

# A law of large numbers for finite-range dependent random matrices

GREG ANDERSON  
*University of Minnesota*

AND

OFER ZEITOUNI  
*University of Minnesota*

## Abstract

We consider random hermitian matrices in which distant above-diagonal entries are independent but nearby entries may be correlated. We find the limit of the empirical distribution of eigenvalues by combinatorial methods. We also prove that the limit has algebraic Stieltjes transform by an argument based on dimension theory of noetherian local rings.

© 2000 Wiley Periodicals, Inc.

## 1 Introduction

Study of the empirical distribution of eigenvalues of random hermitian (or real symmetric) matrices has a long history, starting with the seminal work of Wigner [Wig55] and Wishart [Wis28]. Except in cases where the joint distribution of eigenvalues is explicitly known, most results available are asymptotic in nature and based on one of the following approaches: (i) the *moment method*, i. e., evaluation of expectations of traces of powers of the matrix; (ii) appropriate recursions for the resolvent, as introduced in [PM67]; or (iii) the free probability method (especially the notion of asymptotic freeness) originating with Voiculescu [Voi91]. A good review of the first two approaches can be found in [Ba99]. For the third, see [Voi00], and for a somewhat more combinatorial perspective, [Sp98]. These approaches have been extended to situations in which the matrix analyzed neither possesses i.i.d. entries above the diagonal (as in the Wigner case) nor is it the product of matrices with i.i.d. entries (as in the Wishart case). We mention in particular the papers [MPK92], [KKP96], [Sh96] and [Gu02] for results pertaining to the model of “random band matrices”, all with independent above-diagonal entries.

In our recent work [AZ06] we studied convergence of the empirical distribution of eigenvalues of random band matrices, and developed a combinatorial approach, based on the moment method, to identify the limit (and also to provide central limit theorems for linear statistics). Here we develop the method further to handle a class

of matrices with local dependence among entries (we postpone the precise definition of the class to Section 2). To each random matrix of the class we associate a random band matrix with the same limit of empirical distribution of eigenvalues by a process of “Fourier transformation”, thus making it possible to describe the limit in terms of our previous work (see Theorem 2.3). We also prove that the Stieltjes transform of the limit is algebraic (see Theorem 2.4), and deduce in turn from well-known facts about algebraic functions a strong regularity result for  $\mu$  (see Theorem 2.5). Algebraicity we prove by a general “soft” (i. e., nonconstructive) method based on the theory of noetherian local rings (see Theorem 6.1) which ought to be applicable to many more random matrix problems.

To get the flavor of our results, the reader should imagine a Wigner matrix (i.e., an  $N$ -by- $N$  real symmetric random matrix with i.i.d. above-diagonal entries, each of mean 0 and variance  $1/N$ ), on which a local “filtering” operation is performed: each entry not near the diagonal or an edge is replaced by half the sum of its four neighbors to northeast, southeast, southwest and northwest. At the end of Section 3 (Theorem 2.3 taken for granted) we analyze the “(NE+SE+SW+NW)-filtered Wigner matrix” described above. We find that the limit measure is the free multiplicative convolution of the semicircle law (density  $\propto \mathbf{1}_{|x|<2}\sqrt{4-x^2}$ ) and the arcsine law (density  $\propto \mathbf{1}_{0<x<2}/\sqrt{x(2-x)}$ ). The appearance in this example of a free multiplicative convolution has a simple explanation (see Proposition 3.4). We also write down the quartic equation satisfied by the Stieltjes transform of the limit measure.

Recently other authors have considered the empirical distribution of eigenvalues for matrices with dependent entries, see [GoT05],[Ch06],[SSB07]. Their class of models does not overlap significantly with ours. In particular, in all these works and unlike in our model, the limit of the empirical measure is always the same as that of a semicircle law multiplied by a random or deterministic constant.

Closest to our work is the recent paper by [HLN05], which builds upon earlier work by [BDM96] and [Gi90]. They consider Gram matrices of the form  $X_N X_N^*$  where  $X_N$  is a sequence of (non-symmetric) Gaussian matrices obtained by applying a filtering operation to a matrix with (complex) Gaussian independent entries. They also consider the case  $(X_N + A_N)(X_N + A_N)^*$  with  $A_N$  deterministic and Toeplitz. The Gaussian assumption allows them to directly approximate the matrix  $X_N$  by a unitary transformation of a Gaussian matrix with independent (but not identically distributed) entries, to which the results of [Gi90] and [AZ06] apply. An advantage of their approach is that they do not need to assume finiteness of the filter; they need only a summability condition. We note also that the approach in [HLN05] and [BDM96] is based on the study of resolvents rather than moments.

We mention now motivation from electrical engineering. The analysis of the limiting empirical distribution of eigenvalues of random matrices has recently played an important role in the analysis of communication systems, see [TV04] for an extensive review. In particular, when studying multi-antenna systems, one often

makes the (unrealistic) assumption that gains between different pairs of antennas are uncorrelated. The models studied in this work would allow correlation between neighboring antenna pairs. We do not develop this application further here.

The structure of the paper is as follows. In the next section we describe the class of matrices we treat, and state our main results, namely Theorem 2.3 (asserting a law of large numbers) and Theorem 2.4 (asserting algebraicity of a Stieltjes transform). We also prove Theorem 2.5 (which is essentially folkloric and explains the regularity implied by algebraicity). In Section 3 we discuss the limit measure in detail, and in particular write down algebro-integral equations for its Stieltjes transform, which we call *color equations*. Section 4 provides a computation of limits of traces of powers of the matrices under consideration. In Section 5 we complete the proof of Theorem 2.3 by a variance computation. In Section 6 we set up the algebraic machinery needed to prove Theorem 2.4. We finish proving the theorem in Section 7 by analyzing the color equations.

## 2 Formulation of the main result

After defining a class of kernels in §2.1, we define in §2.2 the class of matrices dealt with in this paper and in §2.3 describe the subclass of *filtered Wigner matrices*. To each kernel we associate a measure in §2.4. Finally, we state our main results, Theorems 2.3 and 2.4.

### 2.1 Kernels

#### Color space (motivation)

As in [AZ06], the spatial change in the variance structure of the entries of a random matrix is captured by an auxiliary variable, which we call “color”. The difference between the derivation in [AZ06] and the present paper is that we append to color space an additional variable (related to the local averaging mentioned in the introduction), that one should think of as a Fourier variable. The precise definition follows.

#### Color space (formal definition)

Let  $C = [0, 1] \times S^1$ , where  $S^1$  is the unit circle in the complex plane. We call  $C$  *color space*. We declare  $C$  to be a probability space under the product of uniform probability measures on  $[0, 1]$  and  $S^1$ , denoted  $P$ . We say that a  $C$ -valued random variable is *uniformly distributed* if its law is  $P$ . In the sequel, all  $L^p(C)$  (resp.,  $L^p(C \times C)$ ) spaces,  $p \geq 1$ , are taken with respect to the measure  $P$  (resp.,  $P \times P$ ).

#### The kernel $s$

We fix a kernel

$$s : C \times C \rightarrow \mathbb{R}$$

which will govern the local covariance structure of our random matrix model. We impose on  $s$  the following conditions.

