

Lectures 5 and 6: Wick's Theorem and the Semicircle Law

Pair partitions, and two proofs of the same theorem

OUTLINE Wick's theorem $\rightarrow$ moments of the spectrum $\rightarrow$ non-crossing pair partitions $\rightarrow$ the Catalan recursion $\rightarrow$ the semicircle law

Schur complement + Wick $\rightarrow m^2 + zm + 1 = 0 \rightarrow$ the semicircle law again, with its density

HOW TO READ Sections marked * are not lectured. They are complete, they are examinable, and they carry most of the exercises. Exercises are typed A, B and C; each week's quiz has one of each.

1 Wick's theorem

In this lecture we compute every moment of a Gaussian vector.

Definition 1.1 (Pair partition). A *pair partition* of $\{1, \dots, 2q\}$ is a partition into q blocks of size two.

Theorem 1.2 (Wick). Let $X \sim \mathcal{N}(0, \Sigma)$ on $\mathbb{R}^n$ and $a_1, \dots, a_m \in \{1, \dots, n\}$. If m is odd then $\mathbb{E}[X_{a_1} \cdots X_{a_m}] = 0$. If $m = 2q$ then

$$\mathbb{E}[X_{a_1} X_{a_2} \cdots X_{a_{2q}}] = \sum_{\text{pair partitions } P \text{ of } \{1, \dots, 2q\}} \prod_{\{r, s\} \in P} \Sigma_{a_r a_s}$$

Proof. By Corollary 1.4 of Lectures 1–2 every linear functional of X is a one-dimensional Gaussian, so X has moments of every order, φ_X is infinitely differentiable, and its derivatives may be taken inside the expectation. Since $\partial_{t_j} e^{i\langle t, x \rangle} = i x_j e^{i\langle t, x \rangle}$, multiplication by X_j acts on the transform as $-i \partial_{t_j}$:

$$\mathbb{E}[X_{a_1} \cdots X_{a_m}] = (-i)^m \partial_{t_{a_1}} \cdots \partial_{t_{a_m}} \varphi_X(t) \Big|_{t=0}. \quad (1)$$

By Definition 1.2, $\varphi_X(t) = \exp(-\frac{1}{2} t^T \Sigma t)$ — for every $\Sigma \succeq 0$, degenerate cases included — and differentiating it once,

$$\partial_{t_i} \varphi_X(t) = - \sum_j \Sigma_{ij} t_j \varphi_X(t). \quad (2)$$

Apply $\partial_{t_{a_2}} \cdots \partial_{t_{a_m}}$ to (2) taken at $i = a_1$. The factor t_j is linear, so Leibniz leaves terms of two kinds: those in which every derivative falls on φ_X , which still carry a factor t_j ; and, for each $s \in \{2, \dots, m\}$, the one in which $\partial_{t_{a_s}}$ falls on t_j and produces δ_{ja_s} . Setting $t = 0$ kills the first kind, leaving

$$\partial_{t_{a_1}} \cdots \partial_{t_{a_m}} \varphi_X \Big|_0 = - \sum_{s=2}^m \Sigma_{a_1 a_s} \partial_{\{t_{a_r} : r \neq 1, s\}} \varphi_X \Big|_0.$$

Rewrite both sides with (1) and multiply through by $(-i)^m$; since $(-i)^2 = -1$ the two minus signs cancel and

$$\mathbb{E}[X_{a_1} \cdots X_{a_m}] = \sum_{s=2}^m \Sigma_{a_1 a_s} \mathbb{E} \left[\prod_{r \neq 1, s} X_{a_r} \right] : \quad (3)$$

pair the first factor with each of the others in turn, and recurse. Induct on m . Both sides vanish for $m = 1$, and for $m = 2$ the claim is the definition of Σ . For $m > 2$, (3) drops the number of factors by two, so parity is preserved: an odd product reduces to $\mathbb{E}[X_a] = 0$, and an even one to a sum over pair partitions of the remaining $2q - 2$ indices, each prefixed by $\Sigma_{a_1 a_s}$ recording the block containing 1. Summing over s enumerates every pair partition of $\{1, \dots, 2q\}$ once. ■

Remark 1.3. Reading (3) with $F(X) = X_{a_2} \cdots X_{a_m}$ and noting that $\partial_j F$ deletes the factor X_j gives

$$\mathbb{E}[X_i F(X)] = \sum_j \Sigma_{ij} \mathbb{E}[\partial_j F(X)],$$

Gaussian integration by parts, which for $n = 1$ and $\Sigma = 1$ is $\mathbb{E}[gF(g)] = \mathbb{E}[F'(g)]$ — Lemma 1.5 of Lectures 1–2, and Stein’s lemma in Lecture 4.

