3

More on Gaussian Matrices*

In the previous chapter, we dealt with the simplest of all Gaussian matrix ensembles, where entries are real, Gaussian random variables, and global symmetry is imposed. It was pointed out by Dyson that there exist precisely three division rings that contain the real numbers, namely, the real themselves, the complex numbers and the quaternions. He showed that this fact implies that there are only three acceptable ensembles of Gaussian random matrices with real eigenvalues: GOE, GUE and GSE. Each is associated with a Dyson index called β (1, 2 and 4, respectively) and except for this difference in β almost all of the results in this book (and many more) apply to the three ensembles. In particular their moments and eigenvalue density are the same as $N \to \infty$, while correlations and deviations from the asymptotic formulas follow families of laws with β as a parameter. In this chapter we will review the other two ensembles (GUE and GSE),[1] and also discuss the general moments of Gaussian random matrices, for which some interesting mathematical tools are available, that are useful beyond RMT.

3.1 Other Gaussian Ensembles

3.1.1 Complex Hermitian Matrices

For matrices with complex entries, the analog of a symmetric matrix is a (complex) Hermitian matrix. It satisfies $\mathbf{A}^\dagger = \mathbf{A}$ where the dagger operator is the combination of matrix transposition and complex conjugation. There are two important reasons to study complex Hermitian matrices. First they appear in many applications, especially in quantum mechanics. There, the energy and other observables are mapped into Hermitian operators, or Hermitian matrices for systems with a finite number of states. The first large N result of random matrix theory is the Wigner semi-circle law. As recalled in the introduction to Chapter 2, it was obtained by Wigner as he modeled the energy levels of complex heavy nuclei as a random Hermitian matrix.

[1] More recently, it was shown how ensembles with an arbitrary value of β can be constructed, see Dumitriu and Edelman [2002], Allez et al. [2012].

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

The other reason Hermitian matrices are important is mathematical. In the large N limit, the three ensembles (real, complex and quaternionic (see below)) behave the same way. But for finite N, computations and proofs are much simpler in the complex case. The main reason is that the Vandermonde determinant which we will introduce in Section 5.1.4 is easier to manipulate in the complex case. For this reason, most mathematicians discuss the complex Hermitian case first and treat the real and quaternionic cases as extensions. In this book we want to stay close to applications in data science and statistical physics, so we will discuss complex matrices only in the present chapter. In the rest of the book we will indicate in footnotes how to extend the result to complex Hermitian matrices.

A complex Hermitian matrix $\mathbf{A}$ has real eigenvalues and it can be diagonalized with a suitable unitary matrix $\mathbf{U}$. A unitary matrix satisfies $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$. So $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^\dagger$, with Λ the diagonal matrix containing its N eigenvalues.

We want to build the complex Wigner matrix: a Hermitian matrix with IID Gaussian entries. We will choose a construction that has unitary invariance for every N. Let us study the unitary invariance of complex Gaussian vectors. First we need to define a complex Gaussian variable.

We say that the complex variable z is centered Gaussian with variance σ^2 if $z = x_\mathrm{r} + \mathrm{i}\, x_\mathrm{i}$ where x_r and x_i are centered Gaussian variables of variance $\sigma^2/2$. We have

$$\mathbb{E}[|z|^2] = \mathbb{E}[x_\mathrm{r}^2] + \mathbb{E}[x_\mathrm{i}^2] = \sigma^2. \tag{3.1}$$

A white complex Gaussian vector $\mathbf{x}$ is a vector whose components are IID complex centered Gaussians. Consider $\mathbf{y} = \mathbf{U}\mathbf{x}$ where $\mathbf{U}$ is a unitary matrix. Each of the components is a linear combination of Gaussian variables so $\mathbf{y}$ is Gaussian. It is relatively straightforward to show that each component has the same variance σ^2 and that there is no covariance between different components. Hence $\mathbf{y}$ is also a white Gaussian vector. The ensemble of a white complex Gaussian vector is invariant under unitary transformation.

To define the Hermitian Wigner matrix, we first define a (non-symmetric) square matrix $\mathbf{H}$ whose entries are centered complex Gaussian numbers and let $\mathbf{X}$ be the Hermitian matrix defined by

$$\mathbf{X} = \mathbf{H} + \mathbf{H}^\dagger. \tag{3.2}$$

If we repeat the arguments of Section 2.2.2, we can show that the ensemble of $\mathbf{X}$ is invariant under unitary transformation: $\mathbf{U}\mathbf{X}\mathbf{U}^\dagger \overset{\text{in law}}{=} \mathbf{X}$.

We did not specify the variance of the elements of $\mathbf{H}$. We would like $\mathbf{X}$ to be normalized as $\tau(\mathbf{X}^2) = \sigma^2 + O(1/N)$. Choosing the variance of the $\mathbf{H}$ as $\mathbb{E}[|\mathbf{H}_{ij}|^2] = 1/(2N)$ achieves precisely that.