**Assumption 2.1.**(I)  $s$  is a nonnegative symmetric function, i.e.

$$s(c, c') = s(c', c) \geq 0.$$

(II)  $s$  has a Fourier expansion

$$s(c, c') = \sum_{i, j \in \mathbb{Z}} s_{ij}(x, y) \xi^i \eta^j \quad (c = (x, \xi), c' = (y, \eta))$$

where all but finitely many of the coefficients

$$s_{ij} : [0, 1] \times [0, 1] \rightarrow \mathbb{C}$$

vanish identically.

(III) There is a finite partition  $\mathcal{I}$  of  $[0, 1]$  into subintervals of positive length such that every coefficient function  $s_{ij}$  is constant on every set of the form  $I \times J$  with  $I, J \in \mathcal{I}$ .(IV)  $s$  is nondegenerate:  $\|s\|_{L^1(C \times C)} > 0$ .**The sets  $Q_K^{(N)}$** For each  $N \in \mathbb{N}$  (here and below  $\mathbb{N}$  denotes the set of positive integers) and  $K > 0$ , we define

$$Q_K^{(N)} = \left\{ i \in \{1, \dots, N\} \mid \min_{x \in \partial \mathcal{I}} |i - Nx| > K \right\}$$

where  $\partial \mathcal{I} \subset [0, 1]$  is the finite set consisting of all endpoints of all intervals belonging to the family  $\mathcal{I}$ .**Remarks**

We have

$$(2.1) \quad \bar{s}_{ij} = s_{-i, -j}, \quad s_{ji}(y, x) = s_{ij}(x, y)$$

because  $s$  is real-valued and symmetric. Assumptions (I, II, III) imply that

$$(2.2) \quad 0 \leq s \leq \|s\|_{L^\infty(C \times C)} < \infty$$

holds everywhere (not just  $P \times P$ -a.e.).**2.2 The model**For each  $N \in \mathbb{N}$ , let

$$X^{(N)} = [X_{ij}^{(N)}]_{i, j=1}^N$$

be an  $N$ -by- $N$  random hermitian matrix. We impose the following conditions, where  $s$  satisfies Assumption 2.1.**Assumption 2.2.**

- (I) (a)  $\forall N \in \mathbb{N} \quad \forall i, j \in \{1, \dots, N\} \quad EX_{ij}^{(N)} = 0.$   
 (b)  $\forall k \in \mathbb{N} \quad \sup_{N=1}^{\infty} \max_{i,j=1}^N E|X_{ij}^{(N)}|^k < \infty.$
- (II) *There exists  $K > 0$  such that for all  $N \in \mathbb{N}$ , the following hold:*
- (a)  $\forall i, j \in \mathbb{Z} \quad \max(|i|, |j|) > K \Rightarrow s_{ij} \equiv 0.$
- (b) *For all nonempty subsets*  
 $A, B \subset \{(i, j) \in \{1, \dots, N\}^2 \mid 1 \leq i \leq j \leq N\}$   
*such that*  

$$\min_{(i,j) \in A} \min_{(k,\ell) \in B} \max(|i-k|, |j-\ell|) > K,$$
*the  $\sigma$ -fields*  
 $\sigma(\{X_{ij}^{(N)} \mid (i, j) \in A\}), \quad \sigma(\{X_{k\ell}^{(N)} \mid (k, \ell) \in B\})$   
*are independent.*
- (c)  $\forall i, j, k, \ell \in Q_K^{(N)}$  s.t.  $\min(j-i, \ell-k) > K$   
 $EX_{ij}^{(N)} \overline{X}_{k\ell}^{(N)} = s_{i-k, \ell-j}(i/N, j/N).$

### The empirical distribution of eigenvalues

Let

$$\lambda_1^{(N)} \leq \lambda_2^{(N)} \leq \dots \leq \lambda_N^{(N)}$$

denote the eigenvalues of the hermitian matrix  $X^{(N)}/\sqrt{N}$ , and let

$$L^{(N)} = N^{-1} \sum_{i=1}^N \delta_{\lambda_i^{(N)}}$$

denote the corresponding empirical distribution of the eigenvalues. We are concerned with the limiting behavior of  $L^{(N)}$  as  $N \rightarrow \infty$ .

### Remarks

(i) The existence of  $K$  satisfying Assumption 2.2(IIa) is assured by Assumption 2.1(IIb). (ii) Assumption 2.2(IIb) says that the on-or-above diagonal entries of  $X^{(N)}$  form a finite-range dependent random field, with  $K$  a bound for the range of dependence. This explains the title of the paper. (iii) Assumption 2.2(IIc) fixes variances at each site and also short-range local correlations for sites in sufficiently general position. (iv) None of our assumptions rule out the possibility that all matrices  $X^{(N)}$  are real. In other words, we can handle hermitian and real symmetric cases uniformly under Assumption 2.2. (v) The relations (2.1), which are equivalent to the reality and symmetry of the kernel  $s$ , play a key role in our analysis of

the limiting behavior of  $L^{(N)}$ . New methods would be needed were the reality and symmetry conditions to be relaxed.

### Kernels of pure spatial type

Let  $s : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$  be a kernel. If the kernel  $\tilde{s} : C \times C \rightarrow \mathbb{R}$  defined by the formula  $\tilde{s}((x, \xi), (y, \eta)) = s(x, y)$  satisfies Assumption 2.1, then by abuse of terminology we say that  $s$  is a kernel of *pure spatial type* satisfying Assumption 2.1, and with the evident modification of Assumption 2.2(IIc) we can use  $s$  to govern the covariance structure of our model. In the special case of kernels of pure spatial type our model essentially contains the model of [AZ06] in the special case in which color space is a finite set.

### Kernels of pure Fourier type

Let  $s : S^1 \times S^1 \rightarrow \mathbb{R}$  be a kernel. If the kernel  $\tilde{s} : C \times C \rightarrow \mathbb{R}$  defined by the formula  $\tilde{s}((x, \xi), (y, \eta)) = s(\xi, \eta)$  satisfies Assumption 2.1 with  $\mathcal{S} = \{[0, 1]\}$ , then by abuse of terminology we say that  $s$  is a kernel of *pure Fourier type* satisfying Assumption 2.1, and with the evident modification of Assumption 2.2(IIc) we can use  $s$  to govern the covariance structure of our model. We suggest that the reader focus on the pure Fourier case when first approaching this paper because little would be lost in terms of grasping the main ideas. The main reason for us to work at a higher level of generality is to make sure that our theory handles not only “filtered Wigner matrices” (for which  $|\mathcal{S}| = 1$ ) but also “filtered Wishart matrices” (for which  $|\mathcal{S}| = 2$ ). (Here and below  $|S|$  denotes the cardinality of a set  $S$ .) We then might as well allow  $|\mathcal{S}| > 2$  as a possibility because there is no gain in simplicity by excluding it.

## 2.3 Filtered Wigner matrices

We describe now a natural class of random matrices fitting into the framework of Assumptions 2.1 and 2.2. This class should be considered the main motivation for the paper. Members of the class arise by “filtering” Wigner matrices. The corresponding kernels are of pure Fourier type and depend in a simple way on the “filter”.