Corollary 1.4. *There are $(2q - 1)!! = (2q - 1)(2q - 3) \cdots 3 \cdot 1$ pair partitions of $2q$ points, and $\mathbb{E}[g^{2q}] = (2q - 1)!!$ for $g \sim \mathcal{N}(0, 1)$.*

Example 1.5. For $q = 2$,

$$\mathbb{E}[X_1 X_2 X_3 X_4] = \Sigma_{12} \Sigma_{34} + \Sigma_{13} \Sigma_{24} + \Sigma_{14} \Sigma_{23},$$

one term for each of the three pair partitions of four points.

The exercises below and in §2.9 are due at the start of Lecture 7.

Exercise 1.6 (TYPE A). Prove Corollary 1.4 — the first point has $2q - 1$ choices of partner; recurse — and check the counts for $2q = 2, 4, 6, 8$. Then write down $\mathbb{E}[g^{20}]$.

Exercise 1.7 (TYPE B). Let $X \sim \mathcal{N}(0, \Sigma)$ on $\mathbb{R}^n$ and A symmetric. Using Theorem 1.2 on $\mathbb{E}[X_i X_j X_k X_l]$, show $\mathbb{E}[X^\top A X] = \text{tr}(A \Sigma)$ and $\text{Var}(X^\top A X) = 2 \text{tr}((A \Sigma)^2)$; deduce $\text{Var} \|X\|^2 = 2 \text{tr} \Sigma^2$. This is the estimate §2.6 needs for $a^\top R a$, with $A = R$.

1.1 The general formula

Theorem 1.2 is the Gaussian instance of a statement holding for any random vector with moments.

Definition 1.8 (Joint cumulants). For $Y_1, \dots, Y_m$ with all moments finite,

$$\kappa(Y_1, \dots, Y_m) = \frac{\partial^m}{\partial t_1 \cdots \partial t_m} \log \mathbb{E}[e^{t_1 Y_1 + \cdots + t_m Y_m}] \Big|_{t=0}.$$

Moments are derivatives of the transform; cumulants are derivatives of its logarithm.

Stated without proof: the moment–cumulant formula. For any $Y_1, \dots, Y_m$ with moments,

$$\mathbb{E}[Y_1 Y_2 \cdots Y_m] = \sum_{\pi \in \mathcal{P}(m)} \prod_{B \in \pi} \kappa(Y_i : i \in B),$$

where $\mathcal{P}(m)$ is the set of *all* partitions of $\{1, \dots, m\}$ into non-empty blocks, of which there are 1, 2, 5, 15, 52, ...

For $m = 2$ this is $\mathbb{E}[Y_1 Y_2] = \mathbb{E} Y_1 \mathbb{E} Y_2 + \text{Cov}(Y_1, Y_2)$, the two partitions of $\{1, 2\}$.

By Definition 1.2 with t in place of it , $\log \mathbb{E}[e^{\langle t, X \rangle}] = \frac{1}{2} t^\top \Sigma t$ is exactly quadratic, so for a centred Gaussian vector

$$\kappa(X_a) = 0, \quad \kappa(X_a, X_b) = \Sigma_{ab}, \quad \kappa(X_{a_1}, \dots, X_{a_k}) = 0 \quad (k \geq 3).$$

A partition contributes nothing unless all of its blocks have size two, and those are the pair partitions. Of the fifteen partitions of a four-element set, three survive.

Remark 1.9. For a non-Gaussian vector the larger blocks do not vanish, $\log Z$ is no longer a polynomial, and the recursion (3) acquires a tail that never terminates. Cumulants then measure the distance from Gaussian, and unlike moments they are additive over independent sums. Lecture 7 takes them as its subject.

2 The semicircle law, twice

Q. Where are the eigenvalues of a large random symmetric matrix?

Theorem 1.2 answers this without any linear algebra, because a trace of a power is a moment of the spectrum.