The Hermitian matrix $\mathbf{X}$ has real diagonal elements with $\mathbb{E}[\mathbf{X}_{ii}^2] = 1/N$ and off-diagonal elements that are complex Gaussian with $\mathbb{E}[|\mathbf{X}_{ij}|^2] = 1/N$. In other words the real and imaginary parts of the off-diagonal elements of $\mathbf{X}$ have variance $1/(2N)$. We can put

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

all this information together in the joint law of the matrix elements of the Hermitian matrix **H**:

$$P(\{X_{ij}\}) \propto \exp\left\{-\frac{N}{2\sigma^2}\operatorname{Tr}\mathbf{X}^2\right\}. \tag{3.3}$$

This law is identical to the real symmetric case (Eq. 2.17) up to a factor of 2. We can then write both the symmetric and the Hermitian case as

$$P(\{X_{ij}\}) \propto \exp\left\{-\frac{\beta N}{4\sigma^2}\operatorname{Tr}\mathbf{X}^2\right\}, \tag{3.4}$$

where β is 1 or 2 respectively.

The complex Hermitian Wigner ensemble is called the *Gaussian unitary ensemble* or GUE.

The results of the previous chapter apply equally to the real symmetric and the complex Hermitian case. Both the self-consistent equation for the Stieltjes transform and the counting of non-crossing pair partitions (see next section) rely on the independence of the elements of the matrix and on the fact that $\mathbb{E}[|\mathbf{X}_{ij}|^2] = 1/N$, true in both cases. We then have that the Stieltjes transform of the two ensembles is the same and they have exactly the same semi-circle distribution of eigenvalues in the large N limit. The same will be true for the quaternionic case ($\beta = 4$) in the next section, and in fact for all values of β provided $N\beta \to \infty$ when $N \to \infty$, see Section 5.3.1:

$$\rho_\beta(\lambda) = \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2}, \quad -2\sigma \le \lambda \le 2\sigma. \tag{3.5}$$

3.1.2 Quaternionic Hermitian Matrices

We will define here the quaternionic Hermitian matrices and the GSE. There are many fewer applications of quaternionic matrices than the more common real or complex matrices. We include this discussion here for completeness. In the literature the link between symplectic matrices and quaternions can be quite obscure for the novice reader. Except for the existence of an ensemble of matrices with $\beta = 4$ we will never refer to quaternionic matrices after this section, which can safely be skipped.

Quaternions are non-commutative extensions of the real and complex numbers. They are written as real linear combinations of the real number 1 and three abstract non-commuting objects $(\mathrm{i},\mathrm{j},\mathrm{k})$ satisfying

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1 \quad \Rightarrow \quad \mathrm{ij} = -\mathrm{ji} = \mathrm{k}, \quad \mathrm{jk} = -\mathrm{kj} = \mathrm{i}, \quad \mathrm{ki} = -\mathrm{ik} = \mathrm{j}. \tag{3.6}$$

So we can write a quaternion as $h = x_\mathrm{r} + \mathrm{i}\, x_\mathrm{i} + \mathrm{j}\, x_\mathrm{j} + \mathrm{k}\, x_\mathrm{k}$. If only x_r is non-zero we say that h is real. We define the quaternionic conjugation as $1^* = 1, \mathrm{i}^* = -\mathrm{i}, \mathrm{j}^* = -\mathrm{j}, \mathrm{k}^* = -\mathrm{k}$ so that the norm $|h|^2 := hh^* = x_\mathrm{r}^2 + x_\mathrm{i}^2 + x_\mathrm{j}^2 + x_\mathrm{k}^2$ is always real and non-negative. The abstract objects i, j and k can be represented as 2×2 complex matrices:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathrm{i} = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}, \quad \mathrm{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathrm{k} = \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \tag{3.7}$$

where the i in the matrices is now the usual unit imaginary number.

Quaternions share all the algebraic properties of real and complex numbers except for commutativity (they form a *division ring*). Since matrices in general do not commute, matrices built out of quaternions behave like real or complex matrices.

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

A Hermitian quaternionic matrix is a square matrix $\mathbf{A}$ whose elements are quaternions and satisfy $\mathbf{A} = \mathbf{A}^\dagger$. Here the dagger operator is the combination of matrix transposition and quaternionic conjugation. They are diagonalizable and their eigenvalues are real. Matrices that diagonalize Hermitian quaternionic matrices are called *symplectic*. Written in terms of quaternions they satisfy $\mathbf{S}\mathbf{S}^\dagger = \mathbf{1}$.