### Wigner matrices

Let

$$\{Y_{ij}\}_{-\infty < i < j < \infty}$$

be an i.i.d. family of real random variables. Assume that  $Y_{01}$  has absolute moments of all orders. Assume that  $EY_{01} = 0$  and  $EY_{01}^2 = 1$ . Put

$$Y_{ii} = 0 \text{ for } -\infty < i < \infty, \quad Y_{ij} = Y_{ji} \text{ for } -\infty < j < i < \infty.$$

Then

$$(2.3) \quad Y_{ij} = Y_{ji}, \quad EY_{ij} = 0, \quad EY_{ij}Y_{kl} = (\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk})(1 - \delta_{ij})(1 - \delta_{kl})$$

for all  $i, j, k, \ell \in \mathbb{Z}$ . Let  $Y^{(N)} = [Y_{ij}^{(N)}]_{i,j=1}^N$  be the  $N$ -by- $N$  matrix with entries  $Y_{ij}$ . Then  $Y^{(N)}/\sqrt{N}$  (in the terminology of [AZ06]) is a *Wigner matrix*, and hence the empirical distribution of its eigenvalues for  $N \rightarrow \infty$  tends to the semicircle law. In particular, if the  $Y_{ij}$  are standard normal random variables and one makes a suitable adjustment to the diagonal of  $Y^{(N)}/\sqrt{N}$ , the result is a Wigner matrix in the standard sense, i.e., a member of the Gaussian orthogonal ensemble.

### The filter, its Fourier transform, and associated kernel

Let a *filter*

$$h : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{R}$$

be given, with  $H$  denoting its Fourier transform, that is

$$H(\xi, \eta) = \sum_{i,j \in \mathbb{Z}} h(i, j) \xi^i \eta^j \quad (\xi, \eta \in S^1).$$

We assume that  $h$  does not vanish identically. We assume that there exists  $K > 0$  such that

$$(2.4) \quad \max(|i|, |j|) > K/2 \Rightarrow h(i, j) = 0,$$

and hence  $H$  is well-defined. We assume that  $h$  satisfies the symmetry condition

$$(2.5) \quad h(-i, -j) = h(j, i),$$

which implies the symmetry condition

$$\overline{H(\xi, \eta)} = H(\eta, \xi).$$

Put

$$s(\xi, \eta) = |H(\xi, \eta)|^2 = \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} s_{ij} \xi^i \eta^j \quad (s_{ij} \in \mathbb{R}).$$

Then  $s : S^1 \times S^1 \rightarrow \mathbb{R}$  is a kernel of pure Fourier type satisfying Assumption 2.1. In particular,  $\|s\|_{L^1(C \times C)} = \|h\|_{L^2(\mathbb{Z} \times \mathbb{Z})}^2 > 0$ .

### Filtered Wigner matrices (definition)

For  $i, j \in \{1, \dots, N\}$  set

$$X_{ij}^{(N)} = \sum_{k=1}^N \sum_{\ell=1}^N Y_{k\ell} h(i-k, \ell-j),$$

thus defining an  $N$ -by- $N$  random matrix  $X^{(N)}$  which in view of the symmetry (2.5) is real symmetric. We call  $X^{(N)}/\sqrt{N}$  a *filtered Wigner matrix*, with *filter*  $h$ . We think of  $X^{(N)}/\sqrt{N}$  as the result of filtering the Wigner matrix  $Y^{(N)}/\sqrt{N}$  by  $h$ .

### Local covariance structure of $X^{(N)}$

From (2.3) and (2.4) we deduce that

$$(2.6) \quad i, j, k, \ell \in \mathcal{Q}_K^{(N)} \text{ \& } \min(j-i, \ell-k) > K \Rightarrow EX_{ij}^{(N)} X_{k\ell}^{(N)} = s_{i-k, \ell-j},$$

after a straightforward (extremely tedious) calculation. This is the main point in verifying that the real symmetric random matrices  $X^{(N)}$  satisfy Assumption 2.2 with respect to the kernel  $s$ . The remaining details needed to check Assumption 2.2 are easy to supply. Thus we can put filtered Wigner matrices into the framework of our model.

### The (NE+SE+SW+NW)-filtered Wigner matrix

Taking

$$h = (1/2)\mathbf{1}_{\{(1,-1), (1,1), (-1,1), (-1,-1)\}}, \text{ and hence } s(e^{i\theta_1}, e^{i\theta_2}) = 4 \cos^2 \theta_1 \cos^2 \theta_2,$$

we get a precisely defined version of the model mentioned in the introduction, which presently we will analyze in detail. (We call this the (NE+SE+SW+NW)-filtered Wigner matrix because each above-the-diagonal entry in the matrix not too near the diagonal or the edges can be viewed as the average of the four entries in a standard Wigner matrix which are to its immediate northeast, southeast, southwest and northwest.) Note that with  $\theta$  uniformly distributed in  $[0, 2\pi)$ , the law of  $2 \cos^2 \theta$  has density  $\propto \mathbf{1}_{0 < x < 2} / \sqrt{x(2-x)}$ .

### Remark

In [AZ06] we handled (real) Wishart matrices  $Z^T Z$  (and more general matrices formed from matrices  $Z$  with independent real but perhaps not i.i.d. entries) in terms of band-type matrices  $\begin{bmatrix} 0 & Z \\ Z^T & 0 \end{bmatrix}$ . A similar trick in the present setting puts “filtered Wishart matrices” into the framework of our model. For the kernels arising in that connection, the associated partition  $\mathcal{S}$  consists of two intervals. We do not discuss the Wishart case further here.

## 2.4 The measure $\mu_s$

We make the last preparation to state our main results. Let  $s$  be a kernel satisfying Assumption 2.1. For each positive integer  $N$ , let

$$C^{(N)} = \{c_1^{(N)}, \dots, c_{N^2}^{(N)}\} \subset C$$

be the set of pairs  $(x, \xi) \in C$  where  $x \in [0, 1) \cap \frac{1}{N}\mathbb{Z}$  and  $\xi^N = 1$ . Then the empirical distribution  $\frac{1}{N^2} \sum_{c \in C^{(N)}} \delta_c$  tends weakly as  $N \rightarrow \infty$  to the uniform probability measure  $P$ . Let  $C^{(\infty)}$  be the union of the sets  $C^{(N)}$ . Let

$$\{\tilde{Y}_e\}_{e \in C^{(\infty)}} \text{ s.t. } e \text{ has cardinality 1 or 2}$$

be an i.i.d. family of standard normal (mean 0 and variance 1) random variables. Let  $\tilde{X}^{(N)}$  be the  $N^2$ -by- $N^2$  real symmetric random matrix with entries

$$\tilde{X}_{ij}^{(N)} = 2^{\delta_{ij}/2} \sqrt{s(c_i^{(N)}, c_j^{(N)})} \tilde{Y}_{\{c_i^{(N)}, c_j^{(N)}\}}.$$

Let  $\tilde{\lambda}_1^{(N)} \leq \dots \leq \tilde{\lambda}_{N^2}^{(N)}$  be the eigenvalues of  $\tilde{X}^{(N)}/N$  and let  $\tilde{L}^{(N)} = \frac{1}{N^2} \sum_{i=1}^{N^2} \delta_{\tilde{\lambda}_i^{(N)}}$  be the empirical distribution of the eigenvalues. By [AZ06, Thm. 3.2] the empirical distribution  $\tilde{L}^{(N)}$  tends weakly in probability as  $N \rightarrow \infty$  to a limit  $\mu_s$  with bounded support. (The strange-looking factor  $2^{\delta_{ij}/2}$  in the definition of  $\tilde{X}^{(N)}$  could be dropped without changing the limit of  $L^{(N)}$ . More generally, within the theory of [AZ06], one has many ways to construct models with limiting measure  $\mu_s$ . We made our concrete choice to simplify the proof of Proposition 3.4 below.) In Section 3 we will provide a combinatorial description of the moments of  $\mu_s$  and write down algebro-integral equations (which we call *color equations*) satisfied by the Stieltjes transform of  $\mu_s$  and certain auxiliary functions.