2.1 Wigner matrices, and the trace as a moment

Definition 2.1 (Gaussian Wigner matrix). M is symmetric $N \times N$ with $\{M_{ij}\}_{i \leq j}$ independent centred Gaussians, $\mathbb{E} M_{ij}^2 = 1/N$ for $i < j$ and $\mathbb{E} M_{ii}^2 = 2/N$. Equivalently, for all indices,

$$\mathbb{E}[M_{ij} M_{kl}] = \frac{1}{N} (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}).$$

The diagonal convention is chosen to make that formula exact; the N diagonal entries are a vanishing fraction of the N^2 and do not affect anything below.

Write $\lambda_1, \dots, \lambda_N$ for the eigenvalues and $\hat{\rho} = \frac{1}{N} \sum_i \delta_{\lambda_i}$ for their empirical distribution. Then

$$\frac{1}{N} \operatorname{tr} M^k = \frac{1}{N} \sum_i \lambda_i^k = \int x^k d\hat{\rho}(x), \quad (4)$$

so Theorem 1.2 applies to spectra.

Imported: the method of moments. A probability law on $\mathbb{R}$ with bounded support is determined by its moments, and if $\int x^k d\mu_N \rightarrow \int x^k d\mu$ for every k with μ compactly supported, then $\mu_N \Rightarrow \mu$. We use this and do not prove it.

2.2 The fourth moment

By (4), $\frac{1}{N} \mathbb{E} \operatorname{tr} M^2 = \frac{1}{N} \sum_{i,j} \mathbb{E} M_{ij}^2 = 1 + \frac{1}{N}$ (Homework 5) — the $\frac{1}{N}$ being the N diagonal entries, which are a vanishing fraction of the N^2 and will be ignored from here on. The next moment is where something happens, and it is where the following distinction starts to matter.

Definition 2.2 (Crossing). Draw the $2q$ points of a pair partition on a line and join each block by an arc above it. Two blocks $\{a, b\}$ and $\{c, d\}$ *cross* if $a < c < b < d$. A pair partition with no crossing blocks is *non-crossing*; there are 1, 2, 5, 14, ... of them on 2, 4, 6, 8 points, against 1, 3, 15, 105 pair partitions in total.

Proposition 2.3. $\frac{1}{N} \mathbb{E} \operatorname{tr} M^4 = 2 + O(1/N)$.

Proof. Expand and apply Theorem 1.2 to each term:

$$\frac{1}{N} \mathbb{E} \operatorname{tr} M^4 = \frac{1}{N} \sum_{i,j,k,l} \mathbb{E}[M_{ij}M_{jk}M_{kl}M_{li}],$$

three pair partitions per term. A pair partition forces index identifications through Definition 2.1; if f indices survive, it contributes $\frac{1}{N} \cdot N^f \cdot \frac{1}{N^2} = N^{f-3}$.

First with second, third with fourth. $\mathbb{E}[M_{ij}M_{jk}]$ is $\frac{1}{N}\delta_{ik}$ up to lower order, and likewise $\mathbb{E}[M_{kl}M_{li}] = \frac{1}{N}\delta_{ki} + \dots$. The single constraint is $i = k$, leaving $f = 3$ and a contribution of 1.

First with fourth, second with third. The same computation gives the constraint $j = l$, so again $f = 3$ and a contribution of 1.

First with third, second with fourth — the crossing one. Now

$$\mathbb{E}[M_{ij}M_{kl}] \mathbb{E}[M_{jk}M_{li}] = \frac{1}{N^2} (\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) (\delta_{jl}\delta_{ki} + \delta_{ji}\delta_{kl}),$$

whose largest term is $\delta_{ik}\delta_{jl}$: two constraints rather than one, so $f = 2$ and the contribution is N^{-1} . Every remaining term forces all four indices equal and contributes N^{-2} . ■

Remark 2.4. Setting $i = k$ and $j = l$ makes $M_{jk} = M_{ji} = M_{ij}$, $M_{kl} = M_{ij}$ and $M_{li} = M_{ij}$: all four factors become the same entry of M . That is what a crossing costs.

Remark 2.5 (Real versus complex). The two δ 's in Definition 2.1 come from $M_{ij} = M_{ji}$. A complex Hermitian matrix has $\mathbb{E}[H_{ij}H_{kl}] = \frac{1}{N}\delta_{il}\delta_{jk}$, one term instead of two, and then the crossing block leaves $f = 1$ and costs N^{-2} . The figure usually quoted is the complex one; Homework 5 does the count.

2.3 The proof, in steps

Proposition 2.3 is the whole mechanism; what remains is bookkeeping and one identification. The argument runs in six steps, of which this lecture does the first two and Homework 5 does the rest.