Given representation of quaternions as 2×2 complex matrices, an $N \times N$ quaternionic Hermitian matrix $\mathbf{A}$ can be written as a $2N \times 2N$ complex matrix $\mathbf{Q}(\mathbf{A})$. We choose a representation where

$$\mathbf{Z} := \mathbf{Q}(\mathbf{1}\mathrm{j}) = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}. \tag{3.8}$$

For a $2N \times 2N$ complex matrix $\mathbf{Q}$ to be the representation of a quaternionic Hermitian matrix it has to have two properties. First, quaternionic conjugation acts just like Hermitian conjugation so $\mathbf{Q}^\dagger = \mathbf{Q}$. Second, it has to be expressible as a real linear combination of unit quaternions. One can show that such matrices (and only them) satisfy

$$\mathbf{Q}^{\mathrm{R}} := \mathbf{Z}\mathbf{Q}^T\mathbf{Z}^{-1} = \mathbf{Q}^\dagger, \tag{3.9}$$

where $\mathbf{Q}^{\mathrm{R}}$ is called the dual of $\mathbf{Q}$. In other words an $N \times N$ Hermitian quaternionic matrix corresponds to a $2N \times 2N$ self-dual Hermitian matrix (i.e. $\mathbf{Q} = \mathbf{Q}^\dagger = \mathbf{Q}^{\mathrm{R}}$). In this $2N \times 2N$ representation symplectic matrices are complex matrices satisfying

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{S}\mathbf{S}^{\mathrm{R}} = \mathbf{1}. \tag{3.10}$$

To recap, a $2N \times 2N$ Hermitian self-dual matrix $\mathbf{Q}$ can be diagonalized by a symplectic matrix $\mathbf{S}$. Its $2N$ eigenvalues are real and they occur in pairs as they are the N eigenvalues of the equivalent Hermitian quaternionic $N \times N$ matrix.

We can now define the third Gaussian matrix ensemble, namely the *Gaussian symplectic ensemble* (GSE) consisting of Hermitian quaternionic matrices whose off-diagonal elements are quaternions with Gaussian distribution of zero mean and variance $\mathbb{E}[|X_{ij}|^2] = 1/N$. This means that each of the four components of each X_{ij} is a Gaussian number of zero mean and variance $1/(4N)$. The diagonal elements of $\mathbf{X}$ are real Gaussian numbers with zero mean and variance $1/(2N)$. As usual $\mathbf{X}_{ij} = \mathbf{X}^*_{ji}$ so only the upper (or lower) triangular elements are independent. The joint law for the elements of a GSE matrix with variance $\tau(\mathbf{X}^2) = \sigma^2$ is given by

$$P(\{X_{ij}\}) \propto \exp\left\{-\frac{N}{\sigma^2}\operatorname{Tr}\mathbf{X}^2\right\}, \tag{3.11}$$

which we identify with Eq. (3.4) with $\beta = 4$. This parameter $\beta = 4$ is a fundamental property of the symplectic group and will consistently appear in contrast with the orthogonal and unitary cases, $\beta = 1$ and $\beta = 2$ (see Section 5.1.4).

The parameter β can be interpreted as the randomness in the norm of the matrix elements. More precisely, we have

$$|X_{ij}|^2 = \begin{cases} x_{\mathrm{r}}^2 & \text{for real symmetric,} \\ x_{\mathrm{r}}^2 + x_{\mathrm{i}}^2 & \text{for complex Hermitian,} \\ x_{\mathrm{r}}^2 + x_{\mathrm{i}}^2 + x_{\mathrm{j}}^2 + x_{\mathrm{k}}^2 & \text{for quaternionic Hermitian,} \end{cases} \tag{3.12}$$

where $x_{\mathrm{r}}, x_{\mathrm{i}}, x_{\mathrm{j}}, x_{\mathrm{k}}$ are real Gaussian numbers such that $\mathbb{E}[|X_{ij}|^2] = 1$. We see that the fluctuations of $|X_{ij}|^2$ decrease with β (precisely $\mathbb{V}[|X_{ij}|^2] = 2/\beta$). By the law of large numbers (LLN), in the $\beta \to \infty$ limit (if such an ensemble existed) we would have $|X_{ij}|^2 = 1$ with no fluctuations.

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

Exercise 3.1.1 Quaternionic matrices of size one

The four matrices in Eq. (3.7) can be thought of as the 2×2 complex representations of the four unit quaternions.

(a) Define $\mathbf{Z} :=$ j and compute $\mathbf{Z}^{-1}$.
(b) Show that for all four matrices $\mathbf{Q}$, we have $\mathbf{Z}\mathbf{Q}^T\mathbf{Z}^{-1} = \mathbf{Q}^\dagger$ where the dagger here is the usual transpose plus complex conjugation.
(c) Convince yourself that, by linearity, any $\mathbf{Q}$ that is a real linear combination of the 2×2 matrices i, j, k and 1 must satisfy $\mathbf{Z}\mathbf{Q}^T\mathbf{Z}^{-1} = \mathbf{Q}^\dagger$.
(d) Give an example of a matrix $\mathbf{Q}$ that does not satisfy $\mathbf{Z}\mathbf{Q}^T\mathbf{Z}^{-1} = \mathbf{Q}^\dagger$.

