On some integrals over the $U(N)$ unitary group and their large N limit

2003 J. Phys. A: Math. Gen. 36 3173

(http://iopscience.iop.org/0305-4470/36/12/318)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 3173–3193 PII: S0305-4470(03)53470-9

On some integrals over the $U(N)$ unitary group and their large N limit

P Zinn-Justin[1] **and J-B Zuber**[2]

[1] Laboratoire de Physique Théorique et Modèles Statistiques, Université Paris-Sud, Bâtiment 100, F-91405 Orsay Cedex, France
[2] Service de Physique Théorique de Saclay, CEA/DSM/SPhT, Unité de recherche associée au CNRS, CEA-Saclay, F-91191 Gif sur Yvette Cedex, France

Received 12 December 2002
Published 12 March 2003
Online at stacks.iop.org/JPhysA/36/3173

Abstract
The integral over the $U(N)$ unitary group $I = \int \mathrm{D}U \exp \operatorname{Tr} AUBU^\dagger$ is re-examined. Various approaches and extensions are first reviewed. The second half of the paper deals with more recent developments: relation with integrable Toda lattice hierarchy, diagrammatic expansion and combinatorics, and what they teach us on the large N limit of $\log I$.

PACS numbers: 02.20.−a, 02.10.−v, 05.90.+m

1. Introduction and notation

1.1. Aim and plan of the paper

The techniques of integration over large matrices are important in several contexts of physics—from QCD [1] to quantum gravity [2], from disordered systems to mesoscopic physics [3]—and of mathematics—enumerative combinatorics [4], integrable systems [5], free probability theory [6], statistics [7], etc. In most applications, a central role is played by the statistics of the eigenvalues of the random matrices, spectrum and correlations, or more generally of their invariants $\operatorname{Tr} A^p$. This is justified once 'angular' variables have been integrated over, such as those appearing in $\operatorname{Tr} AB$.

It is in this general context that the integral

$$I = \int_{U(N)} \mathrm{D}U \exp \operatorname{Tr} AUBU^\dagger$$

where A and B are Hermitian, was studied more than 20 years ago, and an exact expression was derived [8]. Soon after, it was realized that this result had been obtained long before by Harish-Chandra [9] as a corollary of a more general problem. The purpose of this paper is to return to this integral (sometimes called the Harish-Chandra–Itzykson–Zuber integral in the physics literature), to review the known facts and to present some of its known extensions, before turning to what remains a challenge: to find a good and systematic description of the large N limit of its *logarithm*. We shall report on some recent progress made on the latter issue.

Our paper is organized as follows. In the rest of section 1, we introduce the notation and some basic results. The derivation of the expression of I is then reviewed in section 2 using various methods: heat equation, character expansion and Duistermaat–Heckman theorem. Section 3 briefly discusses the extensions in various directions, in particular the case of rectangular matrices. The connection with integrable hierarchies is then presented in section 4, with special attention devoted to its dispersionless limit and what can be learnt from it. The resulting expressions for the large N limit of $\log I$ are then confronted in section 5 with those obtained by (what we believe to be) a novel diagrammatic expansion of $\log I$ and by a purely combinatorial analysis [10]. Finally section 6 contains a summary of results and tables.

1.2. Notation

Let A and B be two $N \times N$ matrices. We shall assume that they are Hermitian, even though many properties that we shall derive do not require it. The subject of study is the integral

$$I(A, B; s) = \int_{U(N)} \mathrm{D}U \exp\left(\frac{N}{s} \mathrm{Tr}\, AUBU^\dagger\right) \tag{1.1}$$

where $\mathrm{D}U$ is the Haar measure on the unitary group $U(N)$, normalized to $\int \mathrm{D}U = 1$; and the large N limit (in a sense defined below) of its logarithm $\frac{1}{N^2} \log I(A, B; s)$. At the possible price of a redefinition of U, one may always assume that A and B are diagonal, $A = \mathrm{diag}(a_i)_{1\leqslant i\leqslant N}$ and $B = \mathrm{diag}(b_i)_{1\leqslant i\leqslant N}$. Clearly, the real parameter s could be scaled away. We find it convenient to keep it as an indicator of the homogeneity in the as and bs.

To any order of their $1/s$ expansions, both $I(A, B; s)$ and its logarithm are completely symmetric polynomials in the eigenvalues a_i and in the b_i independently, as we shall see in the next section. We want to express them in terms of the elementary symmetric functions of the as and of the bs:

$$\theta_p := \frac{1}{N}\sum_{i=1}^N a_i^p \qquad \bar\theta_p := \frac{1}{N}\sum_{i=1}^N b_i^p. \tag{1.2}$$

For finite N, only a finite number of these functions are independent, but this constraint disappears as $N \to \infty$. By large N limit we therefore mean that we consider sequences of matrices A and B of size N such that the moments of their spectral distributions θ_p and $\bar\theta_p$ converge as $N \to \infty$ (see [11] for some rigorous results on this limit). By abuse of notation, we still denote such objects by A and B, and write

$$F(A, B; s) = \lim_{N\to\infty} \frac{1}{N^2} \log I(A, B; s). \tag{1.3}$$

We shall use various combinations of the symmetric functions θ_p and $\bar\theta_p$. In particular, for $\alpha \vdash n$, i.e. α a partition of $n = \sum_p p\alpha_p$, which we also write as $\alpha = [1^{\alpha_1} \cdots n^{\alpha_n}]$, we define

$$\mathrm{tr}_\alpha A := \left(\frac{1}{N}\mathrm{Tr}\, A\right)^{\alpha_1} \cdots \left(\frac{1}{N}\mathrm{Tr}\, A^n\right)^{\alpha_n} = \prod_{p=1}^n \theta_p^{\alpha_p}. \tag{1.4}$$

We shall also make use of the characters of the irreducible holomorphic representations of the linear group $GL(N)$, labelled by Young diagrams λ with rows of length $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_N$. They read

$$\chi_\lambda = \frac{\Delta_\lambda(a)}{\Delta(a)} \tag{1.5}$$

in terms of Vandermonde determinants

$$\Delta(a) = \prod_{1\leqslant j<i\leqslant N} (a_i - a_j) = \det\big(a_i^{j-1}\big) \tag{1.6}$$

and their generalizations

$$\Delta_\lambda(a) = \det\big(a_i^{\lambda_j+j-1}\big). \tag{1.7}$$

Frobenius formula [12] relates these sets of symmetric functions: if $\alpha \vdash n$

$$N^{\Sigma\alpha_j} \operatorname{tr}_\alpha A = \sum_{\substack{\lambda\\ |\lambda|=n}} \chi_\lambda(A)\hat{\chi}_\lambda(\alpha) \tag{1.8}$$

where the sum runs over all Young tableaux λ with $|\lambda| = n$ boxes, and $\hat{\chi}_\lambda(\alpha)$ denotes the character of the symmetric group $\mathfrak{S}_n$, for the representation labelled by λ and for the class labelled by α.

The expression of integrals and differential operators over Hermitian matrices in terms of their eigenvalues involves a Jacobian, and problems of normalization appear. We discuss these questions shortly.

If $M = UAU^\dagger$, A diagonal, U unitary, we have $\mathrm{d}M = U\,\mathrm{d}AU^\dagger + [\mathrm{d}X, M]$, where $\mathrm{d}X := \mathrm{d}UU^\dagger$ is anti-Hermitian. Then $\mathrm{Tr}(\mathrm{d}M)^2 = \sum_i \mathrm{d}a_i^2 + 2\sum_{i<j} |\mathrm{d}X_{ij}|^2|a_i - a_j|^2$ defines the metric tensor $g_{\alpha\beta}$ in the coordinates $\xi^\alpha = (a_i, X_{ij})$. This determines first the measure $\mathrm{D}M = \sqrt{\det g}\prod \mathrm{d}\xi^\alpha = 2^{N(N-1)/2}\prod_i \mathrm{d}M_{ii}\prod_{i<j}\mathrm{d}\,\mathrm{Re}\,M_{ij}\,\mathrm{d}\,\mathrm{Im}\,M_{ij} = 2^{N(N-1)/2}\Delta^2(a)\prod \mathrm{d}a_i \prod \mathrm{d}X_{ij} = C\Delta^2(a)\prod_i \mathrm{d}a_i\,\mathrm{D}U$. The constant C is fixed by computing the integral of a $U(N)$ invariant function of M in two different ways, for example a Gaussian

$$1 = \int \mathrm{D}M\,\frac{\mathrm{e}^{-\frac{1}{2}\operatorname{Tr} M^2}}{(2\pi)^{N^2/2}} = \frac{C}{(2\pi)^{N(N-1)/2}}\int \prod_{i=1}^N \frac{\mathrm{d}a_i}{(2\pi)^{\frac{1}{2}}}\,\mathrm{e}^{-\frac{1}{2}a_i^2}\prod_{i<j}(a_i - a_j)^2.$$

The latter integral equals $\prod_{p=1}^N p!$ thus

$$C = \frac{(2\pi)^{N(N-1)/2}}{\prod_{p=1}^N p!} \qquad \mathrm{D}M = \frac{(2\pi)^{N(N-1)/2}}{\prod_{p=1}^N p!}\Delta^2(a)\prod_{i=1}^N \mathrm{d}a_i\,\mathrm{D}U. \tag{1.9}$$

From the above metric, one also computes the Laplacian

$$\begin{aligned}\Delta_M &= \frac{1}{\sqrt{g}}\frac{\partial}{\partial\xi^\alpha}g^{\alpha\beta}\sqrt{g}\frac{\partial}{\partial\xi^\beta} = \frac{1}{\prod_{i<j}(a_i - a_j)^2}\sum_k \frac{\partial}{\partial a_k}\prod_{i<j}(a_i - a_j)^2\frac{\partial}{\partial a_k} + \Delta_X\\ &= \Delta(a)^{-1}\sum_k\left(\frac{\partial}{\partial a_k}\right)^2\Delta(a) + \Delta_X \end{aligned}\tag{1.10}$$

where the last equality results from the vanishing of $\sum_k\left[\partial_{a_k},\left[\partial_{a_k},\Delta(a)\right]\right] = 0$, as this completely antisymmetric function of the as is a polynomial of degree $N(N-1)/2 - 2$.

2. The exact expression of $I(A, B; s)$

Assume that all the eigenvalues of A and B are distinct. One finds [8, 9]

$$I(A, B; s) = \left(\prod_{p=1}^{N-1} p!\right) (N/s)^{-N(N-1)/2} \frac{\det\left(\mathrm{e}^{\frac{N}{s}a_i b_j}\right)_{1\leqslant i,j\leqslant N}}{\Delta(a)\Delta(b)}. \tag{2.1}$$

Note that both the numerator and the denominator of the rhs are completely antisymmetric functions of the as and of the bs independently, and that the limit where some eigenvalues coalesce is well defined.

This expression (2.1) may be obtained by several different routes.

2.1. Heat equation

For two Hermitian $N \times N$ matrices M_A and M_B, let us consider

$$K(M_A, M_B; s) = \left(\frac{N}{2\pi s}\right)^{N^2/2} \exp\left(-\frac{N}{2s}\operatorname{Tr}(M_A - M_B)^2\right). \tag{2.2}$$

$K(M_A, M_B; s)$ satisfies the heat equation

$$\left(\frac{\partial}{\partial s} - \frac{1}{2}\Delta_{M_A}\right) K(M_A, M_B; s) = 0 \tag{2.3}$$

where Δ_M is the Laplacian over M, together with the boundary condition that for $t \to 0$, $K(M_A, M_B; s) \to \delta(M_A - M_B)$.