1. **Trace to moment.** $\frac{1}{N} \operatorname{tr} M^{2k} = \int x^{2k} d\hat{\rho}$, which is (4).
2. **Wick.** Expanding the trace and applying Theorem 1.2 turns $\frac{1}{N} \mathbb{E} \operatorname{tr} M^{2k}$ into a sum over pair partitions of the $2k$ factors.
3. **Count indices, by reading the trace as a walk.** An index tuple $(i_1, \dots, i_{2k})$ is a closed walk of length $2k$, its r th step traversing the edge $\{i_r, i_{r+1}\}$. Since $\mathbb{E}[M_{ab}M_{cd}] \neq 0$ only when $\{a, b\} = \{c, d\}$, a surviving pair partition matches steps traversing the *same* edge, so the walk uses at most k edges, each twice. Its edge graph is connected, so $v \leq e + 1 \leq k + 1$, and a walk with v free indices contributes N^{v-k-1} : **only walks along a tree with k edges survive.** The counting is nothing but *a tree is the sparsest connected graph.* (Homework 5.)
4. **Limit in expectation.** Hence $\frac{1}{N} \mathbb{E} \operatorname{tr} M^{2k} \rightarrow C_k$, the number of such tree walks up to relabelling — equivalently the number of non-crossing pair partitions of $2k$ points. Odd moments vanish.

- 5. Identify C_k .** This is the step to do carefully, and it is not done by recognizing 1, 2, 5, 14. Condition on the block containing the point 1. If that block is $\{1, b\}$ then the points strictly between 1 and b must pair among themselves, since a block leaving that interval would cross $\{1, b\}$; so b is even, say $b = 2j + 2$, and the remaining points split into an interval of $2j$ and an interval of $2(k - 1 - j)$, each to be filled non-crossingly and independently. Therefore

$$C_k = \sum_{j=0}^{k-1} C_j C_{k-1-j}, \quad C_0 = 1,$$

so the generating function $C(w) = \sum_{k \geq 0} C_k w^k$ satisfies $C = 1 + wC^2$, giving $C(w) = \frac{1 - \sqrt{1 - 4w}}{2w}$ and $C_k = \frac{1}{k+1} \binom{2k}{k}$: the Catalan numbers. (*Homework 5 also counts the tree walks of step 3 directly, by encoding each as a Dyck path.*) Separately, the substitution $x = 2 \sin \theta$ evaluates

$$\int_{-2}^2 x^{2k} \rho_{sc}(x) dx = \frac{1}{k+1} \binom{2k}{k},$$

so the two sides agree. (*Homework 5 does both computations.*)

- 6. From expectation to probability, and from moments to the law.** $\text{Var}(\frac{1}{N} \text{tr } M^{2k}) = O(N^{-2})$, so the empirical moments converge in probability and not merely on average; and the imported moment method converts convergence of all moments into weak convergence.

Theorem 2.6 (Wigner). $\hat{\rho} \Rightarrow \rho_{sc}$ in probability, where $\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{4 - x^2}$ on $[-2, 2]$.

Remark 2.7. Gluing the edges of a $2k$ -gon according to P builds a surface whose faces are the surviving indices, and Euler's formula makes that count maximal exactly for planar gluings. This is the usual proof of step 3. We will not use it.

2.4 What that proof cannot do

- (i) It never produced a **density**. It produced a sequence of numbers that we recognized; perturb the model and there is nothing to recognize.
- (ii) It requires **all moments finite**.
- (iii) It gives **no local information** — nothing about one eigenvalue, a gap, or the edge.
- (iv) It does not survive **deformation**: unequal variances, or a rank-one spike, and the combinatorics does not adapt.

All four are repaired below.

2.5 The Stieltjes transform

Definition 2.8. For a probability law μ on $\mathbb{R}$ and $z \in \mathbb{C}$ with $\Im z > 0$,

$$m_\mu(z) = \int \frac{d\mu(\lambda)}{\lambda - z}.$$

It is analytic on the upper half plane, satisfies $|m_\mu(z)| \leq 1/\Im z$, and $m_\mu(z) = -1/z + O(|z|^{-2})$ as $z \rightarrow \infty$ — the last of which will pin down a branch.

Stieltjes inversion. If μ has a continuous density ρ near λ then $\rho(\lambda) = \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \Im m_\mu(\lambda + i\varepsilon)$. Homework 6 proves it and checks it on two examples.