3.1.3 The Ginibre Ensemble

The Gaussian orthogonal ensemble is such that all matrix elements of $\mathbf{X}$ are IID Gaussian, but with the strong constraint that $X_{ij} = X_{ji}$, which makes sure that all eigenvalues of $\mathbf{X}$ are real. What happens if we drop this constraint and consider a square matrix $\mathbf{H}$ with independent entries? In this case, one may choose two different routes, depending on the context.

- One route is simply to allow eigenvalues to be complex numbers. One can then study the eigenvalue distribution in the complex plane, so the distribution becomes a two-dimensional density. Some of the tools introduced in the previous chapter, such as the Sokhotski–Plemelj formula, can be generalized to complex eigenvalues. The final result is called the Girko circular law: the density of eigenvalues is constant within a disk centered at zero and of radius σ (see Fig. 3.1). In the general case where $\mathbb{E}[H_{ij}H_{ji}] = \rho\sigma^2$, the eigenvalues are confined within an ellipse of half-width $(1+\rho)\sigma$ along the real axis and $(1-\rho)\sigma$ in the imaginary direction, interpolating between a circle for $\rho = 0$ (independent entries) and a line segment on the real axis of length 4σ for $\rho = 1$ (symmetric matrices).
- The other route is to focus on the singular values of $\mathbf{H}$. One should thus study the real eigenvalues of $\mathbf{H}^T\mathbf{H}$ when $\mathbf{H}$ is a square random matrix made of independent Gaussian elements. This is precisely the Wishart problem that we will study in Chapter 4, for the special parameter value $q = 1$. Calling s the square-root of these real eigenvalues, the final result is a quarter-circle:

$$\rho(s) = \frac{\sqrt{4\sigma^2 - s^2}}{\pi\sigma^2}; \qquad s \in (0, 2\sigma). \tag{3.13}$$

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

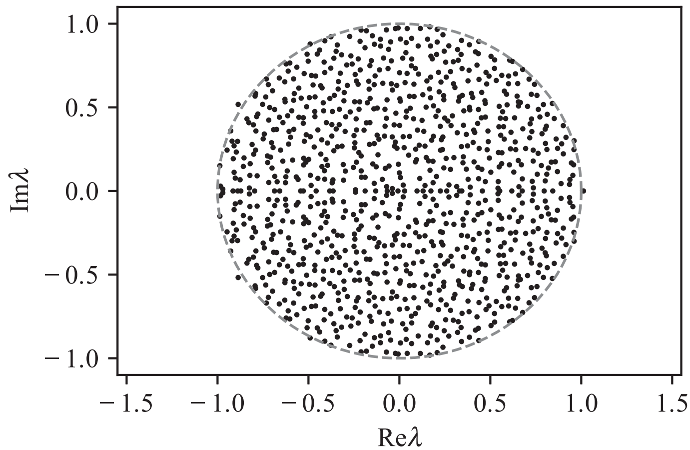

Figure 3.1 Complex eigenvalues of a random $N = 1000$ matrix taken from the Gaussian Ginibre ensemble, i.e. a non-symmetric matrix with IID Gaussian elements with variance $\sigma^2 = 1/N$. The dash line corresponds to the circle $|\lambda|^2 = 1$. As $N \to \infty$ the density becomes uniform in the complex unit disk. This distribution is called the circle law or sometimes, more accurately, the disk law.

Exercise 3.1.2 Three quarter-circle laws

Let $\mathbf{H}$ be a (non-symmetric) square matrix of size N whose entries are IID Gaussian random variable of variance σ^2/N. Then as a simple consequence of the above discussion the following three sets of numbers are distributed according to the quarter-circle law (3.13) in the large N limit. Define

$$w_i = |\lambda_i| \text{ where } \{\lambda_i\} \text{ are the eigenvalues of } \frac{\mathbf{H}+\mathbf{H}^T}{\sqrt{2}},$$

$$r_i = 2|\operatorname{Re}\lambda_i| \text{ where } \{\lambda_i\} \text{ are the eigenvalues of } \mathbf{H},$$

$$s_i = \sqrt{\lambda_i} \text{ where } \{\lambda_i\} \text{ are the eigenvalues of } \mathbf{H}\mathbf{H}^T.$$

(a) Generate a large matrix $\mathbf{H}$ with say $N = 1000$ and $\sigma^2 = 1$ and plot the histogram of the three above sets.

(b) Although these three sets of numbers converge to the same distribution there is no simple relation between them. In particular they are not equal. For a moderate N (10 or 20) examine the three sets and realize that they are all different.

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

3.2 Moments and Non-Crossing Pair Partitions

3.2.1 Fourth Moment of a Wigner Matrix

We have stated in Section 2.2.1 that for a Wigner matrix we have $\tau(\mathbf{X}^4) = 2\sigma^4$. We will now compute this fourth moment directly and then develop a technique to compute all other moments.