The heat kernel $K(M_A, M_B; s)$ is invariant under the simultaneous adjoint action on M_A and M_B by the same unitary matrix U. If we diagonalize $M_A = U_A A U_A^\dagger$ and $M_B = U_B B U_B^\dagger$, $K(M_A, M_B; s) = K(A, UBU^\dagger; s)$, where $U = U_A^\dagger U_B$. Upon integration

$$\begin{aligned}\widetilde{K}(A, B; s) &:= \int \mathrm{D}U\, K(M_A, U M_B U^\dagger; s) = \int \mathrm{D}U\, K(A, UBU^\dagger; s) \\ &= \left(\frac{N}{2\pi s}\right)^{N^2/2} \mathrm{e}^{-\frac{N}{2s}\operatorname{Tr}(A^2+B^2)} I(A, B; s)\end{aligned} \tag{2.4}$$

is again a solution of the heat equation (2.3), but depends only on the eigenvalues a_i of A (and b_i of B). Using the explicit form of the Laplacian in terms of the eigenvalues a_i (see (1.10)), $\widetilde{K}$ satisfies

$$\left(\frac{\partial}{\partial s} - \frac{1}{2}\sum_k \left(\frac{\partial}{\partial a_k}\right)^2\right) \Delta(a)\widetilde{K}(M_A, M_B; s) = 0. \tag{2.5}$$

The product $\Delta(a)\Delta(b)\widetilde{K}(A, B; s)$, an antisymmetric function of the as and of the bs, is a solution of the heat equation with the flat Laplacian, and satisfies the boundary conditions that for $s \to 0$, $C\Delta(a)\Delta(b)\widetilde{K}(A, B; s) \to \frac{1}{N!}\sum_{P\in\mathfrak{S}_N} \epsilon_P \prod_i \delta(a_i - b_{Pi})$, with C being the constant computed in (1.9). In physical terms it is the Green function of N independent free fermions, and it is thus given by the Slater determinant

$$C\Delta(a)\Delta(b)\widetilde{K}(A, B; s) = \frac{1}{N!}\left(\frac{N}{2\pi s}\right)^{N/2} \det\left[\exp\left(-\frac{N}{2s}(a_i - b_j)^2\right)\right] \tag{2.6}$$

which is consistent with (2.1).

2.2. Character expansion

We may expand the exponential in (1.1) to get

$$I(A,B;s)=\sum_{n=0}^{\infty}\frac{(N/s)^n}{n!}\int \mathrm{D}U(\mathrm{Tr}\,AUBU^\dagger)^n. \tag{2.7}$$

Frobenius formula (1.8) evaluated for the partition $[1^n]$ gives for any matrix X of $GL(N)$

$$\mathrm{Tr}^n X=\sum_{\substack{\lambda\\|\lambda|=n}}\hat{d}_\lambda\chi_\lambda(X) \tag{2.8}$$

where $\hat{d}_\lambda=\hat{\chi}_\lambda([1^n])$ is the dimension of the λ-representation of $\mathfrak{S}_n$. Integration over the unitary group then yields

$$\int \mathrm{D}U\,\chi_\lambda(AUBU^\dagger)=\frac{\chi_\lambda(A)\chi_\lambda(B)}{\chi_\lambda(I)}=\frac{\chi_\lambda(A)\chi_\lambda(B)}{d_\lambda} \tag{2.9}$$

and a well-known formula [12] gives

$$\frac{\hat{d}_\lambda}{d_\lambda}=n!\prod_{p=1}^{N}\frac{(p-1)!}{(\lambda_p+p-1)!}. \tag{2.10}$$

Using (1.5) and putting everything together, we find

$$\Delta(a)\Delta(b)I(A,B;s)=\left(\prod_{p=1}^{N-1}p!\right)\sum_{n=0}^{\infty}(N/s)^n\sum_{\substack{\lambda\\|\lambda|=n}}\frac{1}{\prod_p(\lambda_p+p-1)!}\Delta_\lambda(a)\Delta_\lambda(b) \tag{2.11a}$$

$$=\left(\prod_{p=1}^{N-1}p!\right)\sum_{0\leqslant\ell_1<\cdots<\ell_N}\prod_{q=1}^{N}\frac{(N/s)^{\ell_q-q+1}}{(\ell_q)!}\det\left(a_i^{\ell_j}\right)\det\left(b_i^{\ell_j}\right). \tag{2.11b}$$

Equation (2.11a) is interesting on its own, while an extension of the Binet–Cauchy theorem[3] enables one to resum (2.11b) into

$$\Delta(a)\Delta(b)I(A,B;s)=\left(\prod_{p=1}^{N-1}p!\right)(N/s)^{-N(N-1)/2}\det\left(\mathrm{e}^{\frac{N}{s}a_ib_j}\right) \tag{2.12}$$

which is precisely (2.1).

2.3. Duistermaat–Heckman theorem

Let us compute the integral (1.1) by the stationary phase method. The stationary points U_0 of the 'action' $\mathrm{Tr}\,AUBU^\dagger$ satisfy $\mathrm{Tr}\,\delta UU_0^\dagger[U_0BU_0^\dagger,A]=0$ for arbitrary anti-Hermitian $\delta UU_0^\dagger$,

[3] Recall [13] that if $f(x)=\sum_{\ell\geqslant 0}f_\ell x^\ell$,

$$\begin{aligned}\sum_{0\leqslant\ell_1<\ell_2<\cdots<\ell_N}f_{\ell_1}\cdots f_{\ell_N}\det a_i^{\ell_j}\det b_i^{\ell_j}&=\frac{1}{N!}\sum_{\ell_i\geqslant 0}\sum_{P,P'}\epsilon_P\epsilon_{P'}f_{\ell_1}\cdots f_{\ell_N}\prod_i a_{Pi}^{\ell_i}b_{P'i}^{\ell_i}\\&=\frac{1}{N!}\sum_{P,P'}\epsilon_{P\cdot P'}\prod_i\sum_{\ell_i}f_{\ell_i}(a_{P\cdot P'i}b_i)^{\ell_i}\\&=\det f(a_ib_j).\end{aligned}$$

hence $[U_0 B U_0^\dagger, A] = 0$. For diagonal matrices A and B with distinct eigenvalues, this implies that $U_0 B U_0^\dagger$ is diagonal and therefore that the saddle point U_0 are permutation matrices.

Gaussian fluctuations around the stationary point $U_0 = P$ may be computed by writing $U = \mathrm{e}^X P$, where X is anti-Hermitian, and by integrating over X after expanding the action to second order. Summing over all stationary points thus gives the 'one-loop approximation' to the integral (1.1):

$$
\begin{aligned}
I(A,B;s) &= C' \sum_{P\in\mathfrak{S}_N} \mathrm{e}^{\frac{N}{s}\sum_{i=1}^N a_i b_{Pi}} \int \prod_{i<j} \mathrm{d}^2 X_{ij} \exp\left(-\frac{N}{s}\sum_{i<j} |X_{ij}|^2 (a_i - a_j)(b_{Pi} - b_{Pj})\right) \\
&= C' \sum_{P\in\mathfrak{S}_N} \mathrm{e}^{\frac{N}{s}\sum_{i=1}^N a_i b_{Pi}} \frac{\left(\frac{\pi s}{N}\right)^{N(N-1)/2}}{\prod_{i<j}(a_i - a_j)(b_{Pi} - b_{Pj})} \\
&= C' \left(\frac{s\pi}{N}\right)^{N(N-1)/2} \frac{1}{\Delta(a)\Delta(b)} \sum_{P\in\mathfrak{S}_N} \epsilon_P\, \mathrm{e}^{\frac{N}{s}\sum_{i=1}^N a_i b_{Pi}} \\
&= C' \left(\frac{s\pi}{N}\right)^{N(N-1)/2} \frac{\det\left(\mathrm{e}^{\frac{N}{s} a_i b_j}\right)}{\Delta(a)\Delta(b)}
\end{aligned}
\qquad (2.13)
$$

which reproduces the previous result up to a constant C'. The latter may be determined, for example by considering the $s \to 0$ limit, and the result reproduces (1.1). Thus the stationary phase approximation of the original integral (1.1) or (2.4) turns out to give the exact result! This well-known empirical fact turned out to be a particular case of a general situation analysed later by Duistermaat and Heckman [14]: if a classical system has only periodic trajectories with the same period, the stationary phase (or saddle point) method is exact.

In more precise mathematical terms, let $\mathcal{M}$ be a $2n$-dimensional symplectic manifold with symplectic form ω, and suppose that it is invariant under a $U(1)$ action. Let H be the Hamiltonian corresponding to this action (i.e. $\mathrm{d}H = \mathrm{i}_V \omega$, where V is the vector field of infinitesimal $U(1)$ action). Assume also that the fixed point set (the critical points) is discrete. Then the theorem of Duistermaat–Heckman asserts that the stationary phase method is exact, i.e. that

$$
\int \frac{\omega^n}{n!}\,\mathrm{e}^{\mathrm{i}tH} = \left(\frac{2\pi}{t}\right)^n \sum_{P_c} \mathrm{e}^{\mathrm{i}\frac{\pi}{4}\mathrm{sign}(\mathrm{Hess}(P_c))}\, \mathrm{e}^{\mathrm{i}tH(P_c)} \frac{\sqrt{\det \omega(P_c)}}{\sqrt{|\det \mathrm{Hess}(P_c)|}} \qquad (2.14)
$$

where the sum is over (isolated) critical points P_c; the phase involves the signature $\mathrm{sign}(\mathrm{Hess}(P_c))$, i.e. the number of positive minus the number of negative eigenvalues, of the Hessian matrix $\mathrm{Hess}_{ij} = \partial^2 H/\partial\xi^i \partial\xi^j|_{P_c}$.

The integral (1.1) satisfies the conditions of the above theorem. The integration runs over the orbit $\mathcal{O} = \{M = UBU^\dagger\}$ of B under the coadjoint action of U: this orbit, homeomorphic to the manifold $U(N)/U(1)^N$, has the even dimension $N(N-1)$ and is in fact a symplectic manifold. On two tangent vectors $V_i = [X_i, M]$, $i = 1, 2$ (X_i anti-Hermitian), tangent to $\mathcal{O}$ on M, the symplectic form reads $\omega(V_1, V_2) = \mathrm{Tr}\, M[X_1, X_2]$. The Hamiltonian $H = \mathrm{Tr}\, AM$ defines a periodic flow $M(t) = \mathrm{e}^{\mathrm{i}At} M(0)\, \mathrm{e}^{-\mathrm{i}At}$ if all eigenvalues of A are relatively rational. As the latter configurations form a dense set among diagonal matrices A, this constraint may in fact be removed and this justifies *a posteriori* the above stationary phase calculation. For further details on the Duistermaat–Heckman theorem in the present context, the reader may also consult [15, 16].