The following are our main results. In these results we fix a kernel  $s$  satisfying Assumption 2.1 and a family  $[X^{(N)}]_{N=1}^{\infty}$  of random hermitian matrices satisfying Assumption 2.2 with respect to  $s$ . As defined in §2.2, let  $L^{(N)}$  be the empirical distribution of the eigenvalues of  $X^{(N)}/\sqrt{N}$ . Let  $\mu = \mu_s$  be the measure associated to  $s$  by the procedure of §2.4.

**Theorem 2.3.**  $L^{(N)}$  converges weakly in probability to  $\mu$ .

**Theorem 2.4.** The Stieltjes transform  $S(\lambda) = \int \frac{\mu(dx)}{\lambda - x}$  is an algebraic function of  $\lambda$ , i. e., there exists some not-identically-vanishing polynomial  $F(X, Y)$  in two variables with complex coefficients such that  $F(\lambda, S(\lambda))$  vanishes for all complex numbers  $\lambda$  not in the support of  $\mu$ .

To prove Theorem 2.3 we first prove ‘‘convergence in moments’’ in Section 4, and then finish the proof in Section 5 by considering variances. To prove Theorem 2.4 we first set up a general method for proving algebraicity in Section 6, and then finish the proof in Section 7 by analyzing the color equations.

## 2.5 Remarks on algebraicity

(i) In the setting of Theorem 2.4, the polynomial  $F(X, Y)$  is not unique, but if we impose the further condition of irreducibility, then  $F(X, Y)$  is unique up to a nonzero constant factor. (See [Ahl79, Subsec. 2.2, Chap. 8].) When  $F(X, Y)$  is so specified we call it the *irreducible polynomial* for  $S(\lambda)$ .

(ii) The method of proof of Theorem 2.4 does not yield an explicit irreducible polynomial  $F(X, Y)$  for  $S(\lambda)$ . But the elimination of variables necessary to make  $F(X, Y)$  explicit could in principle be carried out on a computer using Gröbner basis techniques. For an introduction to the latter see the very accessible book [CLO07]. Our result serves as a theoretical guarantee that  $F(X, Y)$  can indeed be

found. The extent to which such Gröbner basis calculations could be streamlined enough to be useful in applications remains to be investigated.

(iii) In recent work Rao and Edelman [RE06] explain how knowledge of an algebraic equation for the Stieltjes transform can be exploited for numerical computation of limiting spectra in various random matrix models. Furthermore, in the same setting, they explain how to streamline the computation of free additive and multiplicative convolution.

(iv) The general methods of Section 6, and specifically Theorem 6.1, apply to the Stieltjes transforms of the limiting measures arising from the model of [AZ06] in the case of a finite color space, yielding algebraicity in all those cases. They also apply to the equations for the limit of the Stieltjes transform in [HLN05] and the systems of equations in [Gi90, Corollaries 9.3.2 and 10.1.2], under appropriate hypotheses of “finiteness of filter”, as well as many of the pre-limit “Kx canonical equations” in [Gi01],  $x \in \{1, \dots, 30\}$ , and in particular to his K1 canonical equations.

(v) Questions of algebraicity have attracted attention for some time in areas bordering on random matrix theory. For example, one may ask if the Green function for random walk on a free group is algebraic. A general result asserting this algebraicity under quite mild hypotheses was proved in [Ao84]. These algebraicity questions were recently revisited in [GB07], and a wide variety of striking connections were discussed, e.g., a connection with the theory of context-free languages. An analogous problem in free probability is to determine the spectra of convolution operators on a free group (and thus through the theory of asymptotic freeness to determine the limit measures for random matrix models involving several independent Haar-distributed unitary matrices in the large  $N$  limit). In the paper [Voi92], which was primarily devoted to laying out the operator-valued version of the notion of freeness, an analysis of such spectra was presented as a sample application.

(vi) Algebraicity implies a rather strong regularity property for  $\mu$  which is of theoretical interest independent of computational issues. We state and prove a theorem immediately below to explain this point in detail. We note that other techniques have been employed in the study of regularity of the limiting spectra of random matrices, see for example [SiC95].

## 2.6 Setup for the regularity theorem

We declare a real-valued function  $h$  defined in a bounded open interval  $(a, b)$  to be of *rational beta type* under the following conditions:

- $h$  is real-analytic and nonnegative on  $(a, b)$ .
- There exist positive rational numbers  $c$  and  $d$  such that both limits

$$\lim_{x \downarrow a} (x - a)^{1-c} h(x), \quad \lim_{x \uparrow b} (b - x)^{1-d} h(x)$$

exist and are positive.

Note that  $h \in L^1(a, b)$  and that  $h$  can have only finitely many zeroes in  $(a, b)$ . For example  $x^{c-1}(1-x)^{d-1}$  is of rational beta type in  $(0, 1)$  for all positive rational numbers  $c$  and  $d$  (which explains the terminology).

Now let  $\mu$  be a probability measure on the real line with compact support  $K$  and algebraic Stieltjes transform  $S(z) = \int \frac{\mu(dx)}{z-x}$ . (We take  $z$  as our complex variable instead of  $\lambda$  to emphasize that we are now concerned only with complex analysis, not RMT.) Let  $F(X, Y)$  be a not-identically-vanishing polynomial such that  $F(z, S(z)) = 0$  for all  $z \in \mathbb{C} \setminus K$  and furthermore the discriminant  $D(X)$  of  $F(X, Y)$  with respect to  $Y$  is not-identically-vanishing. (We may, for example, take  $F(X, Y)$  to be the irreducible polynomial of  $S(z)$ , because in that case the discriminant  $D(X)$  cannot vanish identically. But there is no need to insist on irreducibility of  $F(X, Y)$ ; only the condition  $D(X) \neq 0$  turns out to be important.) Let  $n$  be the degree of  $F(X, Y)$  in  $Y$  and write  $F(X, Y) = \sum_{i=0}^n F_i(X)Y^i$ . Let  $A$  be the (finite) set of complex zeroes of  $F_n(X)D(X)$ .