We have

$$\tau(\mathbf{X}^4) = \frac{1}{N}\mathbb{E}[\mathrm{Tr}(\mathbf{X}^4)] = \frac{1}{N}\sum_{i,j,k,l}\mathbb{E}[\mathbf{X}_{ij}\mathbf{X}_{jk}\mathbf{X}_{kl}\mathbf{X}_{li}]. \tag{3.14}$$

Recall that $(\mathbf{X}_{ij} : 1 \le i \le j \le N)$ are independent Gaussian random variables of mean zero. So for the expectations in the above sum to be non-zero, each $\mathbf{X}$ entry needs to be equal to another $\mathbf{X}$ entry.[2] There are two possibilities. Either all four are equal or they are equal pairwise. In the following we will not distinguish between diagonal and off-diagonal terms; as there are many more off-diagonal terms these terms always dominate.

(1) If $\mathbf{X}_{ij} = \mathbf{X}_{jk} = \mathbf{X}_{kl} = \mathbf{X}_{li}$, then

$$\mathbb{E}[\mathbf{X}_{ij}\mathbf{X}_{jk}\mathbf{X}_{kl}\mathbf{X}_{li}] = \frac{3\sigma^4}{N^2}, \tag{3.15}$$

and there are N^2 of them. Thus the total contribution from these terms is

$$\frac{1}{N}N^2\frac{3\sigma^4}{N^2} = \frac{3\sigma^4}{N} \to 0. \tag{3.16}$$

(2) Suppose there are two different pairs. Then there are three possibilities (see Fig. 3.2):

(i) $\mathbf{X}_{ij} = \mathbf{X}_{jk}$, $\mathbf{X}_{kl} = \mathbf{X}_{li}$, and $\mathbf{X}_{ij}$ is different than $\mathbf{X}_{li}$ (i.e. $j \neq l$). Then

$$(3.14) = \frac{1}{N}\sum_{i,j\neq l}\mathbb{E}[\mathbf{X}_{ij}^2\mathbf{X}_{il}^2] = \frac{1}{N}(N^3 - N^2)\left(\frac{\sigma^2}{N}\right)^2 \to \sigma^4 \tag{3.17}$$

as $N \to \infty$.

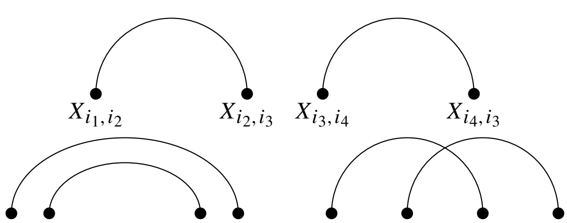

Figure 3.2 Graphical representation of the three terms contributing to $\tau(\mathbf{X}^4)$. The last one is a crossing partition and has a zero contribution.

[2] When we say that $\mathbf{X}_{ij} = \mathbf{X}_{kl}$, we mean that they are the same random variable; given that $\mathbf{X}$ is a symmetric matrix it means either ($i = k$ and $j = l$) or ($i = l$ and $j = k$).

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

(ii) $\mathbf{X}_{ij} = \mathbf{X}_{li}$, $\mathbf{X}_{jk} = \mathbf{X}_{kl}$, and $\mathbf{X}_{ij}$ is different than $\mathbf{X}_{jk}$ (i.e. $i \neq k$). Then

$$(3.14) = \frac{1}{N} \sum_{i \neq k, j} \mathbb{E}[\mathbf{X}_{ij}^2 \mathbf{X}_{jk}^2] = \frac{1}{N}(N^3 - N^2)\left(\frac{\sigma^2}{N}\right)^2 \to \sigma^4 \tag{3.18}$$

as $N \to \infty$.

(iii) $\mathbf{X}_{ij} = \mathbf{X}_{kl}$, $\mathbf{X}_{jk} = \mathbf{X}_{li}$, and $\mathbf{X}_{ij}$ is different than $\mathbf{X}_{jk}$ (i.e. $i \neq k$). Then we must have $i = l$ and $j = k$ from $\mathbf{X}_{ij} = \mathbf{X}_{kl}$, and $i = j$ and $k = l$ from $\mathbf{X}_{jk} = \mathbf{X}_{li}$. This gives a contradiction: there are no such terms.

In sum, we obtain that

$$\tau(\mathbf{X}^4) \to \sigma^4 + \sigma^4 = 2\sigma^4 \tag{3.19}$$

as $N \to \infty$, where the two terms come from the two non-crossing partitions, see Figure 3.2.

In the next technical section, we generalize this calculation to arbitrary moments of $\mathbf{X}$. Odd moments are zero by symmetry. Even moments $\tau(\mathbf{X}^{2k})$ can be written as sums over non-crossing diagrams (non-crossing pair partitions of $2k$ elements), where each such diagram contributes σ^{2k}. So

$$\tau(\mathbf{X}^{2k}) = C_k \sigma^{2k}, \tag{3.20}$$

where C_k are Catalan numbers, the number of such non-crossing diagrams. They satisfy

$$C_k = \sum_{j=1}^{k} C_{j-1} C_{k-j} = \sum_{j=0}^{k-1} C_j C_{k-j-1}, \tag{3.21}$$

with $C_0 = C_1 = 1$, and can be written explicitly as

$$C_k = \frac{1}{k+1}\binom{2k}{k}, \tag{3.22}$$

see Section 3.2.3.