3. Generalizations and extensions

3.1. Other groups

As already mentioned, the integral (2.1) appeared first in the work of Harish-Chandra [9], as a simple application to compact groups of a general discussion of invariant differential operators on Lie algebras. Following Harish-Chandra, let G be a compact Lie group. Denote by Ad the adjoint action of G on itself and on its Lie algebra $\mathfrak{g}$, by $\langle\cdots\rangle$ the invariant bilinear form on $\mathfrak{g}$, and by $\mathfrak{h} \subset \mathfrak{g}$ the Cartan subalgebra. If h_1 and $h_2 \in \mathfrak{h}$

$$\Delta(h_1)\Delta(h_2)\int_G \mathrm{d}g \exp\langle \mathrm{Ad}(g)h_1, h_2\rangle = \mathrm{const}\sum_{w\in W} \epsilon(w)\exp\langle w(h_1), h_2\rangle \tag{3.1}$$

where w is summed over the Weyl group W, $\epsilon(w) = (-1)^{\ell(w)}$, $\ell(w)$ is the number of reflections generating w, and $\Delta(h) = \prod_{\alpha>0}\langle\alpha, h\rangle$, a product over the positive roots of $\mathfrak{g}$. In the case of $U(N)$, if we take $h_1 = \mathrm{i}A$ and $h_2 = \mathrm{i}B$, (3.1) reduces to (2.1).

This extension to general compact G may also be derived by constructing the heat kernel $K(g_1, \mathrm{Ad}(g)g_2; s)$ on G in terms of characters, by using Weyl's formulae for characters and by averaging K over G [17]. Let us sketch the derivation. One may always assume that $g_j = \mathrm{e}^{h_j}$, $h_j \in \mathfrak{h}$, $j = 1, 2$, and one writes

$$K(g_1, g_2; s) = \sum_\lambda d_\lambda \chi_\lambda\left(g_1 g_2^{-1}\right)\mathrm{e}^{-\frac{1}{2}sC_\lambda} \tag{3.2}$$

where C_λ is the value of the quadratic Casimir for the representation of weight λ and d_λ is the dimension of the latter. Integration over the adjoint action gives

$$\int \mathrm{d}g K(g_1, \mathrm{Ad}(g)g_2; s) = \sum_\lambda \chi_\lambda(g_1)\chi_\lambda\left(g_2^{-1}\right)\mathrm{e}^{-\frac{1}{2}sC_\lambda}. \tag{3.3}$$

Weyl's formula for the characters reads

$$\chi_\lambda(\mathrm{e}^{\mathrm{i}h}) = \frac{\sum_{w\in W}\epsilon_w\,\mathrm{e}^{\mathrm{i}\langle\lambda+\rho, w(h)\rangle}}{\text{same for } \lambda = 0} \tag{3.4}$$

where ρ is the Weyl vector, sum of all fundamental weights. As $C_\lambda = |\lambda+\rho|^2 - |\rho|^2$, the rhs of (3.3) is the exponential of a quadratic form in λ. The summation over λ may be extended from the Weyl chamber to the full weight lattice, Poisson summation formula is then used and after taking a limit of infinitesimal s, h_1 and h_2, one is led to (3.1).

For completeness we also mention the generalization of (1.1) involving Grassmannian coordinates and the integration over the supergroup $U(N_1|N_2)$ (see [18, 19]). The case of integration over a pseudounitary group $U(N_1, N_2)$ has also been discussed (see [20] and further references therein).

3.2. Rectangular matrices

Consider the integral

$$I^{(2)}(A, B; s) = \int_{U(N_2)} \mathrm{D}U \int_{U(N_1)} \mathrm{D}V \exp\left(\frac{N}{s}\,\mathrm{Tr}(AUBV^\dagger + \mathrm{h.c.})\right) \tag{3.5}$$

where A is a complex $N_1 \times N_2$ matrix, B a complex $N_2 \times N_1$ matrix and $N = \min(N_1, N_2)$. Without loss of generality, one may assume that $N_1 \geqslant N_2$ and that the $N_2 \times N_2$ matrices $A^\dagger A$ and $BB^\dagger$ are diagonal, with real non-negative eigenvalues a_i and b_i, respectively. We assume again that the as on the one hand, the bs on the other, are all distinct, so that neither of the Vandermonde $N_2 \times N_2$ determinants, $\Delta(a)$ or $\Delta(b)$, vanishes. Then the methods of heat equation or of character expansion presented in section 2 yield the following expression:

$$\Delta(a)\Delta(b)I^{(2)}(A,B;s) = \frac{\prod_{p=1}^{N_2-1} p!\prod_{q=1}^{N_1-1} q!}{\prod_{r=1}^{N_1-N_2-1} r!}(s/N)^{N_2(N_1-1)}\frac{\det I_{N_1-N_2}(2N\sqrt{a_i b_j}/s)}{\prod_{i=1}^{N_2}(a_i b_i)^{\frac{1}{2}(N_1-N_2)}} \tag{3.6}$$

with $I_\nu(z)$ being the Bessel function $I_\nu(z) = \sum_{n=0}^{\infty}\frac{1}{n!(n+\nu)!}\left(\frac{z}{2}\right)^{2n+\nu}$. This expression has been obtained for $N_1 = N_2$ in [13, 19] and in the general case in [21].

Integrals (1.1) and (3.5) are the cases $K = 1, 2$ of an infinite set of unitary integrals which are exactly calculable:

$$\begin{aligned} I^{(K)}(A_k, B_k; s) &= \int_{U_k\in U(N_k)} \prod_{k=1}^{K} \mathrm{D}U_k \exp\frac{N}{s}\sum_{k=1}^{K}\mathrm{Tr}\,A_k U_{k+1} B_k U_k^\dagger \\ &= \mathrm{const}\frac{\det\phi(a_i b_j (N/s)^K)}{\Delta(a)\Delta(b)} \end{aligned} \tag{3.7}$$

(index k is cyclic modulo K) where $N = \min_k(N_k)$, and A_k and $B_k^\dagger$ are $N_k \times N_{k+1}$ matrices, $1 \leqslant k \leqslant K$; the generalized hypergeometric series ϕ is given by $\phi(x) = \sum_{n=0}^{\infty} x^n \Big/ \prod_{k=1}^{K}(n + N_k - N)!$; and, assuming $N_1 = N$, the a_i (resp. b_i) are the N (distinct) eigenvalues of $A_1 A_2 \cdots A_K$ (resp. $B_K \cdots B_2 B_1$).

In what follows, we shall concentrate on integrals (1.1) and (3.5); however, the analysis applies equally well to the more general integral (3.7).

3.3. Correlation functions

Returning to the unitary group and the integral (1.1), it is also natural to consider the correlation functions associated with it, i.e. to compute the integrals

$$\int_{U(N)} \mathrm{D}U\, U_{i_1 j_1}\cdots U_{i_m j_m} {U^\dagger}_{k_1\ell_1}\cdots {U^\dagger}_{k_m\ell_m} \exp\left(\frac{N}{s}\mathrm{Tr}\,AUBU^\dagger\right). \tag{3.8}$$

Partial results have been obtained in [22] and in [23]. We still lack explicit and general expressions for these correlation functions.

4. Connection with integrable hierarchies

The integral (1.1) turns out to provide a non-trivial solution of the two-dimensional Toda lattice hierarchy [24]. This stems from the following observation: define

$$\tau_N = \frac{\det\left(\mathrm{e}^{\frac{1}{\hbar}a_i b_j}\right)_{1\leqslant i,j\leqslant N}}{\Delta(a)\Delta(b)} \tag{4.1}$$

where $\hbar = s/N$. Comparing with equation (2.1), we see that $\tau_N = \hbar^{-N(N-1)/2}\prod_{p=0}^{N-1}(p!)^{-1} I$. Then the following formula holds: (see also [24, 25] for a two-matrix-model formulation)

$$\tau_N = \det\left(\oint\oint \frac{\mathrm{d}u}{2\pi\mathrm{i}u}\frac{\mathrm{d}v}{2\pi\mathrm{i}v}u^j v^i \exp\left(\frac{1}{\hbar}\left(\sum_{q\geqslant 1} t_q u^q + \sum_{q\geqslant 1}\bar{t}_q v^q + u^{-1}v^{-1}\right)\right)\right)_{0\leqslant i,j\leqslant N-1} \tag{4.2}$$

where the integration contours are small enough circles around the origin, and with the traditional notation:

$$t_q = \hbar\frac{1}{q}\sum_{i=1}^{N} a_i^q = \frac{s}{q}\theta_q \qquad \bar{t}_q = \hbar\frac{1}{q}\sum_{i=1}^{N} b_i^q = \frac{s}{q}\bar{\theta}_q \qquad q \geqslant 1. \tag{4.3}$$

Formula (4.2) can be easily proved by noting that $\mathrm{e}^{\frac{1}{\hbar}\sum_q t_q u^q} = \prod_{i=1}^N (1-ua_i)^{-1}$, $\mathrm{e}^{\frac{1}{\hbar}\sum_q \bar{t}_q v^q} = \prod_{i=1}^N (1-vb_i)^{-1}$, and expanding the contours to catch the poles at $u = a_i^{-1}$, $v = b_i^{-1}$. This makes τ_N a tau function of the 2D Toda lattice hierarchy, as we shall discuss now.

4.1. Biorthogonal polynomials and 2D Toda lattice hierarchy

Noting that the parameter $\hbar = s/N$ can always be scaled away we set $\hbar = 1$ throughout this section. We take equation (4.2) to be the definition of τ_N as a function of the two infinite sets of times $(t_q, \bar{t}_q)$, $q \geqslant 1$. We also set $\tau_0 = 1$.

Formula (4.2) suggests introducing a non-degenerate bilinear form on the space of polynomials by

$$\langle q|p\rangle = \oint\oint \frac{\mathrm{d}u}{2\pi\mathrm{i}u}\frac{\mathrm{d}v}{2\pi\mathrm{i}v} p(u)q(v)\exp\left(\sum_{q\geqslant 1} t_q u^q + \sum_{q\geqslant 1}\bar{t}_q v^q + u^{-1}v^{-1}\right) \tag{4.4}$$

and normalized biorthogonal polynomials $q_n(v)$ and $p_n(u)$ with respect to the above bilinear form, that is polynomials of the form $p_n(u) = h_n^{-1}u^n + \cdots$ and $q_n(v) = h_n^{-1}v^n + \cdots$, such that $\langle q_m|p_n\rangle = \delta_{mn}$ for all $m, n \geqslant 1$. One can now replace monomials in equation (4.2) with biorthogonal polynomials and obtain immediately

$$\tau_N = \prod_{i=0}^{N-1} h_i^2. \tag{4.5}$$

Next we introduce the matrices of multiplication by u and v in the basis of biorthogonal polynomials:

$$L_{mn} = \langle q_m|u|p_n\rangle \qquad \bar{L}_{mn} = \langle q_m|v|p_n\rangle \qquad m, n \geqslant 1. \tag{4.6}$$

Note that

$$L_{mn} = 0 \quad m > n+1 \qquad \bar{L}_{mn} = 0 \quad n > m+1. \tag{4.7}$$

A standard calculation [24] leads to the following evolution equations for L and $\bar{L}$ with respect to variations of the $t_q, \bar{t}_q$:

$$\frac{\partial L}{\partial t_q} = -[(L^q)_+, L] \qquad \frac{\partial L}{\partial \bar{t}_q} = [(\bar{L}^q)_-, L] \tag{4.8a}$$

$$\frac{\partial \bar{L}}{\partial t_q} = -[(L^q)_+, \bar{L}] \qquad \frac{\partial \bar{L}}{\partial \bar{t}_q} = [(\bar{L}^q)_-, \bar{L}]. \tag{4.8b}$$

Here $(\cdot)_\pm$ denotes the lower/upper triangular part plus one half of the diagonal part. Equations (4.8) are the standard form of the two-dimensional Toda lattice hierarchy [26] (up to a choice of sign of the t_q).

