Some matrix integrals related to random matrix theory

Michael E. Taylor

Abstract. We present a "naive" derivation of a formula of [BDK] for $\int_{U(n)} |\mathrm{Tr}\, M^j|^2\, dM$, of use in random matrix theory. We also calculate a more refined object, $\int_{U(n)} M^j \otimes M^{-j}\, dM$, which in turn yields a formula for $\int_{U(n)} f(M) \otimes g(M)\, dM$.

1. Inner products of trace functions on $U(n)$

Let $f : S^1 \to \mathbb{C}$ be a bounded Borel function. Given $M \in U(n)$, we define $f(M) \in \mathrm{End}(\mathbb{C}^n)$ by the spectral representation, and consider the trace function

$$X_f : U(n) \to \mathbb{C}, \quad X_f(M) = \mathrm{Tr}\, f(M). \tag{1.1}$$

We are interested in formulas for

$$\int\limits_{U(n)} X_f(M) X_g(M)\, dM, \tag{1.2}$$

recovering some formulas from [BDK], involving identities of Dyson [D], also presented in [M]. We look for alternatives to the method of Dyson.

Note that (1.2) is equal to the trace of

$$\int\limits_{U(n)} f(M) \otimes g(M)\, dM. \tag{1.3}$$

We use Fourier series:

$$f(M) = \sum_{j=-\infty}^{\infty} \hat{f}(j) M^j, \quad \hat{f}(j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ij\theta}\, d\theta. \tag{1.4}$$

We see that (1.3) is equal to

$$\sum_{j,k} \hat{f}(j) \hat{g}(k) A_{jk}, \quad A_{jk} = \int\limits_{U(n)} M^j \otimes M^k\, dM. \tag{1.5}$$

Performing the measure-preserving transformation $M \mapsto e^{i\psi}M$ on $U(n)$, we see that

$$A_{jk} = e^{i(j+k)\psi} A_{jk} \tag{1.6}$$

for all $\psi \in \mathbb{R}$, and hence $A_{jk} = 0$ for $j \neq -k$. Hence

$$\int\limits_{U(n)} f(M) \otimes g(M)\, dM = \sum_j \hat{f}(j)\hat{g}(-j)A_j, \tag{1.7}$$

with

$$A_j = \int\limits_{U(n)} M^j \otimes M^{-j}\, dM. \tag{1.8}$$

Note that performing the measure-preserving transformation $M \mapsto M^{-1}$ on $U(n)$ yields

$$A_j = A_{-j}. \tag{1.9}$$

From (1.7) we deduce that

$$\int\limits_{U(n)} X_f(M)X_g(M)\, dM = \sum_j \hat{f}(j)\hat{g}(-j)a_{jn}, \tag{1.10}$$

where

$$a_{jn} = \operatorname{Tr} A_j = \int\limits_{U(n)} |\operatorname{Tr} M^j|^2\, dM. \tag{1.11}$$

Clearly

$$a_{0n} = n^2. \tag{1.12}$$

We will show in §2 that

$$\begin{aligned} a_{jn} = j \quad &\text{for} \ \ 1 \le j \le n, \\ n \quad &\text{for} \ \ j \ge n. \end{aligned} \tag{1.13}$$

By (1.9) we have $a_{-j,n} = a_{jn}$. This provides an analysis of (1.2).

For the operator analysis of (1.3), we need a more precise analysis of the operators A_j, given by (1.8). This analysis is carried out in §3.

2. The square-norm of Tr M^j

Here we establish the following identity:

Proposition 2.1. *We have*

$$\int_{U(n)} |Tr\, M^j|^2\, dM = j \quad for \;\; 1 \le j \le n, \tag{2.1}$$
$$n \quad for \;\; j \ge n.$$

As mentioned in §1, this was proven in [BDK], using Dyson's formula. Another proof was given in [DE] (Theorem 2.1, part (b)), using an identity between power-sum symmetric functions and Schur functions. Here we give a "naive" proof, based on Weyl's integration formula, which implies that whenever $\varphi : U(n) \to \mathbb{C}$ is invariant under conjugation, then

$$\int_{U(n)} \varphi(M)\, dM = C_n (2\pi)^{-n} \int_{\mathbb{T}^n} \varphi(D(\theta)) J(\theta)\, d\theta_1 \cdots d\theta_n, \tag{2.2}$$

where $D(\theta)$ is the diagonal matrix with diagonal entries $e^{i\theta_1}, \dots, e^{i\theta_n}$, and

$$J(\theta) = \prod_{k<\ell} |e^{i\theta_k} - e^{i\theta_\ell}|^2. \tag{2.3}$$

We will verify in calculations below that

$$C_n = \frac{1}{n!}. \tag{2.4}$$

In particular, (2.2) gives

$$\int_{U(n)} |\mathrm{Tr}\, M^j|^2\, dM = C_n (2\pi)^{-n} \int_{\mathbb{T}^n} |e^{ij\theta_1} + \cdots + e^{ij\theta_n}|^2 J(\theta)\, d\theta. \tag{2.5}$$