3.2.2 Catalan Numbers: Counting Non-Crossing Pair Partitions

We would like to calculate all moments of $\mathbf{X}$. As written above, all the odd moments $\tau(\mathbf{X}^{2k+1})$ vanish (since the odd moments of a Gaussian random variable vanish). We only need to compute the even moments:

$$\tau(\mathbf{X}^{2k}) = \frac{1}{N}\mathbb{E}\left[\mathrm{Tr}(\mathbf{X}^{2k})\right] = \frac{1}{N} \sum_{i_1, \ldots, i_{2k}} \mathbb{E}\left(\mathbf{X}_{i_1 i_2} \mathbf{X}_{i_2 i_3} \ldots \mathbf{X}_{i_{2k} i_1}\right). \tag{3.23}$$

Since we assume that the elements of $\mathbf{X}$ are Gaussian, we can expand the above expectation value using Wick's theorem using the covariance of the $\{\mathbf{X}_{ij}\}$'s. The matrix $\mathbf{X}$ is symmetric, so we have to keep track of the fact that $\mathbf{X}_{ij}$ is the same variable as $\mathbf{X}_{ji}$. For this reason, using Wick's theorem proves quite tedious and we will not follow this route here.

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

From the Taylor series at infinity of the Stieltjes transform, we expect every even moment of $\mathbf{X}$ to converge to an $O(1)$ number as $N \to \infty$. We will therefore drop any $O(1/N)$ or smaller term as we proceed. In particular the difference of variance between diagonal and off-diagonal elements of $\mathbf{X}$ does not matter to first order in $1/N$.

In Eq. (3.23), each $\mathbf{X}$ entry must be equal to at least one another $\mathbf{X}$ entry, otherwise the expectation is zero. On the other hand, it is easy to show that for the partitions that contain at least one group with > 2 (actually ≥ 4) $\mathbf{X}$ entries that are equal to each other, their total contribution will be of order $O(1/N)$ or smaller (e.g. in case (1) of the previous section). Thus we only need to consider the cases where each $\mathbf{X}$ entry is paired to exactly one other $\mathbf{X}$ entry, which we also referred to as a pair partition.

We need to count the number of types of pairings of $2k$ elements that contribute to $\tau(\mathbf{X}^{2k})$ as $N \to \infty$. We associate to each pairing a diagram. For example, for $k = 3$, we have $5!! = 5 \cdot 3 \cdot 1 = 15$ possible pairings (see Fig. 3.3).

To compute the contribution of each of these pair partitions, we will compute the contribution of *non-crossing* pair partitions and argue that pair partitions with crossings do not contribute in the large N limit. First we need to define what is a non-crossing pair partition of $2k$ elements. A pair partition can be draw as a diagram where the $2k$ elements are points on a line and each point is joined with its pair partner by an arc drawn above that line. If at least two arcs cross each other the partition is called *crossing*, and *non-crossing* otherwise. In Figure 3.3 the five partitions on the left are non-crossing while the ten others are crossing.

In a non-crossing partition of size $2k$, there is always at least one pairing between consecutive points (the smallest arc). If we remove the first such pairing we get a non-crossing pair partition of $2k - 2$ elements. We can proceed in this way until we get to a paring of only two elements: the unique (non-crossing) pair partition contributing to (Fig. 3.4)

$$\tau(\mathbf{X}^2) = \sigma^2. \tag{3.24}$$

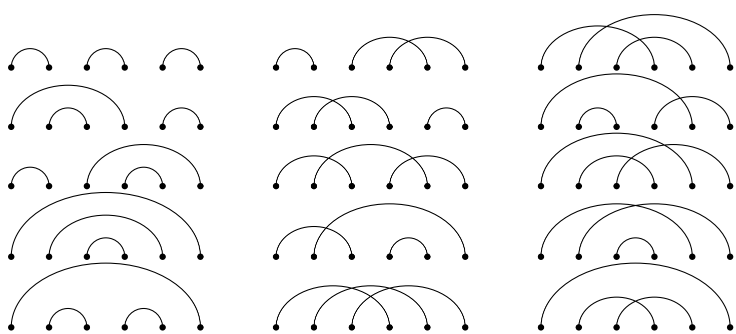

Figure 3.3 Graphical representation of the 15 terms contributing to $\tau(\mathbf{X}^6)$. Only the five on the left are non-crossing and have a non-zero contribution as $N \to \infty$.

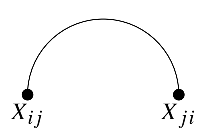

Figure 3.4 Graphical representation of the only term contributing to $\tau(\mathbf{X}^2)$. Note that the indices of two terms are already equal prior to pairing.