It is often more convenient to write these equations as quadratic equations in the set of (τ_N). These are the 'bilinear' Hirota equations. They are obtained by picking two sets of times $(x_q, \bar{x}_q)$ and $(y_q, \bar{y}_q)$ and writing in two different ways

$$\oint\oint \frac{\mathrm{d}u}{2\pi\mathrm{i}u}\frac{\mathrm{d}v}{2\pi\mathrm{i}v} p_{n,x,\bar{x}}(u)q_{m,y,\bar{y}}(v)\exp\left(\sum_{q\geqslant 1} x_q u^q + \sum_{q\geqslant 1}\bar{y}_q v^q + u^{-1}v^{-1}\right)$$

where $p_{n,x,\bar{x}}$ is the right biorthogonal polynomial associated with times $(x_q, \bar{x}_q)$, whereas $q_{m,y,\bar{y}}$ is the left biorthogonal polynomial associated with times $(y_q, \bar{y}_q)$. One finds

$$\oint \frac{\mathrm{d}u}{2\pi \mathrm{i} u} u^{n-m} \tau_n \left(x_q - \frac{1}{q} u^{-q}, \bar{x}_q \right) \tau_{m+1} \left(y_q + \frac{1}{q} u^{-q}, \bar{y}_q \right) \mathrm{e}^{\sum_{q \geqslant 1} (x_q - y_q) u^q}$$
$$= \oint \frac{\mathrm{d}v}{2\pi \mathrm{i} v} v^{m-n} \tau_m \left(y_q, \bar{y}_q - \frac{1}{q} v^{-q} \right) \tau_{n+1} \left(x_q, \bar{x}_q + \frac{1}{q} v^{-q} \right) \mathrm{e}^{\sum_{q \geqslant 1} (\bar{y}_q - \bar{x}_q) v^q}. \tag{4.9}$$

Expanding in powers of $x_q - y_q$ and $\bar{x}_q - \bar{y}_q$ results in an infinite set of partial differential equations satisfied by the (τ_N).

Example. Expand to first order in $x_1 - y_1$, and set $m = n + 1$. The result is

$$\tau_{n+1} \tau_{n-1} = \tau_n \partial \bar{\partial} \tau_n - \partial \tau_n \bar{\partial} \tau_n \qquad \forall n \geqslant 1 \tag{4.10}$$

(with $\partial = \partial / \partial t_1$, $\bar{\partial} = \partial / \partial \bar{t}_1$) which is a form of the Toda lattice equation. It is of course also the Desnanot–Jacobi determinant identity applied to equation (4.2).

Finally, the matrices L and $\bar{L}$ satisfy an additional relation, the so-called string equation, which takes the form [24]

$$[L^{-1}, \bar{L}^{-1}] = 1. \tag{4.11}$$

4.2. Large N limit as dispersionless limit

We now restore the parameter $\hbar$, which is required for the large N limit. Indeed, as $N \to \infty$, $\hbar$ must be sent to zero, keeping $s = \hbar N$ fixed. We define

$$f(t_q, \bar{t}_q; s) = \lim_{N \to \infty} \hbar^2 \log \tau_N(t_q, \bar{t}_q) \tag{4.12}$$

where $\tau_N(t_q, \bar{t}_q)$ is defined by equation (4.2). In a region of the space of parameters $(t_q, \bar{t}_q)$ which includes a neighbourhood of the origin $t_q = \bar{t}_q = 0$, $f(t_q, \bar{t}_q; s)$ is governed by 'dispersionless' equations we shall describe now. In this context, $f(t_q, \bar{t}_q; s)$ is called the dispersionless tau function.

First, comparing with expression (1.3) of the free energy $F(\theta_q s^{-q/2}, \bar{\theta}_q s^{-q/2})$ (where we have explicitly stated the dependence of F on the θ_q and the $\bar{\theta}_q$ and scaled away the parameter s), using equation (2.1) and definitions (4.3), we see that

$$f(t_q, \bar{t}_q; s) = -\tfrac{1}{2} s^2 \log s + \tfrac{3}{4} s^2 + s^2 F(q t_q s^{-q/2-1}, q \bar{t}_q s^{-q/2-1}). \tag{4.13}$$

This is the scaling form of the dispersionless tau function[4].

The dispersionless Toda hierarchy is a classical limit of the Toda hierarchy: it is obtained by replacing commutators with Poisson brackets defined by

$$\{g(z,s), h(z,s)\} = z \frac{\partial g}{\partial z} \frac{\partial h}{\partial s} - z \frac{\partial h}{\partial z} \frac{\partial g}{\partial s}. \tag{4.14}$$

Here z is the classical analogue of the shift operator $Z_{ij} = \delta_{ij+1}$, which justifies that $\{\log z, s\} = 1$.

The Lax equations (4.8) become

$$\frac{\partial \ell}{\partial t_q} = -\{(\ell^q)_+, \ell\} \qquad \frac{\partial \ell}{\partial \bar{t}_q} = \{(\bar{\ell}^q)_-, \ell\} \tag{4.15a}$$

[4] In fact, one can get rid of one more parameter since A and B can be scaled independently, but we choose not to do so for symmetry reasons.

$$\frac{\partial\bar{\ell}}{\partial t_q} = -\{(\ell^q)_+, \bar{\ell}\} \qquad \frac{\partial\bar{\ell}}{\partial\bar{t}_q} = \{(\bar{\ell}^q)_-, \bar{\ell}\} \tag{4.15b}$$

where $(\cdot)_\pm$ now refers to the positive and negative parts of the Laurent expansion in z, and $\ell(z, s, t)$ and $\bar{\ell}(z, s, t)$ have the following z dependence:

$$\ell = rz + \sum_{k=0}^{\infty} \lambda_k z^{-k} \tag{4.16a}$$

$$\bar{\ell} = rz^{-1} + \sum_{k=0}^{\infty} \bar{\lambda}_k z^k \tag{4.16b}$$

related to the structure (4.7) of L and $\bar{L}$.

These equations imply that the differential forms $\mathrm{d}\varphi(\ell, s, t)$ and $\mathrm{d}\bar{\varphi}(\bar{\ell}, s, t)$ are closed:

$$\mathrm{d}\varphi = m\, \mathrm{d}\ell/\ell + \log z\, \mathrm{d}s + \sum_{q\geqslant 1} (\ell^q)_+ \mathrm{d}t_q + \sum_{q\geqslant 1} (\bar{\ell}^q)_- \mathrm{d}\bar{t}_q \tag{4.17a}$$

$$\mathrm{d}\bar{\varphi} = \bar{m}\, \mathrm{d}\bar{\ell}/\bar{\ell} + \log z\, \mathrm{d}s + \sum_{q\geqslant 1} (\ell^q)_+ \mathrm{d}t_q + \sum_{q\geqslant 1} (\bar{\ell}^q)_- \mathrm{d}\bar{t}_q \tag{4.17b}$$

where m and $\bar{m}$ are Orlov–Shulman functions, which can be characterized by

$$m = \sum_{q\geqslant 1} q t_q \ell^q + s + \sum_{q\geqslant 1} \frac{\partial f}{\partial t_q} \ell^{-q} \tag{4.18a}$$

$$\bar{m} = \sum_{q\geqslant 1} q \bar{t}_q \bar{\ell}^q + s + \sum_{q\geqslant 1} \frac{\partial f}{\partial \bar{t}_q} \bar{\ell}^{-q} \tag{4.18b}$$

and satisfy the 'dressed' Poisson bracket relations

$$\{\log\ell, m\} = 1 \qquad \{\log\bar{\ell}, \bar{m}\} = 1. \tag{4.19}$$

In the present case, we have the following constraints, which determine uniquely the solution of the dispersionless Toda hierarchy:

$$m = \bar{m} = (\ell\bar{\ell})^{-1}. \tag{4.20}$$

Equations (4.20) are directly related (via equations (4.19)) to the classical limit of the string equation (4.11), i.e. $\{\ell^{-1}, \bar{\ell}^{-1}\} = 1$.

Let us denote $a = \ell^{-1}$ and $b = \bar{\ell}^{-1}$. The fact that $m = \bar{m}$ implies[5] that $\varphi(a)$ and $\bar{\varphi}(b)$ are related by Legendre transform, or that their derivatives $b(a) = \frac{\mathrm{d}}{\mathrm{d}a}\varphi(a)$ and $a(b) = \frac{\mathrm{d}}{\mathrm{d}b}\bar{\varphi}(b)$ are functional inverses of each other (the latter fact was derived by *ad hoc* methods in [28] and [29]).

4.3. Application of the dispersionless formalism

In what follows we set $s = 1$, so that $\theta_q = qt_q$. In order to explore the structure of the function $f(t_q, \bar{t}_q; s)$, we now assume that only a finite number of t_q and $\bar{t}_q$ are nonzero. Note that this cannot happen if the eigenvalues are real; however, here we are interested in the properties of the integral (1.1) as a formal power series in the t_q and therefore we do not worry about the

[5] The constraint $m = \bar{m}$ is true for a more general class of solutions of Toda, see, for example, section 4.4 and [27].

actual support of the eigenvalues. Then, according to equations (4.16) and (4.18), the Laurent expansion of $a(z)$ and $b(z)$ is finite, of the form:

$$a = \sum_{q=1}^{\bar n+1} \alpha_q z^{-q} \qquad b = \sum_{q=1}^{n+1} \beta_q z^q \tag{4.21}$$

where $n = \max\{q|t_q \neq 0\}$, $\bar n = \max\{q|\bar t_q \neq 0\}, \alpha_1 = \beta_1 = 1/r$, $\alpha_{\bar n+1} = \bar n \bar t_{\bar n} r^{\bar n+1}$ and $\beta_{n+1} = n t_n r^{n+1}$ (r is defined by equations (4.16)). Then, it is easy to show that a and b satisfy an algebraic equation, some coefficients of which can be worked out explicitly:

$$b^{\bar n+1}a^{n+1} - b^{\bar n}a^n - b^{\bar n}(a^{n-1}\theta_1 + \cdots + \theta_n) - a^n(b^{\bar n-1}\bar\theta_1 + \cdots + \bar\theta_{\bar n}) + \text{lower order terms} = 0. \tag{4.22}$$

Plugging equation (4.21) into equation (4.22) leads to algebraic equations for the coefficients α_q, β_q as a function of the $\theta_q, \bar\theta_q$.