We re-state this as follows. Set $\zeta_j = e^{i\theta_j}$, so

$$|e^{ij\theta_1} + \cdots + e^{ij\theta_n}|^2 = |\zeta_1^j + \cdots + \zeta_n^j|^2 = \sum_{\mu,\nu} \zeta_\mu^j \zeta_\nu^{-j}, \tag{2.6}$$

and

$$\begin{aligned} J(\theta) &= \prod_{k<\ell} |e^{i\theta_k} - e^{i\theta_\ell}|^2 = \prod_{k<\ell} |\zeta_k - \zeta_\ell|^2 \\ &= \prod_{k<\ell} (\zeta_k - \zeta_\ell)(\zeta_k^{-1} - \zeta_\ell^{-1}) \\ &= (\mathrm{sgn}\, \gamma)(\zeta_1 \cdots \zeta_n)^{-(n-1)} \prod_{k<\ell} (\zeta_k - \zeta_\ell)^2, \end{aligned} \tag{2.7}$$

where γ is the permutation on $\{1,\dots,n\}$ such that $\gamma(k)=n+1-k$. We see that $\int_{U(n)}|\mathrm{Tr}\,M^j|^2\,dM$ is the constant term in

$$C_n(\operatorname{sgn}\gamma)(\zeta_1\cdots\zeta_n)^{-(n-1)}\Big(\sum_{\mu,\nu}\zeta_\mu^j\zeta_\nu^{-j}\Big)\prod_{k<\ell}(\zeta_k-\zeta_\ell)^2. \tag{2.8}$$

Thus our task is to identify the constant term in this Laurent polynomial. To work on the last factor, we recognize

$$V(\zeta)=\prod_{k<\ell}(\zeta_k-\zeta_\ell) \tag{2.9}$$

as a Vandermonde determinant; hence

$$V(\zeta)=\sum_{\sigma\in S_n}(\operatorname{sgn}\sigma)\zeta_1^{\sigma(1)-1}\cdots\zeta_n^{\sigma(n)-1}, \tag{2.10}$$

where S_n denotes the group of permutations of $\{1,\dots,n\}$. Hence

$$\prod_{k<\ell}(\zeta_k-\zeta_\ell)^2=V(\zeta)^2=\sum_{\sigma,\tau\in S_n}(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)\zeta_1^{\sigma(1)+\tau(1)-2}\cdots\zeta_n^{\sigma(n)+\tau(n)-2}. \tag{2.11}$$

Before getting back to (2.8), let us first identify the constant term in

$$J(\theta)=(\operatorname{sgn}\gamma)(\zeta_1\cdots\zeta_n)^{-(n-1)}V(\zeta)^2. \tag{2.12}$$

We see this is

$$(\operatorname{sgn}\gamma)\sum(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau), \tag{2.13}$$

where the sum is over all $\sigma,\tau\in S_n$ such that $\sigma(k)+\tau(k)=n+1$, for each $k\in\{1,\dots,n\}$. In other words, we require $\tau=\gamma\sigma$, where γ is as specified in (2.7). Thus the sum in (2.13) is equal to $n!(\operatorname{sgn}\gamma)^2=n!$, which establishes (2.4).

Back to (2.8), i.e., the study of the constant term in

$$C_n(\operatorname{sgn}\gamma)(\zeta_1\cdots\zeta_n)^{-(n-1)}\sum_{\mu,\nu,\sigma,\tau}(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)\zeta_\mu^j\zeta_\nu^{-j}\zeta_1^{\sigma(1)+\tau(1)-2}\cdots\zeta_n^{\sigma(n)+\tau(n)-2}, \tag{2.14}$$

which we write as

$$C_n(\operatorname{sgn}\gamma)(S_1+S_2), \tag{2.15}$$

where S_1 arises from the sum over $\mu=\nu$ and S_2 arises from the sum over $\mu\neq\nu$. Parallel to the analysis of (2.12)–(2.13), we have

$$S_1=n\cdot n!\,(\operatorname{sgn}\gamma), \tag{2.16}$$

or $C_n(\operatorname{sgn}\gamma)S_1 = n$.