The moment method has no analogue of this step: it returns the density itself rather than a sequence of numbers to be recognized. The bridge to matrices is immediate from (4):

$$m_{\hat{\rho}}(z) = \frac{1}{N} \operatorname{tr} (M - z)^{-1}. \quad (5)$$

Remark 2.9. The right side of (5) was computed numerically in Problem Set 1, where it was the effective degrees of freedom of ridge regression. Same object, week one.

2.6 The self-consistent equation

Write $m_N(z) = \frac{1}{N} \mathbb{E} \operatorname{tr} (M - z)^{-1}$. Let $M^{(1)}$ be M with its first row and column deleted, and let $a \in \mathbb{R}^{N-1}$ be the first column of M with its first entry deleted.

Lemma 2.10 (Schur complement). $((M - z)^{-1})_{11} = \frac{1}{M_{11} - z - a^\top (M^{(1)} - z)^{-1} a}$.

This is the block-inversion identity behind Corollary 2.5 of Lectures 1–2.

Now estimate the three pieces. Put $R = (M^{(1)} - z)^{-1}$, which is a function of $M^{(1)}$ alone and therefore independent of a , and note $\mathbb{E}[a_r a_s] = \delta_{rs}/N$.

- (a) $M_{11} = O_{\mathbb{P}}(N^{-1/2})$, since $\operatorname{Var} M_{11} = 2/N$.
- (b) $\mathbb{E}[a^\top R a \mid M^{(1)}] = \frac{1}{N} \operatorname{tr} R$, and by Theorem 1.2 its conditional variance is $\frac{2}{N^2} \operatorname{tr}(R^2)$. Since $\|R\| \leq (\Im z)^{-1}$ we have $|\operatorname{tr}(R^2)| \leq N(\Im z)^{-2}$, so the variance is $O(1/N)$. This is the only place probability enters.
- (c) $\frac{1}{N} \operatorname{tr} R$ differs from $m_N(z)$ by $O(1/N)$, since deleting one row and column of an $N \times N$ matrix moves a normalized resolvent trace by that much.

All diagonal entries of $(M - z)^{-1}$ have the same law, so $m_N(z) = \mathbb{E}[(M - z)^{-1}]_{11}$. Feeding (a)–(c) into Lemma 2.10 and passing to the limit,

$$m(z) = \frac{1}{-z - m(z)}, \quad \text{that is} \quad m^2 + zm + 1 = 0. \quad (6)$$

Asserted. That $\frac{1}{N} \operatorname{tr}(M - z)^{-1}$ concentrates around its mean, so that the limit in (6) may be taken inside, is a concentration estimate we do not prove.

Equation (6) is a fixed-point equation for the answer, derived without knowing the answer. The moment method had to be told what to recognize.

2.7 Solving, and the comparison

Both roots of (6) are $\frac{-z \pm \sqrt{z^2 - 4}}{2}$. As $z \rightarrow \infty$, $\sqrt{z^2 - 4} = z - 2/z + O(|z|^{-3})$, so only the $+$ root satisfies $m \sim -1/z$:

$$m(z) = \frac{-z + \sqrt{z^2 - 4}}{2}.$$

For $\lambda \in (-2, 2)$ we have $\sqrt{\lambda^2 - 4} = i\sqrt{4 - \lambda^2}$, so $\Im m(\lambda + i0) = \frac{1}{2}\sqrt{4 - \lambda^2}$, and Stieltjes inversion returns $\rho_{sc}(\lambda) = \frac{1}{2\pi}\sqrt{4 - \lambda^2}$. That is Theorem 2.6 again.

Remark 2.11 (The two proofs are the same equation). Expanding Definition 2.8 for large $|z|$ gives $m(z) = -\sum_{k \geq 0} C_k z^{-(2k+1)}$, so with $w = z^{-2}$ we have $-zm(z) = C(w)$. Substituting into the Catalan relation $C = 1 + wC^2$ from step 5 returns $-zm = 1 + m^2$, which is (6). *The recursion that counts non-crossing pair partitions and the self-consistent equation for the resolvent are one equation written twice.*

Return to the four complaints of §2.4. The density came out directly. Only the second moment of the entries was used, so nothing needed all moments. Unequal variances $\mathbb{E}M_{ij}^2 = S_{ij}/N$ turn (6) into a system of N coupled equations rather than one, which is a change of size and not of method. A rank-one spike is a rank-one perturbation of the resolvent. And shrinking $\Im z$ toward the real axis, rather than holding it fixed, is precisely what converts a global law into a local one.