Examples

- If $\bar\theta_q = \delta_{q1}\bar\theta_1$, one can set $\bar\theta_1 = 1$ by homogeneity. In this case equation (4.22) is quadratic in b. In terms of b and $\ell = 1/a$ it reads

$$b^2 - b(\ell + \theta_1\ell^2 + \cdots + \theta_n\ell^{n+1}) - \ell P(\ell) = 0 \tag{4.23}$$

where P is a polynomial with $P(0) = 1$. One way to determine P is to note that equation (4.21) provides a rational parametrization of a and b, so that the resulting curve (a, b) has genus zero, which forces the discriminant of equation (4.23) to be of the form $\ell(1 + \ell/(4\psi^2))Q(\ell)^2$, where Q is a polynomial of degree n with $Q(0) = 2$, and ψ is a constant. This yields the expression

$$b(\ell) = \frac{1}{2}\left(\ell + \theta_1\ell^2 + \cdots + \theta_n\ell^{n+1} + \sqrt{\ell\left(1 + \frac{\ell}{4\psi^2}\right)}Q(\ell)\right). \tag{4.24}$$

Considering equation (4.18a) as an asymptotic expansion of $m = b/\ell$ as $\ell \to \infty$ fixes Q: Q is the polynomial part of $2\psi(1 + \theta_1\ell + \cdots + \theta_n\ell^n)/\sqrt{1 + 4\psi^2/\ell}$. Imposing $Q(0) = 2$ leads to

$$\psi = 1 + \sum_{q=1}^{n}(-1)^{q+1}\frac{(2q)!}{(q!)^2}\theta_q\psi^{2q+1}. \tag{4.25}$$

Note that $\psi = r^2 = \partial^2F/\partial\theta_1\partial\bar\theta_1 = \exp(\partial^2F/\partial s^2)$ (the latter equality being the dispersionless limit of Toda equation (4.10)).

By Lagrange inversion, equation (4.25) allows us to get the exact expansion of ψ as well as of F in powers of the θ_q. One finds, restoring the $\bar\theta_1$ dependence:

$$\begin{aligned} \psi &= \sum_{\alpha_1,\ldots,\alpha_n=0}^{\infty} \frac{\left(\sum_{q\geqslant1}(2q+1)\alpha_q\right)!}{\left(\sum_{q\geqslant1}2q\alpha_q+1\right)!}\prod_{q\geqslant1}\frac{1}{\alpha_q!}\left((-1)^{q+1}\frac{(2q)!}{(q!)^2}\theta_q\right)^{\alpha_q}\bar\theta_1^{\sum_q q\alpha_q} \\ F &= \sum_{\alpha_1,\ldots,\alpha_n=0}^{\infty} \frac{\left(\sum_{q\geqslant1}(2q+1)\alpha_q-3\right)!}{\left(\sum_{q\geqslant1}2q\alpha_q\right)!}\prod_{q\geqslant1}\frac{1}{\alpha_q!}\left((-1)^{q+1}\frac{(2q)!}{(q!)^2}\theta_q\right)^{\alpha_q}\bar\theta_1^{\sum_q q\alpha_q}. \end{aligned} \tag{4.26}$$

- If $\theta_q = \delta_{qn}\theta_n$, $\bar\theta_q = \delta_{qn}\bar\theta_n$ (with the same n), one can show using the above formalism that $\psi = \partial^2F/\partial\theta_1\partial\bar\theta_1$ satisfies the following equation:

$$\psi = 1 + (n+1)\theta_n\bar\theta_n\psi^{n+2} \tag{4.27}$$

so that

$$\psi = \sum_{k\geqslant 1}(n+1)^k \frac{((n+2)k)!}{((n+1)k+1)!k!}(\theta_n\bar{\theta}_n)^k \qquad F = \sum_{k\geqslant 1}(n+1)^k \frac{((n+2)k-3)!}{((n+1)k)!k!}(\theta_n\bar{\theta}_n)^k. \tag{4.28}$$

For example, for $n = 2$, $F = \frac{1}{2}(\theta_2\bar{\theta}_2) + \frac{3}{4}(\theta_2\bar{\theta}_2)^2 + \frac{9}{2}(\theta_2\bar{\theta}_2)^3 + \cdots$.

4.4. Case of rectangular matrices

All of the above formalism applies equally well to the integral (3.5) over rectangular matrices. The latter has a 'diagonal' character expansion, i.e.

$$\tau_N^{(2)} = \text{const}\, I^{(2)}(A, B; s) = \sum_\lambda \hbar^{-|\lambda|} \frac{1}{\prod_p (\lambda_p + p - 1)!(\lambda_p + p - 1 + \nu)!} \frac{\Delta_\lambda(a)}{\Delta(a)} \frac{\Delta_\lambda(b)}{\Delta(b)} \tag{4.29}$$

where $\hbar = s/N$, $N = \min(N_1, N_2)$, $\nu = |N_1 - N_2|$.

The form of equation (4.29) alone proves that the $\left(\tau_N^{(2)}\right)$ form a tau function of 2D Toda lattice (for fixed ν). There is also a determinant formula similar to equation (4.2) which we shall not write.

The large N limit is taken keeping $\hbar N = s$ and $\frac{1}{2}\hbar\nu \equiv \xi$ fixed. The times are $t_q = \hbar\frac{1}{q}\operatorname{Tr}(AA^\dagger)^q$ and $\bar{t}_q = \hbar\frac{1}{q}\operatorname{Tr}(BB^\dagger)^q$. Biorthogonal polynomials and the dispersionless equations can be obtained similarly as above.

We have thus constructed a whole family $g_\xi(t_q, \bar{t}_q; s) = \lim_{N\to\infty} \hbar^2 \log \tau_N^{(2)}$ of dispersionless tau functions. The diagonal character expansion (4.29) implies that they belong to the class of solutions that satisfy $m = \bar{m}$. After some calculations, one finds that the constraint analogous to equation (4.20) is

$$m = \bar{m} = \sqrt{(\ell\bar{\ell})^{-1} + \xi^2} - \xi. \tag{4.30}$$

Note that if $\xi = 0$, i.e. $N_1 = N_2$, the constraint (4.30) becomes $m = \bar{m} = (\ell\bar{\ell})^{-1/2}$, which suggests redefining $\ell = \ell_2^2$ and $\bar{\ell} = \bar{\ell}_2^2$. The new functions ℓ_2 and $\bar{\ell}_2$ satisfy the string equation $\left\{\ell_2^{-1}, \bar{\ell}_2^{-1}\right\} = 1$ of the integral (1.1), and of course the same Toda equations (4.15) for even times, with the replacement $t_{2q} \to t_q$ and $\bar{t}_{2q} \to \bar{t}_q$. In fact, one can check that we recover in this case the dispersionless tau function of the usual integral (1.1) with odd moments equal to zero: $g_0(u_q, \bar{u}_q; s) = 2f(t_{2q} = u_q, t_{2q+1} = 0, \bar{t}_{2q} = \bar{u}_q, \bar{t}_{2q+1} = 0; s)$.

5. Diagrammatic expansion and combinatorics

5.1. A Feynman diagram expansion of F

To develop a Feynman diagram expansion of (1.1), we first trade the integration over the unitary group for an integration over $N \times N$ complex matrices X by writing

$$e^{\frac{N}{s}\operatorname{Tr} AUBU^\dagger} = \left(\frac{\pi}{N}\right)^{-N^2} \int DX\, DX^\dagger\, e^{-N \operatorname{Tr} XX^\dagger} \exp\left(\frac{N}{\sqrt{s}} \operatorname{Tr}(AUX^\dagger + XBU^\dagger)\right). \tag{5.1}$$

Thus,

$$I(A, B; s) = \text{const} \int DX\, DX^\dagger \exp(-N \operatorname{Tr}(XX^\dagger) + N^2 W(AXBX^\dagger; s)) \tag{5.2}$$

where

$$e^{N^2 W(X_1 X_2; s)} = \int DU \exp\left(\frac{N}{\sqrt{s}} \operatorname{Tr}(UX_1 + U^\dagger X_2)\right). \tag{5.3}$$

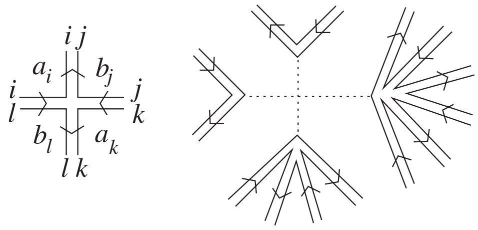


Figure 1. (*a*) A vertex of type [2]. (*b*) A multi-vertex of type $[1^2\,2^1\,3^1]$.

This integral is known exactly in this large N limit [30]: with the same abuse of notation as mentioned in section 1,

$$W(Y;s)=\sum_{n=1}^{\infty}s^{-n}\sum_{\alpha\vdash n}W_\alpha\frac{\mathrm{tr}_\alpha\,Y}{\prod_p(\alpha_p!\,p^{\alpha_p})} \tag{5.4a}$$

$$W_\alpha=(-1)^n\frac{\left(2n+\sum\alpha_p-3\right)!}{(2n)!}\prod_{p=1}^{n}\left(\frac{-(2p)!}{p!(p-1)!}\right)^{\alpha_p} \tag{5.4b}$$

and W_α is an integer, as follows from the recursion formulae discussed in [30].

The form (5.2) is adequate to develop a diagrammatic expansion of $F(A,B;s)$ in the style of Feynman (see [31, 32] for reviews). Recall that double lines are conveniently introduced to encode the conservation of indices [33]. The inverse of the first term in the exponential of (5.2) yields the propagator, $\langle X_{ij}X^\dagger_{kl}\rangle = {}_j^i\!\!\Longrightarrow\!\!{}_k^l = \frac{1}{N}\delta_{i\ell}\delta_{jk}$, while each term in W, i.e. each monomial $\mathrm{tr}_\alpha(AXBX^\dagger)\big/\prod_p(\alpha_p!\,p^{\alpha_p})$ gives rise to a multi-vertex of type α (see figure 1), which comes with a weight $N^{2-\sum\alpha_p}W_\alpha$ times products of as and bs and Kronecker symbols expressing the conservation of indices. When these Kronecker symbols are 'contracted', they leave sums over closed circuits (or faces) of powers of as or bs. A face of side p thus contributes a factor $N\theta_p$, resp. $N\bar\theta_q$ with the notation of (1.2). To put it another way, the Feynman diagrams may be regarded as bicolourable, with faces carrying alternating a or b 'colour'.

As is well known, only connected diagrams contribute to the free energy $\log I$. In the discussion of this connectivity, each multi-vertex just introduced must be regarded as a connected object, and it is useful to keep track of this fact by drawing a tree of dotted lines which connect the various traces which compose it, see figure 1(*b*).

If a diagram contributing to $\log I$ has P propagators, V_α vertices of type α and L loops of indices, and builds a surface of total genus g with c connected components, it carries a power of N equal to

$$\begin{aligned}\#(N)&=-P+\sum_\alpha V_\alpha\left(2-\sum_p\alpha_p\right)+L\\&=\left(\sum_\alpha V_\alpha\sum_p\alpha_p-P+L\right)+2\sum_\alpha V_\alpha\left(1-\sum_p\alpha_p\right)\\&=2-2g+2\left(c-1-\sum_\alpha V_\alpha\left(\sum_p\alpha_p-1\right)\right)\leqslant 2-2g\end{aligned}$$

where the last inequality expresses that to separate the diagram into c connected components, one must cut at most its $\sum_\alpha V_\alpha\left(\sum_p\alpha_p-1\right)$ dotted lines.