It remains to consider S_2. We see that, for a given $\mu \neq \nu$, a pair $\sigma, \tau \in S_n$ contributes to S_2 in the sum (2.14) if and only if $\sigma(k) + \tau(k) = n + 1$ for all but two values of $k \in \{1, \dots, n\}$, namely $k = \mu$ and ν, and

$$\begin{aligned} \sigma(\mu) + \tau(\mu) &= n + 1 - j, \\ \sigma(\nu) + \tau(\nu) &= n + 1 + j. \end{aligned} \tag{2.17}$$

Equivalently, we require $\tau = \psi\gamma\sigma$ where γ is as in (2.7) and $\psi \in S_n$ has the property that $\psi(k) = k$ except for two values of $k \in \{1, \dots, n\}$, namely $k_1 = \gamma\sigma(\mu)$ and $k_2 = \gamma\sigma(\nu)$, and

$$\psi(k_1) = k_1 - j, \quad \psi(k_2) = k_2 + j.$$

This requires $\psi(k_1) = k_2,\ \psi(k_2) = k_2$, with

$$k_1 = k_2 + j. \tag{2.18}$$

Then

$$S_2 = \sum (\operatorname{sgn}\gamma)(\operatorname{sgn}\psi), \tag{2.19}$$

the sum running over such allowable (μ, ν, σ, ψ). Note that (2.18) constrains k_1; we require $j + 1 \leq k_1 \leq n$. Thus if $j \geq n$ the sum in (2.19) is empty and $S_2 = 0$. If $1 \leq j \leq n - 1$, then there are $(n - j) \cdot n!$ terms in the sum (2.19), and each $\operatorname{sgn}\psi = -1$. Hence

$$S_2 = -(n - j) \cdot n!\,(\operatorname{sgn}\gamma), \quad 1 \leq j \leq n - 1, \tag{2.20}$$

and we have (2.1).

3. Operator analysis of A_j

Here we want to compute the operators A_j defined by (1.8), i.e.,

$$A_j = \int_{U(n)} M^j \otimes M^{-j}\, dM. \tag{3.1}$$

Let us define $\sigma : \mathbb{C}^n \otimes \mathbb{C}^n \to \mathbb{C}^n \otimes \mathbb{C}^n$ by

$$\sigma(u \otimes v) = v \otimes u, \tag{3.2}$$

and let I denote the identity operator on $\mathbb{C}^n \otimes \mathbb{C}^n$. Let $j \wedge n$ denote $\min(j, n)$. We will establish the following.

Proposition 3.1. *For $j \geq 1,\ n \geq 2$, we have*

$$A_j = \frac{n^2-(j\wedge n)}{n(n^2-1)}\,\sigma + \frac{(j\wedge n)-1}{n^2-1}\,I. \tag{3.3}$$

Proof. Take $X \in U(n)$, acting on $\mathbb{C}^n \otimes \mathbb{C}^n$ by $X(u\otimes v) = Xu \otimes Xv$. Then

$$X^{-1}A_jX(u\otimes v) = \int_{U(n)} X^{-1}M^jXu \otimes X^{-1}M^{-j}Xv\,dM = A_j(u\otimes v), \tag{3.4}$$

since $M \mapsto X^{-1}MX$ is a measure-preserving transformation on $U(n)$. Thus A_j commutes with the action of $U(n)$ on $\mathbb{C}^n \otimes \mathbb{C}^n$; it follows that

$$A_j = \alpha_{jn}\sigma + \beta_{jn}I, \tag{3.5}$$

for some scalars α_{jn} and β_{jn}. Taking the trace of the right side of (3.5) and comparing this with (1.11)–(1.13), we have

$$n\alpha_{jn} + n^2\beta_{jn} = j\wedge n. \tag{3.6}$$

To obtain a second identity involving α_{jn} and β_{jn}, we proceed as follows. Write (3.5) as

$$A_j(u\otimes v) = \alpha_{jn}\,v\otimes u + \beta_{jn}\,u\otimes v. \tag{3.7}$$

Setting $u = v$ and taking the inner product of both sides with $u\otimes u$ yields

$$\begin{aligned} (\alpha_{jn}+\beta_{jn})|u|^4 &= \int_{U(n)} (M^ju\otimes M^{-j}u, u\otimes u)\,dM \\ &= \int_{U(n)} (M^ju,u)(M^{-j}u,u)\,dM \\ &= \int_{U(n)} |(M^ju,u)|^2\,dM. \end{aligned} \tag{3.8}$$

On the other hand, taking $v \perp u$ and taking the inner product of both sides of (3.7) with $v\otimes u$ yields

$$\begin{aligned} \alpha_{jn}\,|u|^2\,|v|^2 &= \int_{U(n)} (M^ju\otimes M^{-j}v, v\otimes u)\,dM \\ &= \int_{U(n)} (M^ju,v)(v,M^ju)\,dM \\ &= \int_{U(n)} |(M^ju,v)|^2\,dM. \end{aligned} \tag{3.9}$$