2.8 * Counting the walks, and identifying the limit

Not lectured. These exercises supply steps 3, 5 and 6 of §2.3.

Exercise 2.12 (TYPE A). Show $\frac{1}{N}\mathbb{E} \operatorname{tr} M^2 = 1 + \frac{1}{N}$ and $\frac{1}{N}\mathbb{E} \operatorname{tr} M^{2k+1} = 0$. *One line each; the second needs no computation — say why.*

Exercise 2.13 (TYPE B). (Step 3, first half.) Expanding $\frac{1}{N}\mathbb{E} \operatorname{tr} M^{2k}$ and applying Theorem 1.2, use $\mathbb{E}[M_{ab}M_{cd}] \neq 0 \Rightarrow \{a, b\} = \{c, d\}$ to show a pair partition contributes nothing unless each block pairs two steps traversing the **same edge**. Deduce that a contributing walk uses at most k distinct edges, each traversed exactly twice in the leading terms. *Steps with $i_r = i_{r+1}$, and edges used four or more times, both lower the vertex count and are lower order.*

Exercise 2.14 (TYPE B). (Step 3, second half.) Let G be the edge graph of a contributing walk, with v vertices and e edges. The walk is closed, so G is connected and $e \geq v - 1$; combine with $e \leq k$ to get $v \leq k + 1$, with equality exactly when G is a tree with k edges each traversed twice. Then count: $\sim N^v$ index tuples, k covariances of size $1/N$, and a $1/N$ in front, so the contribution is N^{v-k-1} and only tree walks survive.

Exercise 2.15 (TYPE C). (Step 5, one route.) Let C_n be the number of non-crossing pair partitions of $2n$ points and condition on the block $\{1, b\}$ containing the point 1. Show the points strictly between must pair among themselves, so $b = 2j + 2$ is even and the rest splits into intervals of $2j$ and $2(n - 1 - j)$; deduce $C_n = \sum_{j < n} C_j C_{n-1-j}$ with $C_0 = 1$, and compute 1, 1, 2, 5, 14, 42. Then show the generating function satisfies $C = 1 + wC^2$, solve, and take the root finite at $w = 0$. *Expanding it gives $C_n = \frac{1}{n+1} \binom{2n}{n}$; you may quote that.*

Exercise 2.16 (TYPE C). (Step 5, the other route.) Count the tree walks directly. Fix a tree with k edges and a start vertex, and record the walk as $2k$ symbols: $+1$ for a step onto a vertex *not visited before*, -1 otherwise. Show every partial sum is ≥ 0 and the total is 0 — a **Dyck path** — and that the walk is recoverable, so the correspondence is a bijection. Conclude that the number of contributing walks on N labelled vertices is exactly $C_k N(N - 1) \cdots (N - k)$, hence $\frac{1}{N}\mathbb{E} \operatorname{tr} M^{2k} \rightarrow C_k$.

The idea is the encoding: forget where the walk is, record only whether it is going somewhere new or coming back. Check the count numerically for $k \leq 3$, $N \leq 6$; it is exact, not asymptotic.

Both routes land on the same closed form, and so does the density: substituting $x = 2 \sin \theta$ and evaluating a Beta integral gives $\int_{-2}^2 x^{2k} \rho_{sc} = \frac{1}{k+1} \binom{2k}{k}$. Nothing was matched by eye.

Exercise 2.17 (TYPE C). For a *complex Hermitian* Wigner matrix, $\mathbb{E}[H_{ab}H_{cd}] = \frac{1}{N}\delta_{ad}\delta_{bc}$ — one delta term where the real case has two. Redo the crossing pair partition of $\text{tr } H^4$ and show it now leaves one free index, so the correction is $O(N^{-2})$ rather than $O(N^{-1})$. *One extra way to glue two entries together is the whole difference.*

Exercise 2.18 (TYPE A). Plot the empirical spectral distribution at $N = 50$ and $N = 2000$ against ρ_{sc} , then once more with entries $\pm 1/\sqrt{N}$, which are not Gaussian and for which Theorem 1.2 is false. One figure. *Universality, weeks before the course explains it.*

Homework 5 counts the pair partitions, computes $\frac{1}{N}\mathbb{E} \text{tr } M^2$ and the exponent $N^{f(P)-k-1}$, and checks $\int x^{2k}\rho_{sc} = C_k$; Homework 6 proves Stieltjes inversion, supplies the estimates (b) and (c) above, and rederives Marchenko–Pastur by the same route.