This simple counting (see the second reference of [30]) shows that the leading $O(N^2)$ terms contributing to F are obtained as the sum of planar (i.e. genus $g = 0$) and minimally connected Feynman diagrams, such that cutting any dotted line makes them disconnected. They are thus trees in these dotted lines.

These Feynman rules, supplemented by the usual prescriptions for the symmetry factor of each diagram (equal to the inverse of the order of its automorphism group), are what is needed to get a systematic expansion of F in powers of the θ and $\bar{\theta}$:

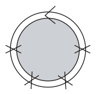
$$F = \sum_{\substack{\text{minimally connected}\\ \text{diagrams } \mathcal{G}}} \frac{s^{-P}}{|\mathrm{Aut}\,\mathcal{G}|} \prod_{\text{vertices}} W_\alpha \prod_{a\text{-faces}} \theta_p \prod_{b\text{-faces}} \bar{\theta}_q. \tag{5.5}$$

Note that these rules imply that the sign attached to each diagram (coming from the signs of the W_α) is also $(-1)^L = (-1)^{\text{power of } \theta + \text{power of } \bar{\theta}}$.

Examples. $W(XX^\dagger; s)$ contains a term linear in $XX^\dagger$ (with coefficient $1/s$), which may be inserted in arbitrary number on any propagator without changing the topology of the diagram: each individual insertion (which raises the corresponding power of a and b by one unit) is depicted by a cross on the propagator.

Terms in F linear in θ and $\bar{\theta}$ come solely from these insertions, which yield

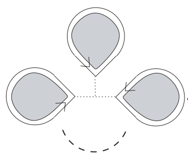
$$\sum_{p=1}^{\infty} \frac{1}{p} \frac{\theta_p \bar{\theta}_p}{s^p} :$$



(a single loop with p insertions has a symmetry factor equal to p).

Terms in F of the form $(\theta_1\bar{\theta}_1/s)^p$ come entirely from a multi-vertex of type $\mathrm{tr}_{[1^p]}\, XX^\dagger$, whence a contribution

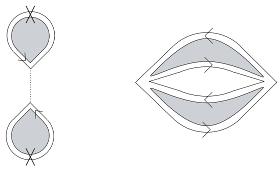
$$2^p \frac{(3p-3)!}{p!(2p)!}$$



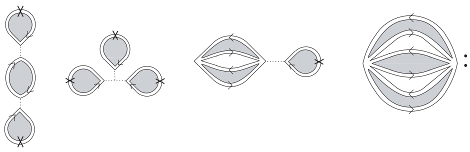
.

More generally, the term in F of the form $\prod_p (\theta_p)^{\alpha_p} \bar{\theta}_1^n$ comes solely from the multi-vertex of type α, $\alpha \vdash n$, in which the contraction of b lines into 'petals' is unique and determines that of a lines. The weight W_α together with the symmetry factor $\prod_p p^{\alpha_p} \alpha_p!$ reproduces the result of (4.26).

Likewise, it is easy to find the diagrams contributing to the $(\theta_2\bar{\theta}_2/s^2)^2$ and $(\theta_2\bar{\theta}_2/s^2)^3$ terms in F



which give $\left(\frac{1}{2} + \left(\frac{1}{2}\right)^2\right)(\theta_2\bar{\theta}_2/s^2)^2 = \frac{3}{4}(\theta_2\bar{\theta}_2/s^2)^2$ and



$$: \quad \left(\frac{1}{2} + \frac{1}{3!} \times 8 + \frac{1}{2} \times 4 + \frac{1}{3!} \times 2^2\right) \left(\frac{\theta_2\bar{\theta}_2}{s^2}\right)^3$$

$$= \frac{9}{2} \left(\frac{\theta_2\bar{\theta}_2}{s^2}\right)^3$$

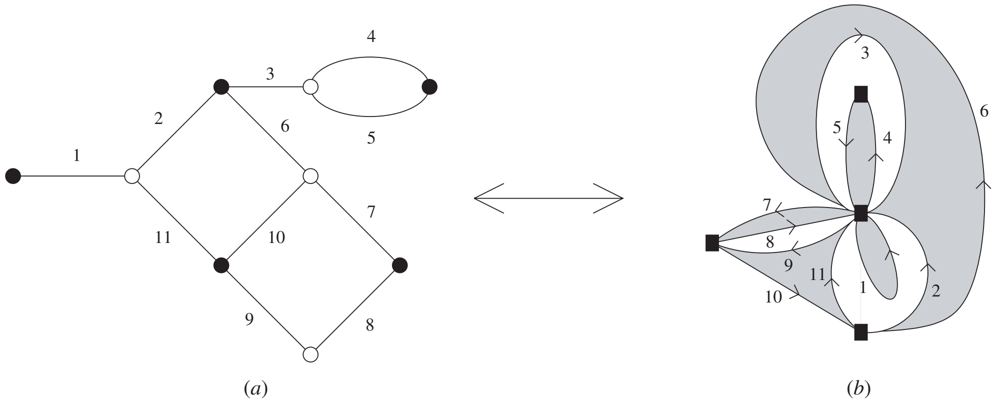


Figure 2. (*a*) A bicoloured planar map and (*b*) its dual. $\sigma = (1\ 2\ 11)\ (3\ 4\ 5)\ (6\ 7\ 10)\ (8\ 9)$, $\tau = (1)\ (2\ 3\ 6)\ (4\ 5)\ (7\ 8)\ (9\ 11\ 10)$, $\sigma\tau = (1\ 3\ 5\ 6\ 8\ 11)\ (2\ 10)\ (4)\ (7\ 9)$.

in agreement with the expansion at the end of section 4. Conversely, one sees how effective the methods of section 4 are to resum the classes of Feynman diagrams. Indeed, it is not obvious how to derive directly from the diagrammatic expansion that $F(\theta_2\bar{\theta}_2)$, or more generally $F(\theta_n\bar{\theta}_n)$ and $\psi(\theta_n\bar{\theta}_n)$, have the simple form given by equation (4.28). In particular, it would be interesting to find a direct combinatorial proof of equation (4.27) satisfied by ψ. Its form suggests a possible connection with decorated rooted trees, perhaps in the spirit of [34].

Remark. Actions of the form of equation (5.2) and their diagrammatic expansion are a generalization of the 'dually weighted graphs' of [35]: the extra ingredient is these graphs are required to be bicoloured and the black/white faces are weighted separately. They are also a generalization of the bicoloured diagrams of [36]—the faces of the latter being weighted only according to their colour and not to their number of edges (this corresponds to the case where A and B are projectors).

5.2. Combinatorics

We now show that this expansion is in fact equivalent to that recently obtained in [10]. What is remarkable is that the methods are orthogonal: [10] is based on manipulations in the symmetric group, and in particular uses some results on the number of solutions of equations for permutations (which are however themselves related to planar constellations [34]). We now explain the results of [10] that are relevant here, and show their equivalence to our Feynman diagram expansion.

First, we explain how to associate with a (not necessarily connected) bicoloured map g, a pair of permutations $\sigma(g)$ and $\tau(g)$ of the set of edges. $\sigma(g)$ (resp. $\tau(g)$) is the permutation which with an edge associates the next edge obtained by clockwise rotation around the white (resp. black) vertex to which it is connected, see the example of figure 2(*a*). This is a one-to-one correspondence.

An important remark is that the permutation $\sigma(g)\tau(g)$ encodes the *faces* of the map. Indeed, cycles of $\sigma(g)\tau(g)$ are obtained by turning clockwise around each face and keeping only the edges going from a white to a black vertex. It is implied that each connected component has its own face at infinity (for which rotation must be made anticlockwise). Consequently, the lengths of the cycles of $\sigma(g)\tau(g)$ are precisely one half of the sizes of the faces.

Also, for ρ a permutation of the set of edges, define Π_ρ to be the partition of the *set* of edges into the orbits (cycles) of ρ. For two such partitions Π and Π', we say that $\Pi \leqslant \Pi'$ iff every block of Π is included in a block of Π'; and define $\Pi \vee \Pi'$ to be the smallest common majorant. Also, denote by $\#\Pi$ the number of blocks of Π. Finally, if n is the number of edges, define $C_\rho \vdash n$ to be the partition of the *integer* n corresponding to Π_ρ, i.e. the lengths of the cycles of ρ.

Then the results of [10, theorems 2.12, 4.2] can be reformulated as follows:

$$F(A, B; s) = \sum_{n=1}^{\infty} \frac{s^{-n}}{n!} \sum_{\substack{\sigma,\tau\in\mathfrak{S}_n \\ \#\Pi_\sigma+\#\Pi_\tau+\#\Pi_{\sigma\tau}-n=2\#(\Pi_\sigma\vee\Pi_\tau)}} \gamma\,(\sigma\tau, \Pi_\sigma \vee \Pi_\tau)\, \mathrm{tr}_{C_\sigma} A\, \mathrm{tr}_{C_\tau} B. \tag{5.6}$$

The coefficient $\gamma(\rho, \Pi)$ is only defined explicitly if $\Pi = \Pi_\rho$, in which case $\gamma(\rho, \Pi_\rho) = W_{C_\rho}$ where W_α is given by equation (5.4*b*). In general, we have the following expression for $\gamma(\rho, \Pi)$:

$$\gamma(\rho, \Pi) = \sum_{\substack{\Pi'\geqslant\Pi_\rho,\#(\Pi\vee\Pi')=1 \\ \#\Pi_\rho-\#\Pi'=\#\Pi-1}} \prod_{\substack{\text{blocks}\, i \\ \text{of}\, \Pi'}} W_{C_{\rho_i}} \tag{5.7}$$

where the ρ_i are the restrictions of ρ to each block i of the partition Π'. This expression looks somewhat complicated but it is very easy to interpret once we relate this formalism to the diagrammatic expansion of the previous section.

In order to proceed with the equivalence, we first associate with the pair of permutations σ and τ a map according to the above construction. Since σ and τ are permutations of $\{1, \ldots, n\}$, this produces a map with *labelled edges* from 1 to n. The quantity $\#\Pi_\sigma + \#\Pi_\tau + \#\Pi_{\sigma\tau} - n$ is simply the Euler–Poincaré characteristic of the map and the condition on the summation simply imposes *planarity* of the map (or more precisely, of each of its connected components).

Secondly, we unlabel the resulting map. That is, since the summand of equation (5.6) does not actually depend on the labelling, one can sum together maps which are only distinguished by the labelling of edges; by definition of the symmetry factor of a map (inverse of the order of the automorphism group, i.e. here number of permutations that commute with both σ and τ), the only modification this produces is to replace the $1/n!$ with the symmetry factor of the unlabelled map. Thus, we rewrite equation (5.6) as

$$F(A, B, s) = \sum_g \frac{s^{-n}}{|\mathrm{Aut}\, g|} \gamma(\sigma(g)\tau(g), \Pi_{\sigma(g)} \vee \Pi_{\tau(g)})\, \mathrm{tr}_{C_{\sigma(g)}} A\, \mathrm{tr}_{C_{\tau(g)}} B \tag{5.8}$$

where the summation is over inequivalent (possibly disconnected) planar maps g, and n is the number of edges.