**Theorem 2.5.** *Notation and assumptions are as above. Let  $I$  be a connected component of  $\mathbb{R} \setminus A$  and let  $\mu|_I$  be the restriction of  $\mu$  to (the Borel subsets of)  $I$ . If  $I \setminus K$  is nonempty, then  $\mu|_I$  vanishes identically. If  $I$  is bounded and  $\mu|_I$  does not vanish identically, then  $\mu|_I$  has density of rational beta type with respect to Lebesgue measure.*

We emphasize that this result is essentially folkloric. We supply a proof only for lack of a suitable reference. We remark that the result does not rule out the possibility of  $\mu$ -atoms at points of  $A \cap \mathbb{R}$ . First we need a lemma. For each positive integer  $N$  let  $r_N$  be a choice of analytic  $N^{\text{th}}$  root of  $z$  defined in the open set  $\mathbb{C} \setminus \{-it \mid t \in [0, \infty)\}$ . Throughout, we use  $\Im z$  for the imaginary part of  $z \in \mathbb{C}$ .

**Lemma 2.6.** *Choose  $\delta > 0$  so small that  $A \cap \mathbb{R} = A \cap \{|\Im z| < \delta\}$ . Let*

$$W = \{|\Im z| < \delta\} \setminus \bigcup_{a \in A \cap \mathbb{R}} \{a - it \mid t \in [0, \infty)\}.$$

*Then: (i) There exists a unique analytic function  $g$  defined in  $W$  which agrees with  $S$  on the set  $\{0 < \Im z < \delta\}$ . (ii) For each  $a \in A \cap \mathbb{R}$  there exists a positive integer  $N$  and  $\varepsilon > 0$  such that  $(z-a)^N g(z)$  admits expansion in the set  $W \cap \{|z-a| < \varepsilon\}$  as a convergent power series in  $r_N(z-a)$ .*

*Proof.* Let  $\hat{F}(X, Y)$  be the irreducible polynomial of  $S(z)$ , which divides  $F(X, Y)$ . Let  $\hat{n}$  be the degree of  $\hat{F}(X, Y)$  in  $Y$  and write  $\hat{F}(X, Y) = \sum_{i=0}^{\hat{n}} \hat{F}_i(X)Y^i$ . Let  $\hat{D}(X)$  be the discriminant of  $\hat{F}(X, Y)$  with respect to  $Y$ . Let  $\hat{A}$  be the set of complex zeroes of  $\hat{F}_{\hat{n}}(X)\hat{D}(X)$ . Then we have  $\hat{A} \subset A$ . Let  $\mathbf{f}$  be the collection of all pairs  $(f, U)$  consisting of an open subset  $U \subset \mathbb{C}$  and an analytic function  $f$  defined in  $U$  such that  $\hat{F}(z, f(z)) = 0$  for all  $z \in U$ . Then  $\mathbf{f}$  is the global analytic and algebraic function naturally associated to  $\hat{F}(X, Y)$ . (Here up to minor changes of notation we follow the definitions of [Ahl79, Secs. 1,2, Chap. 8].) The main point of this construction for us is that the pair  $(S, \mathbb{C} \setminus K)$  belongs to  $\mathbf{f}$ . Every branch of  $\mathbf{f}$  at a point in  $\mathbb{C} \setminus \hat{A}$

admits analytic continuation along any arc not passing through the set  $\hat{A}$ . (See [Ahl79, Subsec. 2.2, Pg. 303, Chap. 8].) Since  $U$  and  $\{0 < \Im z < \delta\}$  are simply connected and disjoint from  $\hat{A}$ , statement (i) follows by the Monodromy Theorem [Ahl79, Theorem 2, Subsec. 1.6, Chap. 8]. Statement (ii) follows from the study of singularities of algebraic functions in [Ahl79, Subsec. 2.3, Chap. 8].  $\square$

*Proof of Theorem 2.5.* Recall the inversion formula

$$\int \phi(x) \mu(dx) = \frac{1}{\pi} \lim_{y \downarrow 0} \int \Im S(x+iy) \phi(x) dx$$

holding for all bounded continuous functions  $\phi$  defined on the real line. With  $g$  as in part (i) of the previous lemma, let  $h = \frac{1}{\pi} \Im g|_{\mathbb{R} \setminus A}$ . From the inversion formula it follows that  $\mu|_I$  has density  $h|_I$  with respect to Lebesgue measure. If  $I \setminus K$  is nonempty (this is in particular the case if  $I$  is unbounded), then  $h|_I$  vanishes identically in some (nonempty) open subinterval of  $I$ , and hence (since real analytic) vanishes identically. Otherwise  $I$  is bounded, and if  $h|_I$  does not vanish identically, then  $h|_I$  is of rational beta type by part (ii) of the preceding lemma along with the fact that  $h|_I \in L^1(I)$ .  $\square$

### 3 The moments and Stieltjes transform of $\mu$

We fix a kernel  $s$  satisfying Assumption 2.1 and put  $\mu = \mu_s$ . We provide a detailed description of the moments and Stieltjes transform of  $\mu$ . We also introduce combinatorial tools needed throughout the paper.

#### 3.1 Quick review of key combinatorial notions

##### Graphs

For us a *graph*  $G = (V, E)$  consists by definition of a finite set  $V$  of *vertices* and a set  $E$  of *edges*, where every element of  $E$  is a subset of  $V$  of cardinality 1 or 2. In other words, we are dealing here with graphs (i) which have finitely many vertices, (ii) which have unoriented edges, (iii) in which a vertex may be joined to itself by an edge, but (iv) in which no two vertices may be joined by more than one edge. A graph  $G$  is a *tree* if  $G$  is connected and  $|V| = |E| + 1$ . We emphasize that every edge of a tree joins two distinct vertices—it is not allowed for a vertex of a tree to be joined to itself by an edge.

##### Set partitions

We say that a set  $\pi \subset 2^{\{1, \dots, k\}}$  is a *set-partition* of  $k$  if  $\pi$  is a disjoint family of nonempty sets with union equal to  $\{1, \dots, k\}$ . The elements of  $\pi$  are called the *parts* of  $\pi$ . For each  $i \in \{1, \dots, k\}$ , let  $\pi(i)$  be the part of  $\pi$  to which  $i$  belongs. For convenience we extend  $i \mapsto \pi(i)$  to a periodic function on  $\mathbb{Z}$  by the rule  $\pi(i) = \pi(i+k)$ .

### The graph associated to a set partition

To each set partition  $\pi$  of  $k$  there is canonically associated a graph  $G_\pi = (V_\pi, E_\pi)$ , where  $V_\pi = \pi$  and

$$E_\pi = \{\{\pi(i), \pi(i+1)\} \mid i = 1, \dots, k\}.$$

By construction  $G_\pi$  comes canonically equipped with a walk, namely

$$\pi(1), \dots, \pi(k), \pi(k+1) = \pi(1),$$

whence in particular it follows that  $G_\pi$  is connected.

### Wigner set partitions

We say that a set partition  $\pi$  of  $k$  is a *Wigner set partition* if the corresponding graph  $G_\pi$  has  $k/2 + 1$  vertices and  $k/2$  edges, in which case  $G_\pi$ , since connected, is a tree. We denote the set of such  $\pi$  by  $\mathscr{W}_k$ . For  $k$  odd the set  $\mathscr{W}_k$  is empty. For  $k = 2\ell$  the set  $\mathscr{W}_k$  is canonically in bijection with the set of rooted planar trees with  $\ell + 1$  nodes and hence, as is well-known [St99], the cardinality of  $\mathscr{W}_k$  is the *Catalan number*  $\frac{1}{\ell+1} \binom{2\ell}{\ell}$ .

