensional R^2 gravity, *Nucl. Phys.* **B471** (1996), 309–333.

[22] B. Eynard and J. Zinn-Justin, "The $O(n)$ model on a random surface: critical points and large-order behaviour", *Nucl. Phys.* **B386** (1992), 558; B. Eynard and C. Kristjansen, "Exact solution of the $O(n)$ model on a random lattice", *Nucl. Phys.* **B455** (1995), 577–618; "More on the exact solution of the $O(n)$ model on a random lattice and an investigation of the case $|n| > 2$", *Nucl. Phys.* **B466** (1996), 463–487.

[23] P. Di Francesco, P. Mathieu and D. Sénéchal, *Conformal field theory*, Graduate Texts in Contemporary Physics, Springer, New York, 1996.

[24] B. in *Phase Transitions and Critical Phenomena*, Eynardvol. **11**, edited by C. Domb and J. L. Lebowitz, Academic Press, 1987.

[25] L. Chekhov and C. Kristjansen, "Hermitian matrix model with plaquette interaction", *Nucl. Phys.* **B479** (1996), 683–696.

[26] J. Jacobsen and J. Kondev, "Field theory of compact polymers on the square lattice", *Nucl. Phys.* **B 532** [FS], (1998), 635–688; "Transition from the compact to the dense phase of two-dimensional polymers", *J. Stat. Phys.* **96**, (1999), 21–48.

[27] F. David and B. Duplantier, "Exact partition functions and correlation functions of multiple Hamiltonian walks on the Manhattan lattice", *J. Stat. Phys.* **51**, (1988), 327–434.

[28] P. Di Francesco and E. Guitter, *Europhys. Lett.* **26** (1994), 455.

[29] P. Di Francesco, B. and E. Guitter, "Coloring random triangulations", *Nucl. Phys.* **B516** [FS] (1998), 543–587.

[30] R. J. Baxter, "Colorings of a hexagonal lattice", *J. Math. Phys.* **11** (1970), 784–789; "q colourings of the triangular lattice", *J. Phys.* **A19** (1986), 2821–2839.

[31] P. Di Francesco, "Folding the square-diagonal lattice", *Nucl. Phys.* **B525** [FS] (1998), 507–548; "Folding transitions of the square-diagonal lattice", *Nucl. Phys.* **B528** [FS] (1998), 453–465; P. Di Francesco and E. Guitter, "Folding transition of the triangular lattice", *Phys. Rev.* **E50** (1994), 4418; M. Bowick, P. Di Francesco, O. Golinelli and E. Guitter, "3D folding of the triangular lattice", *Nucl. Phys.* **B450** [FS] (1995), 463–494.

[32] P. Di Francesco, "Folding and coloring problems in mathematics and physics", Bull. Amer. Math. Soc. **37**:3 (2000), 251–307.

[33] *Mathematical Intelligencer* **19**:4 (1997), 48; **20**:3 (1998), 29.

[34] I. Krichever, O. Lipan, P. Wiegmann and A. Zabrodin, "Quantum integrable systems and elliptic solutions of classical discrete nonlinear equations", *Comm. Math. Phys.* **188** (1997), 267–304.

[35] C. Itzykson and J.-B. Zuber, "The planar approximation II", *J. Math. Phys.* **21** (1980), 411–421.

[36] Harish-Chandra, "Differential operators on a semi-simple Lie algebra", *Amer. Jour. of Math.* **79** (1957), 87–120; J. Duistermaat and G. Heckman, "On the variation of cohomology of the symplectic form of the reduced phase space", *Inv. Math.* **69** (1982), 259–268.

[37] W. Tutte, "A census of planar maps", Canad. Jour. of Math. **15** (1963) 249–271.

Philippe Di Francesco
CEA-Saclay
Service de Physique Théorique
F-91191 Gif sur Yvette Cedex
France
philippe@spht.saclay.cea.fr