Thirdly, we replace these maps g with their dual maps $\hat{g}$. Note that this must be performed separately for each connected component. The correspondence is again one-to-one. An example is given in figure 2. The new maps have bicoloured faces; equivalently, one can orient their edges so that white (resp. black) vertices correspond to clockwise (resp. anticlockwise) faces.

The resulting maps $\hat{g}$ are very similar to the Feynman diagrams of section 5.1; however, they still lack the 'dotted lines' which link together the various connected components. This is where equation (5.7) comes in. Indeed with each term in the sum (5.7) we associate one particular set of dotted lines as follows: since $\Pi' \geqslant \Pi_{\sigma(g)\tau(g)}$, and the cycles of $\sigma(g)\tau(g)$ are associated with faces of g, one can think of Π' as ways of grouping together faces of g, that is vertices of $\hat{g}$: these are precisely the dotted lines. Finally, the other conditions in the summation of equation (5.7) can be interpreted as follows: $\#(\Pi' \vee \Pi_{\sigma(g)} \vee \Pi_{\tau(g)}) = 1$ ensures that the diagram including dotted lines is connected; and $\#\Pi_{\tau(g)\sigma(g)} - \#\Pi' = \#(\Pi_{\sigma(g)} \vee \Pi_{\tau(g)}) - 1$ ensures that it is 'minimally connected', i.e. that it has a tree structure.

Finally, we compare the weights: the $W_{C_{\rho_i}}$ are associated with groups of vertices of $\hat{g}$ linked together by dotted lines, just as in the previous section; as to the $\mathrm{tr}_{C_{\sigma(g)}} A$ (resp. $\mathrm{tr}_{C_{\tau(g)}} B$), they associate with each white (resp. black) vertex of g, that is each clockwise (resp. anticlockwise) face of $\hat{g}$, a weight of θ_p (resp. $\bar{\theta}_p$) where p is the number of edges surrounding it, which is again what is required.

Remark. The special case $\Pi_{\sigma(g)} \vee \Pi_{\tau(g)} = \Pi_{\sigma(g)\tau(g)}$, for which the associated weight is a single $W_{C_{\sigma(g)\tau(g)}}$, occurs when g is a disjoint union of bicoloured trees, or equivalently when all vertices of $\hat{g}$ are linked together by dotted lines.

5.3. Case of rectangular matrices

We sketch here how the diagrammatic expansion of section 5.1 can be generalized to the case of the integral (3.5) over rectangular matrices. Introduce complex rectangular $N_1 \times N_2$ matrices X and Y; then, applying the same trick as above, we find

$$\begin{aligned}\int_{U(N_2)} \mathrm{D}U \int_{U(N_1)} \mathrm{D}V \exp\left(\frac{N}{s}\operatorname{Tr}(AUBV^\dagger + VB^\dagger U^\dagger A^\dagger)\right) \\ = \text{const} \int \mathrm{D}X\,\mathrm{D}X^\dagger\,\mathrm{D}Y\,\mathrm{D}Y^\dagger \exp\big(-N' \operatorname{Tr}(XX^\dagger + YY^\dagger) \\ + N_2^2 W(X^\dagger AA^\dagger Y; s) + N_1^2 W(B^\dagger Y^\dagger XB; s)\big) \qquad (5.9)\end{aligned}$$

where $N = \min(N_1, N_2)$ and $N' = \max(N_1, N_2)$. The diagrammatic expansion is identical to that of section 5.1, except that now there are two types of vertices, weighted by extra factors N_1/N' and N_2/N', respectively. Due to the presence of X and $Y^\dagger$, resp. Y and $X^\dagger$, in these vertices, we see that they must alternate and therefore these diagrams have both faces *and* vertices bicoloured. Up to a factor of 2 corresponding to the two possible colourings of the vertices, this is the same as requiring the diagrams to have bicoloured even-sized faces. Each face of size $2p$ receives a weight $\operatorname{Tr}(AA^\dagger)^p$ or $\operatorname{Tr}(BB^\dagger)^p$ depending on its colour (orientation).

Note that if $N_1 = N_2 = N$, the diagrams are weighted identically as in section 5.1; this means that the integral (5.2) for square matrices can be considered in the large N limit as the particular case of the integral (1.1) for which all odd moments vanish (cf a similar observation in section 4.4).

6. Summary of results, tables

One may compute the first terms of the expansion $F(A, B; s) = \sum_{n=1}^{\infty} \frac{1}{s^n} F_n(\theta, \bar{\theta})$. It is sufficient to tabulate them for $\theta_1 = \bar{\theta}_1 = 0$ since if we write $A = A' + \theta_1 I$, $B = B' + \bar{\theta}_1 I$, with A' and B' traceless, $F(A, B; s) = F(A', B'; s) + \theta_1\bar{\theta}_1/s$. Or alternatively

$$F_n(\theta_1, \bar{\theta}_1, \theta_2, \bar{\theta}_2, \ldots) = \theta_1\bar{\theta}_1\delta_{n1} + F_n(0, 0, \theta_2', \bar{\theta}_2', \ldots)$$

with $\theta_p' = \sum_{q=0}^{p} \binom{p}{q} \theta_q (-\theta_1)^{p-q}$ and likewise for $\bar{\theta}_p$. Up to order 8, we find

$$F_1 = 0 \qquad F_2 = \tfrac{1}{2}\theta_2\bar{\theta}_2 \qquad F_3 = \tfrac{1}{3}\theta_3\bar{\theta}_3$$

$$F_4 = \tfrac{1}{4}\big[\theta_2^2\big(3\bar{\theta}_2^2 - 2\bar{\theta}_4\big) + \theta_4\big(-2\bar{\theta}_2^2 + \bar{\theta}_4\big)\big]$$

$$F_5 = \tfrac{1}{5}[\theta_2\theta_3(20\bar{\theta}_2\bar{\theta}_3 - 5\bar{\theta}_5) + \theta_5(-5\bar{\theta}_2\bar{\theta}_3 + \bar{\theta}_5)]$$

$$
\begin{aligned}
F_6 &= \tfrac{1}{6}\big[\theta_2^3\big(27\bar\theta_2^3 - 16\bar\theta_3^2 - 30\bar\theta_2\bar\theta_4 + 7\bar\theta_6\big) + \theta_3^2\big(-16\bar\theta_2^3 + 6\bar\theta_3^2 + 15\bar\theta_2\bar\theta_4 - 3\bar\theta_6\big)\\
&\qquad + 3\theta_2\theta_4\big(-10\bar\theta_2^3 + 5\bar\theta_3^2 + 10\bar\theta_2\bar\theta_4 - 2\bar\theta_6\big) + \theta_6\big(7\bar\theta_2^3 - 3\bar\theta_3^2 - 6\bar\theta_2\bar\theta_4 + \bar\theta_6\big)\big]\\
F_7 &= \theta_2^2\theta_3\big(66\bar\theta_2^2\bar\theta_3 - 21\bar\theta_2\bar\theta_5 - 20\bar\theta_3\bar\theta_4 + 4\bar\theta_7\big) - \theta_3\theta_4\big(20\bar\theta_2^2\bar\theta_3 - 5\bar\theta_3\bar\theta_4 - 6\bar\theta_2\bar\theta_5 + \bar\theta_7\big)\\
&\qquad - \theta_2\theta_5\big(21\bar\theta_2^2\bar\theta_3 - 6\bar\theta_3\bar\theta_4 - 6\bar\theta_2\bar\theta_5 + \bar\theta_7\big) + \tfrac{1}{7}\theta_7\big(28\bar\theta_2^2\bar\theta_3 - 7\bar\theta_3\bar\theta_4 - 7\bar\theta_2\bar\theta_5 + \bar\theta_7\big)\\
F_8 &= \tfrac{1}{8}\big[3\theta_2^4\big(117\bar\theta_2^4 - 192\bar\theta_2\bar\theta_3^2 - 180\bar\theta_2^2\bar\theta_4 + 25\bar\theta_4^2 + 56\bar\theta_3\bar\theta_5 + 56\bar\theta_2\bar\theta_6 - 10\bar\theta_8\big)\\
&\qquad - 4\theta_2\theta_3{}^2\big(144\bar\theta_2^4 - 176\bar\theta_2\bar\theta_3^2 - 200\bar\theta_2^2\bar\theta_4 + 25\bar\theta_4^2 + 48\bar\theta_3\bar\theta_5 + 56\bar\theta_2\bar\theta_6 - 9\bar\theta_8\big)\\
&\qquad - 4\theta_2^2\theta_4\big(135\bar\theta_2^4 - 195\bar\theta_2^2\bar\theta_4 + 25\bar\theta_4^2 + 54\bar\theta_3\bar\theta_5 + \bar\theta_2\big(-200\bar\theta_3^2 + 56\bar\theta_6\big) - 9\bar\theta_8\big)\\
&\qquad + \theta_4^2\big(75\bar\theta_2^4 - 100\bar\theta_2\bar\theta_3^2 - 100\bar\theta_2^2\bar\theta_4 + 10\bar\theta_4^2 + 24\bar\theta_3\bar\theta_5 + 28\bar\theta_2\bar\theta_6 - 4\bar\theta_8\big)\\
&\qquad + \theta_3\theta_5\big(168\bar\theta_2^4 - 192\bar\theta_2\bar\theta_3^2 - 216\bar\theta_2^2\bar\theta_4 + 24\bar\theta_4^2 + 48\bar\theta_3\bar\theta_5 + 56\bar\theta_2\bar\theta_6 - 8\bar\theta_8\big)\\
&\qquad + \theta_2\theta_6\big(168\bar\theta_2^4 - 224\bar\theta_2\bar\theta_3^2 - 224\bar\theta_2^2\bar\theta_4 + 28\bar\theta_4^2 + 56\bar\theta_3\bar\theta_5 + 56\bar\theta_2\bar\theta_6 - 8\bar\theta_8\big)\\
&\qquad + \theta_8\big(-30\bar\theta_2^4 + 36\bar\theta_2\bar\theta_3^2 + 36\bar\theta_2^2\bar\theta_4 - 4\bar\theta_4^2 - 8\bar\theta_3\bar\theta_5 - 8\bar\theta_2\bar\theta_6 + \bar\theta_8\big)\big].
\end{aligned}
$$

One may slightly simplify this expansion by making the following asymmetric change of variables: we introduce instead of the moments $\bar\theta_q$, the *free cumulants* $\bar\phi_q$ which are defined by [8, 37, 38]

$$
\bar\phi_q = -\sum_{\substack{\alpha_1,\ldots,\alpha_q\geqslant 0\\ \sum_i i\alpha_i = q}} \frac{\big(q + \sum_i \alpha_i - 2\big)!}{(q-1)!}\prod_i \frac{(-\bar\theta_i)^{\alpha_i}}{\alpha_i!}. \tag{6.1}
$$