**Lemma 3.1.** Fix  $\pi \in \mathscr{W}_k$ . (i) For each  $i \in \{1, \dots, k\}$  we have  $\pi(i) \neq \pi(i+1)$ . (ii) For each  $e \in E_\pi$  the equation  $e = \{\pi(i), \pi(i+1)\}$  has exactly two solutions  $i \in \{1, \dots, k\}$ , say  $i_1$  and  $i_2$ , and moreover  $\pi(i_1) \neq \pi(i_2)$ . (iii) For each  $i \in \{1, \dots, k\}$  we have  $\{\pi(i), \pi(i+1)\} = \{\pi(j-1), \pi(j)\}$ , where  $j$  is the least of the integers  $\ell > i$  such that  $\pi(i) = \pi(\ell)$ .

*Proof.* The lemma formalizes facts about the tree  $G_\pi$  and the canonical walk on it which are clear from a graph-theoretic point of view. (i) No edge of  $G_\pi$  connects a vertex to itself. (ii) The canonical walk on  $G_\pi$  visits each edge of  $G_\pi$  exactly twice. More precisely, the canonical walk traverses each edge of  $G_\pi$  exactly once in each direction. (iii) The canonical walk on  $G_\pi$  extended by periodicity returns to a given vertex on the same edge by which it departed.  $\square$

### Tree integrals

Let  $\{\kappa_A\}$  be an i.i.d. family of  $C$ -valued uniformly distributed random variables indexed by finite nonempty sets of positive integers. Expectations with respect to these variables are denoted by  $\mathbb{E}$ . For each  $\pi \in \mathscr{W}_k$  we define a bounded random variable by the formula

$$M_\pi = \prod_{\{A,B\} \in E_\pi} s(\kappa_A, \kappa_B),$$

which is well-defined on account of the symmetry  $s(c, c') = s(c', c)$ . We call the expectation  $\mathbb{E}M_\pi$  a *tree integral*.

**Proposition 3.2** (Combinatorial description of the moments of  $\mu$ ). *We have*

$$(3.1) \quad \langle \mu, x^k \rangle = \sum_{\pi \in \mathscr{W}_k} \mathbb{E}M_\pi$$

for every integer  $k > 0$ .

The proposition is essentially just a special case of [AZ06, Thm. 3.2], but a fair amount of explanation is needed because the set up in this paper is (superficially) incompatible with that of [AZ06]—here we emphasize set partitions, whereas in [AZ06] we emphasized “words” and “spelling”. We can immediately reduce the proposition to the following technical lemma. The lemma is slightly more detailed than needed for the proof of the proposition—part (ii) will be needed for the derivation of algebro-integral equations for the Stieltjes transform of  $\mu$ .

**Lemma 3.3.** *Put  $A = 2\|s\|_{L^\infty(C \times C)}^{1/2}$ . (i) There exists a unique system  $\{\Phi_n, \Psi_n\}_{n \in \mathbb{N}}$  of functions in  $L^\infty(C)$  such that*

$$(3.2) \quad \Psi_n(c) = \int s(c, c') \Phi_n(c') P(dc')$$

*holds P-a.e. for every n and there holds an identity*

$$(3.3) \quad \sum_{n=1}^{\infty} \Phi_n t^n = t \left( 1 - t \sum_{n=1}^{\infty} \Psi_n t^n \right)^{-1}$$

*of formal power series in t with coefficients in  $L^\infty(C)$ . (ii) The bounds*

$$(3.4) \quad 0 \leq \Phi_n \leq A^{n-1}, \quad 0 \leq \Psi_n \leq A^{n+1}/4$$

*hold P-a.e. for every n. (iii) The formula*

$$(3.5) \quad \langle P, \Phi_{k+1} \rangle = \sum_{\pi \in \mathcal{M}_k} \mathbb{E} M_\pi$$

*holds for every integer  $k > 0$ .*

*Proof of Proposition 3.2 granting Lemma 3.3.* According to [AZ06, Lemma 3.2] (taking there  $\sigma = P$ ,  $D = 0$ , and  $s^{(2)} = s$ ), there exists a unique probability measure on the real line with  $k^{\text{th}}$  moment  $\langle P, \Phi_{k+1} \rangle$  for every  $k \geq 0$ . That measure according to [AZ06, Thm. 3.2] is none other than  $\mu$ .  $\square$

*Plan for the proof of Lemma 3.3.* Part (i) of the lemma is nothing but an inductive definition of  $\Phi_n$  and  $\Psi_n$ . So only parts (ii,iii) require proof. In principle, part (iii) follows from [AZ06, Lemmas 6.3 and 6.4], but because of the incompatibility of set-ups noted above, those lemmas cannot be directly applied here—some amplification is needed. Also part (ii) is most easily explained from the point of view of [AZ06, loc. cit.] So, after recalling in §3.2 the needed background from [AZ06], we lightly sketch in §3.3 a proof of parts (ii,iii) of the lemma.

### 3.2 The “verbal” approach

We briefly recall the point of view of [AZ06] and compare it to the present one. The material reviewed here will be used in a substantial way only in Section 3, not in later sections of the paper.

## Words

We fix a set of *letters* and define a *word* to be a finite nonempty sequence  $w = \alpha_1 \cdots \alpha_k$  of letters. We say that words  $w = \alpha_1 \cdots \alpha_k$  and  $x = \beta_1 \cdots \beta_\ell$  are *equivalent* if  $k = \ell$  (the words are the same length) and there exists a one-to-one-correspondence  $\varphi : \{\alpha_i\} \rightarrow \{\beta_j\}$  such that  $\varphi(\alpha_i) = \beta_i$  for  $i = 1, \dots, k$  (each word codes to the other under a simple substitution cipher). Each word  $w = \alpha_1 \cdots \alpha_k$  of length  $k$  gives rise naturally to a set partition of  $k$ , namely the set of equivalence classes for the equivalence relation  $i \sim j \Leftrightarrow \alpha_i = \alpha_j$ . Two words are equivalent if and only if they have the same length and give rise to the same set partition. The upshot is that speaking of equivalence classes of words is equivalent to speaking of set partitions.

## Wigner words

Let  $w$  be a word of at least two letters with same first and last letter, and let  $w'$  be the word obtained by dropping the last letter of  $w$ . The word  $w$  is a Wigner word in the sense of [AZ06] if and only if the set partition associated to  $w'$  is a Wigner set partition in the sense of this paper. In [AZ06] we also declared every one-letter word to be a Wigner word. The Wigner words have a simple inductive characterization [AZ06, Prop. 4.5 and §4.7] which is not so convenient to state in the set partition language. To wit, a word  $w$  is a Wigner word if and only if the following conditions hold:

- The first and last letters of  $w$  are the same.
- No letter appears twice in a row in  $w$ .
- Let  $\alpha$  be the first letter of  $w$ . Write  $w = \alpha w_1 \alpha \cdots \alpha w_r \alpha$ , where  $\alpha$  does not appear in any of the words  $w_i$ . Then each word  $w_i$  is a Wigner word, and moreover for  $i \neq j$  the words  $w_i$  and  $w_j$  have no letters in common. (If  $r = 0$  then  $w$  consists of a single letter and is by definition a Wigner word.)