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

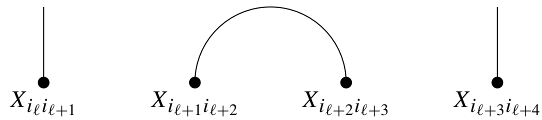

Figure 3.5 Zoom into the smallest arc of a non-crossing partition. The two middle matrices are paired while the other two could be paired together or to other matrices to the left and right respectively. After the pairing of $\mathbf{X}_{i_{\ell+1},i_{\ell+2}}$ and $\mathbf{X}_{i_{\ell+2},i_{\ell+3}}$, we have $i_{\ell+1} = i_{\ell+3}$ and the index $i_{\ell+2}$ is free.

We can use this argument to prove by induction that each non-crossing partition contributes a factor σ^{2k}. In Figure 3.5, consecutive elements $\mathbf{X}_{i_{\ell+1},i_{\ell+2}}$ and $\mathbf{X}_{i_{\ell+2},i_{\ell+3}}$ are paired; we want to evaluate that pair and remove it from the diagram. The variance contributes a factor σ^2/N. We can make two choices for index matching. First consider $i_{\ell+1} = i_{\ell+3}$ and $i_{\ell+2} = i_{\ell+2}$. In that case, the index $i_{\ell+2}$ is free and its summation contributes a factor of N. The identity $i_{\ell+1} = i_{\ell+3}$ means that the previous matrix $\mathbf{X}_{i_\ell,i_{\ell+1}}$ is now linked by matrix multiplication to the following matrix $\mathbf{X}_{i_{\ell+1},i_{\ell+4}}$. In other words we are left with σ^2 times a non-crossing partition of size $2k-2$, which contributes σ^{2k-2} by our induction hypothesis. The other choice of index matching, $i_{\ell+1} = i_{\ell+2} = i_{\ell+3}$, can be viewed as fixing a particular value for $i_{\ell+2}$ and is included in the sum over $i_{\ell+2}$ in the previous index matching. So by induction we do have that each non-crossing pair partition contributes σ^{2k}.

Before we discuss the contribution of crossing pair partitions, let's analyze in terms of powers of N the computation we just did for the non-crossing case. The computation of each term in $\tau(\mathbf{X}^{2k})$ involves $2k$ matrices that have in total $4k$ indices. The trace and the matrix multiplication forces $2k$ equalities among these indices. The normalization of the trace and the k variance terms gives a factor of σ^{2k}/N^{k+1}. To get a result of order 1 we need to be left with $k+1$ free indices whose summation gives a factor of N^{k+1}. Each k pairing imposes a matching between pairs of indices. For the first $k-1$ choice of pairing we managed to match one pair of indices that were already equal. At the last step we matched to pairs of indices that were already equal. Hence in total we added only $k+1$ equality constraints which left us with $k+1$ free indices as needed.

We can now argue that crossing pair partitions do not contribute in the large N limit. For crossing partition it is not possible to choose a matching at every step that matches a pair of indices that are already equal. If we use the previous algorithm of removing at each step the leftmost smallest arc, at some point, the smallest arc will have a crossing and we will be pairing to matrices that share no indices, adding two equality constraints at this step. The result will therefore be down by at least a factor of $1/N$ with respect to the non-crossing case. This argument is not really a proof but an intuition why this might be true.[3]

We can now complete our moments computation. Let

$$C_k := \#\text{ of non-crossing pairings of } 2k \text{ elements.} \tag{3.25}$$

Since every non-crossing pair partition contributes a factor σ^{2k}, summing over all non-crossing pairings we immediately get that

$$\tau(\mathbf{X}^{2k}) = C_k \sigma^{2k}. \tag{3.26}$$

[3] A more rigorous proof can be found in e.g. Anderson et al. [2010], Tao [2012] or Mingo and Speicher [2017]. In this last reference, the authors compute the moments of $\mathbf{X}$ exactly for every N (when $\sigma_{\rm d}^2 = \sigma_{\rm od}^2$).

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

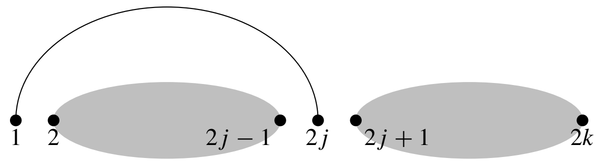

Figure 3.6 In a non-crossing pairing, the paring of site 1 with site $2j$ splits the graph into two disjoint non-crossing parings.