Indeed, if one now expands F as

$$
F = \sum_{q\geqslant 1} \frac{\partial F}{\partial\theta_q}\bigg|_{\theta=0} \theta_q + \frac{1}{2!}\sum_{q,r\geqslant 1}\frac{\partial^2 F}{\partial\theta_q\partial\theta_r}\bigg|_{\theta=0}\theta_q\theta_r + \cdots
$$

then the first three derivatives were computed exactly in [24] using the formalism of dispersionless Toda hierarchy, and turn out to be simply expressible in terms of the $\bar\phi_q$. In particular, $\frac{\partial F}{\partial\theta_q}\big|_{\theta=0} = \frac{1}{q}\bar\phi_q$. The same expansion to order 8 now takes the form:

$$
\begin{aligned}
F_2 &= \tfrac{1}{2}\theta_2\bar\phi_2 \qquad F_3 = \tfrac{1}{3}\theta_3\bar\phi_3 \qquad F_4 = \tfrac{1}{4}\theta_4\bar\phi_4 - \tfrac{1}{2}\theta_2^2\big(\bar\phi_4 + \tfrac{1}{2}\bar\phi_2^2\big)\\
F_5 &= \tfrac{1}{5}\theta_5\bar\phi_5 - \theta_3\theta_2(\bar\phi_5 + \bar\phi_3\bar\phi_2)\\
F_6 &= \tfrac{1}{6}\theta_6\bar\phi_6 - \theta_4\theta_2\big(\bar\phi_6 + \bar\phi_4\bar\phi_2 + \tfrac{1}{2}\bar\phi_3^2\big) - \tfrac{1}{2}\theta_3^2\big(\bar\phi_6 + \bar\phi_4\bar\phi_2 + \bar\phi_3^2 + \tfrac{1}{3}\bar\phi_2^3\big)\\
&\qquad + \tfrac{1}{6}\theta_2^3\big(7\bar\phi_6 + 12\bar\phi_4\bar\phi_2 + 5\bar\phi_3^2 + 2\bar\phi_2^3\big)\\
F_7 &= \tfrac{1}{7}\theta_7\bar\phi_7 - \theta_5\theta_2(\bar\phi_7 + \bar\phi_5\bar\phi_2 + \bar\phi_4\bar\phi_3) - \theta_4\theta_3\big(\bar\phi_7 + \bar\phi_5\bar\phi_2 + 2\bar\phi_4\bar\phi_3 + \bar\phi_3\bar\phi_2^2\big)\\
&\qquad + \theta_3\theta_2^2\big(4\bar\phi_7 + 7\bar\phi_5\bar\phi_2 + 8\bar\phi_4\bar\phi_3 + 5\bar\phi_3\bar\phi_2^2\big)\\
F_8 &= \tfrac{1}{8}\theta_8\bar\phi_8 - \theta_6\theta_2\big(\bar\phi_8 + \bar\phi_6\bar\phi_2 + \bar\phi_5\bar\phi_3 + \tfrac{1}{2}\bar\phi_4^2\big)\\
&\qquad - \theta_5\theta_3\big(\bar\phi_8 + \bar\phi_6\bar\phi_2 + 2\bar\phi_5\bar\phi_3 + \bar\phi_4^2 + \bar\phi_4\bar\phi_2^2 + \bar\phi_3^2\bar\phi_2\big)\\
&\qquad - \tfrac{1}{2}\theta_4^2\big(\bar\phi_8 + \bar\phi_6\bar\phi_2 + 2\bar\phi_5\bar\phi_3 + \tfrac{3}{2}\bar\phi_4^2 + \bar\phi_4\bar\phi_2^2 + 2\bar\phi_3^2\bar\phi_2 + \tfrac{1}{4}\bar\phi_2^4\big)\\
&\qquad + \tfrac{1}{2}\theta_4\theta_2^2\big(9\bar\phi_8 + 16\bar\phi_6\bar\phi_2 + 18\bar\phi_5\bar\phi_3 + 11\bar\phi_4^2 + 11\bar\phi_4\bar\phi_2^2 + 14\bar\phi_3^2\bar\phi_2 + \bar\phi_2^4\big)\\
&\qquad + \theta_3^2\theta_2\big(9\bar\phi_8 + 16\bar\phi_6\bar\phi_2 + 24\bar\phi_5\bar\phi_3 + 11\bar\phi_4^2 + 16\bar\phi_4\bar\phi_2^2 + 20\bar\phi_3^2\bar\phi_2 + 2\bar\phi_2^4\big)\\
&\qquad - \tfrac{3}{8}\theta_2^4\big(10\bar\phi_8 + 24\bar\phi_6\bar\phi_2 + 24\bar\phi_5\bar\phi_3 + 15\bar\phi_4^2 + 24\bar\phi_4\bar\phi_2^2 + 24\bar\phi_3^2\bar\phi_2 + 3\bar\phi_2^4\big).
\end{aligned}
$$

One notes the intriguing feature that in either basis, for all known terms (up to $p = 10$), pF_p has only integer coefficients. In other words, the quantity $s\partial F/\partial s$ has only integer coefficients. The rationale behind this fact and its reinterpretation in terms of the counting of 'rooted' objects have remained elusive to us. The integral (1.1) still has a few mysteries . . . .

In this paper, we have presented the standard lore on the integral (1.1) as well as recent developments. The integral itself is well understood by a variety of methods (section 2) and admits interesting generalizations (section 3). However, the explicit expression, in terms of symmetric functions of the eigenvalues, of the integral itself (not to mention the more complicated problem of the associated correlation functions) is more subtle; section 4 was devoted to the discussion of this problem in the large N limit. These considerations, based on integrable hierarchies, are complemented by other recent works based on combinatorial arguments, as discussed in section 5. It is our feeling that the exact interrelations between these various approaches require further study.

Acknowledgments

It is a pleasure to thank M Bauer for his help in the reading of [14]. One of the authors (J-BZ) wants to acknowledge the hospitality of MSRI, Berkeley, where a part of this work was performed.

References

[1] Verbaarschot J J M and Wettig T 2000 *Ann. Rev. Nucl. Part. Sci.* **50** 343–410 (*Preprint* hep-ph/0003017)
[2] Di Francesco P, Ginsparg P and Zinn-Justin J 1995 *Phys. Rep.* **254** 1–133
[3] Guhr T, Mueller-Groeling A and Weidenmueller H A 1998 *Phys. Rep.* **299** 189–425 (*Preprint* cond-mat/9707301)
[4] Di Francesco P 2001 *MSRI Publications* vol 40, ed P M Bleher and A R Its (Cambridge: Cambridge University Press) pp 111–70 (*Preprint* math-ph/9911002)
[5] Morozov A 1994 *Phys. Usp.* **37** 1–55 (*Preprint* hep-th/9303139)
Adler M and van Moerbeke P 1997 *Commun. Pure Appl. Math.* **50** 241–90 (*Preprint* hep-th/9706182)
Orlov A Yu 2002 Tau functions and matrix integrals *Preprint* math-ph/0210012
[6] Voiculescu D V, Dykema K J and Nica A 1992 *Free Random Variables* (Providence, RI: AMS)
[7] James A T 1964 *Ann. Math. Stat.* **35** 475–501
[8] Itzykson C and Zuber J-B 1980 *J. Math. Phys.* **21** 411–21
[9] Harish-Chandra 1957 *Am. J. Math.* **79** 87–120
[10] Collins B 2002 Moments and cumulants of polynomial random variables on unitary groups, the Itzykson–Zuber integral and free probability *Preprint* math-ph/0205010
[11] Guionnet A and Zeitouni O Large deviations asymptotics for spherical integrals webpage http://www.umpa.ens-lyon.fr/~aguionne (*J. Funct. Anal.* at press)
[12] Weyl H 1948 *The Classical Groups* (Princeton, NJ: Princeton University Press)
[13] Balantekin A B 2000 *Phys. Rev.* D **62** 085017 (*Preprint* hep-th/0007161)
[14] Duistermaat J J and Heckman G J 1982 *Invent. Math.* **69** 259–68
Duistermaat J J and Heckman G J 1983 *Invent. Math.* **72** 153–8
[15] Stone M 1989 *Nucl. Phys.* B **314** 557–86
[16] Szabo R J 1996 Equivariant localization of path integrals *Preprint* hep-th/9608068
Karshon Y 1996–97 Lecture notes on group actions on manifolds webpage http://www.ma.huji.ac.il/~karshon/teaching/1996–97/actions/lecture-notes.html
[17] Altschuler D and Itzykson C 1991 *Ann. Inst. H Poincaré* **54** 1–8
[18] Alfaro J, Medina R and Urrutia L F 1995 *J. Math. Phys.* **36** 3085–93 (*Preprint* hep-th/9412012)
Guhr T 1996 *Commun. Math. Phys.* **176** 555–86
[19] Guhr T and Wettig T 1996 *J. Math. Phys.* **37** 6395–413 (*Preprint* hep-th/9605110)
[20] Fyodorov Y V and Strahov E 2002 *Nucl. Phys.* B **630** 453–91 (*Preprint* math-ph/0201045)
[21] Jackson A D, Şener M K and Verbaarschot J J M 1996 *Phys. Lett.* B **387** 355–60 (*Preprint* hep-th/9605183)
[22] Shatashvili S 1993 *Commun. Math. Phys.* **154** 421–32 (*Preprint* hep-th/9209083)

[23] Morozov A 1992 *Mod. Phys. Lett.* A **7** 3503–8 (*Preprint* hep-th/9209074)
[24] Zinn-Justin P 2002 *Nucl. Phys.* B **634** 417–32 (*Preprint* math-ph/0202045)
[25] Kazakov V A and Migdal A A 1993 *Nucl. Phys.* B **397** 214–38
[26] Ueno K and Takasaki K 1984 *Group representations and systems of differential equations* ed H Morikawa and K Okamoto p 1–95
[27] Kostov I K 2002 *Nucl. Phys.* B **624** 146–62 (*Preprint* hep-th/0107247)
[28] Matytsin A 1994 *Nucl. Phys.* B **411** 805–20 (*Preprint* hep-th/9306077)
[29] Zinn-Justin P 1998 *Commun. Math. Phys.* **194** 631–50
[30] O'Brien H and Zuber J-B 1984 *Phys. Lett.* B **144** 407–8
O'Brien H and Zuber J-B 1985 *Nucl. Phys.* B **253** 621–34
[31] Bessis D, Itzykson C and Zuber J-B 1980 *Adv. Appl. Math.* **1** 109–57
[32] Zvonkin A 1997 *Math. Comput. Modelling* **26** 281–304
[33] 't Hooft G 1974 *Nucl. Phys.* B **72** 461–73
[34] Bousquet-Mélou M and Schaeffer G 2000 *Adv. Appl. Math.* **24** 337–68
Schaeffer G Conjugaison d'arbres et cartes combinatoires aléatoires *Thèse de Doctorat* webpage http://dept-info.labri.u-bordeaux.fr/∼schaeffe/cv/bibli/These.html
[35] Kazakov V A, Staudacher M and Wynter T 1996 *Commun. Math. Phys.* **177** 451–68
Kazakov V A, Staudacher M and Wynter T 1996 *Commun. Math. Phys.* **179** 235–56
Kazakov V A, Staudacher M and Wynter T 1996 *Nucl. Phys.* B **471** 309–33
[36] Di Francesco P 2002 Rectangular matrix models and combinatorics of colored graphs *Preprint* cond-mat/0208037
[37] Brézin E, Itzykson C, Parisi G and Zuber J-B 1978 *Commun. Math. Phys.* **59** 35–51
[38] Speicher R 1994 *Math. Ann.* **298** 611–28