Now let $e_1, \dots, e_n$ denote the standard orthonormal basis of $\mathbb{C}^n$, and set $u = e_1$. Apply (3.8) for $\ell = 1$ and (3.9) with $v = e_\ell$ for $\ell > 1$, and sum, to get

$$\begin{aligned} \alpha_{jn} + \beta_{jn} + (n-1)\alpha_{jn} &= \sum_{\ell=1}^{n} \int_{U(n)} |(M^j e_1, e_\ell)|^2 \, dM \\ &= \int_{U(n)} |M^j e_1|^2 \, dM. \end{aligned} \tag{3.10}$$

Since $|M^j e_1| = 1$, we have our second identity:

$$n\alpha_{jn} + \beta_{jn} = 1. \tag{3.11}$$

Now (3.3) follows.

With (3.3) in hand (and noting that $A_0 = I$), we can finish the computation of (1.7). We obtain

$$\begin{aligned} \int_{U(n)} f(M) \otimes g(M) \, dM = \frac{1}{n^2-1} &\big(h(0) - F_n h(0)\big)\Big(I - \frac{\sigma}{n}\Big) \\ &- \frac{1}{n^2-1}\big(h(0) - \hat{h}(0)\big)(I - n\sigma) \\ &+ \hat{h}(0) I, \end{aligned} \tag{3.12}$$

where

$$h(\theta) = f * \check{g}(\theta) = \frac{1}{2\pi} \int_{S^1} f(\varphi) g(\varphi - \theta) \, d\varphi, \tag{3.13}$$

and where $F_n h$ denotes the nth Fejér mean of the Fourier series of h:

$$F_n h(\theta) = \sum_{|j|<n} \Big(1 - \frac{|j| \wedge n}{n}\Big) \hat{h}(j) e^{ij\theta}. \tag{3.14}$$

Note that $\hat{h}(0)$ is equal to the product of the mean values of f and g.

Remark 3.1. Taking v orthogonal to u and taking the inner product of both sides of (3.7) with $u \otimes v$ yields the identity

$$\int_{U(n)} (M^j u, u)\overline{(M^j v, v)} \, dM = \beta_{jn} \, |u|^2 \, |v|^2, \quad \text{for } \ u \perp v. \tag{3.15}$$

Remark 3.2. It is apparent that the matrix entries of A_j are given in terms of

$$\int_{U(n)} m_{\alpha_1\beta_1}\cdots m_{\alpha_j\beta_j}\overline{m}_{\alpha_{j+1}\beta_{j+1}}\cdots\overline{m}_{\alpha_{2j}\beta_{2j}}\,dM = P^{(j)}_{\alpha\beta}. \tag{3.16}$$

In particular, for $j = 1$ one has $(1/n)\delta_{\alpha_1\alpha_2}\delta_{\beta_1\beta_2}$, yielding a direct calculation of A_1, consistent with that case of (3.3). For general j, the quantities $P^{(j)}_{\alpha\beta}$ are matrix entries of

$$P^{(j)} = \int_{U(n)} \big(\otimes^j M\big)\otimes\big(\otimes^j\overline{M}\big)\,dM, \tag{3.17}$$

the orthogonal projection of $(\otimes^j\mathbb{C}^n)\otimes(\otimes^j\mathbb{C}^n)$ onto the linear subspace $\mathcal{I}$ on which $(\otimes^j M)\otimes(\otimes^j\overline{M})$ acst trivially for all $M\in U(n)$. The first fundamental theorem of invariant theory (cf. [GW]) describes a spanning set of $\mathcal{I}$, but it is not an orthonormal basis. The projection $P^{(2)}$ has rank 2, and examining this yields another calculation of A_2. However, for larger j this approach to calculating A_j does not seem straightforward.

Acknowledgment. Thanks to A.N. Varchenko for encouraging me to seek a pattern in the behavior of the operators A_j by taking a close look at A_2.

References

[BDK] D. Bump, P. Diaconis, and J. Keller, Unitary correlations and the Fejer kernel, Preprint, 2001.

[DE] P. Diaconis and S. Evans, Linear functionals of eigenvalues of random matrices, Trans. AMS 353 (2001), 2615–2633.

[D] F. Dyson, Statistical theory of energy levels of complex systems, I–III, J. Math. Phys. 3 (1962), 140–156; 157–165; 166–175.

[GW] R. Goodman and N. Wallach, Representations and Invariants of the Classical Groups, Cambridge Univ. Press, 1998.

[M] M. Mehta, Random Matrices, Academic Press, New York, 1991.