3.2.3 Recursion Relation for Catalan Numbers

In order to compute the Catalan numbers C_k, we will write a recursion relation for them. Take a non-crossing pairing, site 1 is linked to some even site $2j$ (it is easy to see that 1 cannot link to an odd site in order for the partition to be non-crossing). Then the diagram is split into two smaller non-crossing pairings of sizes $2(j-1)$ and $2(k-j)$, respectively (see Fig. 3.6). Thus we get the inductive relation[4]

$$C_k = \sum_{j=1}^{k} C_{j-1} C_{k-j} = \sum_{j=0}^{k-1} C_j C_{k-j-1}, \tag{3.27}$$

where we let $C_0 = C_1 = 1$. One can then prove by induction that C_k is given by the Catalan number:

$$C_k = \frac{1}{k+1} \binom{2k}{k}. \tag{3.28}$$

Using the Taylor series for the Stieltjes transform (2.22), we can use the Catalan number recursion relation to find an equation for the Stieltjes transform of the Wigner ensemble:

$$\mathfrak{g}(z) = \sum_{k=0}^{\infty} \frac{C_k}{z^{2k+1}} \sigma^{2k}. \tag{3.29}$$

Thus, using (3.27), we obtain that

$$\begin{aligned} \mathfrak{g}(z) - \frac{1}{z} &= \sum_{k=1}^{\infty} \frac{\sigma^{2k}}{z^{2k+1}} \left(\sum_{j=0}^{k-1} C_j C_{k-j-1} \right) \\ &= \frac{\sigma^2}{z} \sum_{j=0}^{\infty} \frac{C_j}{z^{2j+1}} \sigma^{2j} \left(\sum_{k=j+1}^{\infty} \frac{C_{k-j-1}}{z^{2(k-j-1)+1}} \sigma^{2(k-j-1)} \right) \\ &= \frac{\sigma^2}{z} \left(\sum_{j=0}^{\infty} \frac{C_j}{z^{2j+1}} \sigma^{2j} \right) \left(\sum_{\ell=0}^{\infty} \frac{C_\ell}{z^{2\ell+1}} \sigma^{2\ell} \right) = \frac{\sigma^2}{z} \mathfrak{g}^2(z), \end{aligned} \tag{3.30}$$

which gives the same self-consistent equation for $\mathfrak{g}(z)$ as in (2.35) and hence the same solution:

$$\mathfrak{g}(z) = \frac{z - z\sqrt{1 - 4\sigma^2/z^2}}{2\sigma^2}. \tag{3.31}$$

[4] Interestingly, this recursion relation is also found in the problem of RNA folding. For deep connections between the physics of RNA and RMT, see Orland and Zee [2002].

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press

The same result could have been derived by substituting the explicit solution for the Catalan number Eq. (3.28) into (3.29), but this route requires knowledge of the Taylor series:

$$\sqrt{1-x} = 1 - \sum_{k=0}^{\infty} \frac{2}{k+1} \binom{2k}{k} \left(\frac{x}{4}\right)^{k+1}. \tag{3.32}$$

Exercise 3.2.1 Non-crossing pair partitions of eight elements

(a) Draw all the non-crossing pair partitions of eight elements. Hint: use the recursion expressed in Figure 3.6.

(b) If $\mathbf{X}$ is a unit Wigner matrix, what is $\tau(\mathbf{X}^8)$?

Bibliographical Notes

- Again, several books cover the content of this chapter, see for example
  - M. L. Mehta. *Random Matrices*. Academic Press, San Diego, 3rd edition, 2004,
  - G. Blower. *Random Matrices: High Dimensional Phenomena*. Cambridge University Press, Cambridge, 2009,
  - B. Eynard, T. Kimura, and S. Ribault. Random matrices. *preprint arXiv:1510.04430*, 2006,

  and in particular for a detailed discussion of the relation between non-crossing and large matrices, see
  - T. Tao. *Topics in Random Matrix Theory*. American Mathematical Society, Providence, Rhode Island, 2012,
  - G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge University Press, Cambridge, 2010,
  - J. A. Mingo and R. Speicher. *Free Probability and Random Matrices*. Springer, New York, 2017.
- Concerning the construction of generalized beta ensembles, see e.g.
  - I. Dumitriu and A. Edelman. Matrix models for beta ensembles. *Journal of Mathematical Physics*, 43(11):5830–5847, 2002,
  - R. Allez, J. P. Bouchaud, and A. Guionnet. Invariant beta ensembles and the Gauss-Wigner crossover. *Physical Review Letters*, 109(9):094102, 2012.
- On the Ginibre ensemble and the circular law, see
  - V. L. Girko. Circular law. *Theory of Probability and Its Applications*, 29(4):694–706, 1985,
  - H. J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein. Spectrum of large random asymmetric matrices. *Physical Review Letters*, 60(19):1895–1899, 1988,
  - C. Bordenave and D. Chafaï. Around the circular law. *Probability Surveys*, 9, 2012.

- On the connection between non-crossing diagrams and RNA folding, see
  - P.-G. de Gennes. Statistics of branching and hairpin helices for the dAT copolymer. *Biopolymers*, 6(5):715–729, 1968,
  - H. Orland and A. Zee. RNA folding and large N matrix theory. *Nuclear Physics B*, 620(3):456–476, 2002.

https://doi.org/10.1017/9781108768900.005 Published online by Cambridge University Press