## “Verbal” description of tree integrals

Let  $\{\kappa_\alpha\}$  be a letter-indexed i.i.d. family of uniformly distributed  $C$ -valued random variables. Given a Wigner word  $w$ , we define a random variable  $M_w$  inductively by the following procedure:

- Writing  $w = \alpha w_1 \alpha \cdots \alpha w_r \alpha$  as in the inductive characterization of Wigner words, and letting  $\beta_i$  be the initial letter of  $w_i$  for  $i = 1, \dots, r$ , we set  $M_w = \prod_{i=1}^r s(\kappa_\alpha, \kappa_{\beta_i}) M_{w_i}$ . (When  $w$  is one letter long, we put  $M_w = 1$ .)

The formula (3.5) claimed in Lemma 3.3 can be rewritten

$$(3.6) \quad \langle P, \Phi_{k+1} \rangle = \sum_{w \in \mathcal{W}_{k+1}} \mathbb{E} M_w$$

where the sum is extended over a set of representatives  $W_{k+1}$  for equivalence classes of Wigner words of length  $k+1$ . Further and crucially, notation as above in the inductive characterization of the random variables  $M_w$ , we have a relation

$$(3.7) \quad \mathbb{E}(M_w | \kappa_\alpha) = \prod_{i=1}^r \mathbb{E}(s(\kappa_\alpha, \kappa_{\beta_i}) \mathbb{E}(M_{w_i} | \kappa_{\beta_i}) | \kappa_\alpha)$$

among conditional expectations.

### 3.3 Proof of Lemma 3.3(ii,iii)

It is enough to construct an example of a system  $\{\Phi_n, \Psi_n\}$  in  $L^\infty(C)$  satisfying (3.2, 3.3, 3.4, 3.5), and to do so we follow the path of the proofs of [AZ06, Lemmas 6.3 and 6.4]. Fix a letter  $\alpha$ . We may suppose that every word belonging to the set of representatives  $W_{k+1}$  figuring in (3.6) begins with  $\alpha$ . There exist for each integer  $k \geq 0$  well-defined  $\Phi_{k+1}, \Psi_{k+1} \in L^\infty(C)$  such that

$$\Phi_{k+1}(\kappa_\alpha) = \sum_{w \in W_{k+1}} \mathbb{E}(M_w | \kappa_\alpha), \quad \Psi_{k+1}(\kappa_{\alpha'}) = \sum_{w \in W_{k+1}} \mathbb{E}(s(\kappa_{\alpha'}, \kappa_\alpha) M_w | \kappa_{\alpha'}),$$

where  $\alpha'$  is a letter not appearing in any of the words belonging to the set  $W_{k+1}$ . The system  $\{\Phi_n, \Psi_n\}$  has property (3.2) by construction, and has property (3.5) since the latter is equivalent to (3.6). Since

$$|\mathcal{W}_k| = |W_{k+1}| \leq 2^k, \quad 0 \leq M_w \leq \|s\|_{L^\infty(C \times C)}^{k/2} \text{ for } w \in W_{k+1},$$

for all integers  $k \geq 0$ , the system  $\{\Phi_n, \Psi_n\}$  has property (3.4). Finally, from (3.7) and the inductive characterization of Wigner words, we obtain identities

$$\Phi_{k+1} = \sum_{r=0}^{\infty} \sum_{\substack{(\ell_1, \dots, \ell_r) \in \mathbb{N}^r \\ \sum_{i=1}^r (\ell_i + 1) = k}} \prod_{i=1}^r \Psi_{\ell_i}$$

in  $L^\infty(C)$  for all integers  $k \geq 0$  which together imply that the system  $\{\Phi_n, \Psi_n\}$  has property (3.3). The proofs of Lemma 3.3 and Proposition 3.2 are now complete.  $\square$

### 3.4 The color equations

We continue in the setting of Proposition 3.2.

#### Nice functions

We say that a complex-valued function  $f$  on color space is *nice* (with respect to the kernel  $s$  and associated partition  $\mathcal{S}$  of  $[0, 1]$ ) if  $f$  has a Fourier expansion

$$f(c) = \sum_{i \in \mathbb{Z}} f_i(x) \xi^i \quad (c = (x, \xi) \in C)$$

where all but finitely many of the coefficient functions  $f_i : [0, 1] \rightarrow \mathbb{C}$  vanish identically, and every coefficient function  $f_i$  is constant on every interval of the partition  $\mathcal{S}$ . It is not difficult to see that  $\Phi_n$  and  $\Psi_n$  can be “corrected” on a set of  $P$ -measure zero in a unique way to become nice. So we may and we will assume

hereafter without any loss of generality that the functions  $\Phi_n$  and  $\Psi_n$  are nice and that all the relations asserted in Proposition 3.2 to hold  $P$ -a.e. in fact hold without exception.

### An auxiliary function

For all complex numbers  $|\lambda| > A$  and  $c \in C$  put

$$\Psi(c, \lambda) = \sum_{n=1}^{\infty} \Psi_n(c) \lambda^{-n},$$

defining a function depending holomorphically on  $\lambda$  and satisfying estimates

$$(3.8) \quad |\Psi(c, \lambda)| \leq \frac{1}{4} \frac{A^2}{|\lambda| - A}, \quad |\lambda| > 2A \Rightarrow |\Psi(c, \lambda)| < \frac{A^2}{2|\lambda|} \leq \frac{A}{4}$$

uniform in  $c$ .

### The equations

From the system of equations described in Proposition 3.2 we now deduce by the substitution  $t = 1/\lambda$  and application of dominated convergence the relations

$$(3.9) \quad \int \frac{s(c, c') P(dc')}{\lambda - \Psi(c', \lambda)} = \Psi(c, \lambda), \quad \int \frac{P(dc)}{\lambda - \Psi(c, \lambda)} = S(\lambda) = \int \frac{\mu(dx)}{\lambda - x}$$

which hold for all  $c \in C$  and  $|\lambda| > 2A$ . We call the relations (3.9) the *color equations*. Equations of this sort have appeared already in other contexts, see [BK99], [Gi90], [HLN05, eq. 2.2], [KKP96].

**Proposition 3.4.** *In the setting of the color equations, assume further that for some nice nonnegative function  $f$  on color space*

$$s(c, c') = f(c)f(c'), \quad \|f\|_{L^\infty(C)} = A/2, \quad \|f\|_{L^1(C)} = 1.$$

Let  $\mu_f$  be the law of  $f$  viewed as a random variable on  $C$  under  $P$ . Let  $S_f(\lambda)$  be the Stieltjes transform of  $\mu_f$ . Then: (i) There exists a function  $w(\lambda)$  defined and holomorphic for  $|\lambda| \gg 0$  such that

$$(3.10) \quad \lim_{|\lambda| \rightarrow \infty} w(\lambda) \lambda = 1,$$

$$(3.11) \quad \lambda S(\lambda) = 1 + w(\lambda)^2 = \frac{\lambda}{w(\lambda)} S_f\left(\frac{\lambda}{w(\lambda)}\right)$$

for  $|\lambda| \gg 0$ . (We emphasize that we do mean  $S$  on the LHS and  $S_f$  on the RHS.) (ii) The measure  $\mu$  is the free multiplicative convolution of  $\mu_f$  with the semicircle law of mean 0 and variance 1.

*Proof.* (i) Put

$$(3.12) \quad w(\lambda) = \int \frac{f(c)P(dc)}{\lambda - \Psi(c, \lambda)} = \int \Psi(c, \lambda)P(dc),$$

thus defining a holomorphic function in the domain  $|\lambda| > 2A$  such that

$$w(\lambda) = \frac{1}{\lambda} + O\left(\frac{1}{|\lambda|^2}\right), \quad \Psi(c, \lambda) = w(\lambda)f(c).$$

By definition

$$S_f(\lambda) = \int \frac{P(dc)}{\lambda - f(c)},$$

which is a function holomorphic in the domain  $|\lambda| > A/2$ . For  $|\lambda| \gg 1$  we have by (3.9) and the first equality in (3.12) that

$$\begin{aligned} \lambda S(\lambda) &= \int \frac{\lambda P(dc)}{\lambda - w(\lambda)f(c)} = \frac{\lambda}{w(\lambda)} S_f\left(\frac{\lambda}{w(\lambda)}\right), \\ \lambda S(\lambda) - 1 &= \int \frac{w(\lambda)f(c)P(dc)}{\lambda - w(\lambda)f(c)} = w(\lambda)^2, \end{aligned}$$

which proves the result.

(ii) We return to the setting of §2.4. Let  $W^{(N)}$  be the  $N^2$ -by- $N^2$  matrix with entries

$$W_{ij}^{(N)} = 2^{\delta_{ij}/2} \tilde{Y}_{\{c_i^{(N)}, c_j^{(N)}\}} / N.$$

Note that  $W^{(N)}$  belongs to the GOE (the factor  $2^{\delta_{ij}/2}$  is needed for orthogonal invariance). Let  $\Lambda^{(N)}$  be the  $N^2$ -by- $N^2$  deterministic diagonal matrix with diagonal entries

$$\Lambda_{ii}^{(N)} = \sqrt{f(c_i^{(N)})}.$$

Then we have

$$\tilde{X}^{(N)} / N = \Lambda^{(N)} W^{(N)} \Lambda^{(N)}.$$

Furthermore, as  $N \rightarrow \infty$ , the empirical distribution of eigenvalues of  $W^{(N)}$  (resp.,  $(\Lambda^{(N)})^2$ ) tends to the semicircle law (resp.,  $\mu_f$ ). Finally, since  $W^{(N)}$  and  $\Lambda^{(N)}$  are asymptotically freely independent, see [HP00, Corollary 4.3.6] and the discussion on page 157 there concerning the extension from the unitary to orthogonal case,  $\mu$  has the claimed form of free multiplicative convolution.  $\square$

### 3.5 Analysis of (NE+SE+SW+NW)-filtered Wigner matrix

We consider the setup of §2.3 in the special case of (NE+SE+SW+NW)-filtered Wigner matrices. We are thus considering the model mentioned in the introduction. Proposition 3.4 applies, with  $\mu_f$  equal to the law of  $2\cos^2\theta = 1 + \cos 2\theta$  with  $\theta$  distributed uniformly in  $[0, 2\pi)$ . Integrating, we get

$$S_f(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\lambda - 1 - \cos 2\theta} = \frac{1}{\sqrt{\lambda(\lambda - 2)}}.$$

It follows from (3.11) that

$$(1 + w^2)^2 = \frac{\lambda}{\lambda - 2w},$$

and hence (after taking out an irrelevant factor of  $w$ )

$$2w^4 - \lambda w^3 + 4w^2 - 2\lambda w + 2 = 0.$$

In turn, after forming the resultant of lefthand side above and  $1 + w^2 - \lambda S$  with respect to  $w$  and taking out irrelevant factors, we get the equation

$$(3.13) \quad 4\lambda^2 S^4 - \lambda^3 S^3 - \lambda^2 S^2 + \lambda S + 1 = 0.$$

Since the equation is quartic, it can in principle be solved explicitly by root extractions. We omit the details, which are available from the authors. Fortunately explicit formulas for  $S(\lambda)$  are not needed to find out the key features of  $\mu$ . We can use Theorem 2.5 and its proof to get information without having to calculate too much. The discriminant of the left side of (3.13) is

$$-16\lambda^6(8\lambda^4 + 107\lambda^2 - 1024).$$

The only nonzero real roots of the discriminant are

$$(3.14) \quad \pm \frac{1}{4} \sqrt{-107 + 51\sqrt{17}} \quad (\text{approximately } \pm 2.5406).$$

Since  $\mu$  is symmetric and cannot be supported on finitely many points (equation (3.13) cannot be solved by any rational function of  $\lambda$ ), Theorem 2.5 leaves no choice but that the support of  $\mu$  be the interval with endpoints specified in (3.14). To finish our discussion let us verify that that  $d\mu/dx$  has a spike at the origin proportional to  $1/\sqrt{|x|}$ . Let  $\sqrt{\lambda}$  be the unique square root of  $\lambda$  in  $\mathbb{C} \setminus -i[0, \infty)$  which is positive along the positive real axis. For small enough  $\varepsilon > 0$  there are exactly four analytic solutions of (3.13) in the slit disk  $\{|\lambda| < \varepsilon\} \setminus -i[0, \infty)$ , and these are of the form  $i^\nu(\sqrt{i/\sqrt{2\lambda}}) + O_{\lambda \rightarrow 0}(1)$  for  $\nu = 0, 1, 2, 3$ . This is verified by using the fact that such solutions have in any case for  $\varepsilon > 0$  sufficiently small an expansion on  $\{|\lambda| < \varepsilon\} \setminus -i[0, \infty)$  in fractional (possibly negative) powers of  $\lambda$ , and then one calculates the expansions by the method of undetermined coefficients. The only such solution with positive imaginary part on  $(-\varepsilon_1, 0) \cup (0, \varepsilon_1)$  for  $\varepsilon_1 > 0$  sufficiently small corresponds to  $\nu = 0$ ; this gives the spike.

#### 4 The main limit calculation

To the end of proving Theorem 2.3, we first prove the following result.

**Proposition 4.1.** *Let Assumption 2.2 hold. For each positive integer  $k$ ,*

$$\lim_{N \rightarrow \infty} N^{-k/2-1} E \text{trace}((X^{(N)})^k) = \sum_{\pi \in \mathcal{M}_k} \mathbb{E} M_\pi.$$

The proof requires some preparation and will not be completed until §4.5.

### 4.1 Notation, terminology and strategy

#### $(N, k)$ -words

Let  $N$  and  $k$  be positive integers. An  $(N, k)$ -word  $\mathbf{i}$  is by definition a function

$$\mathbf{i}: \{1, \dots, k\} \rightarrow \{1, \dots, N\}.$$

To each  $(N, k)$ -word we attach a random variable

$$X_{\mathbf{i}}^{(N)} = \prod_{\alpha=1}^k X_{\mathbf{i}(\alpha), \mathbf{i}(\eta_k(\alpha))}^{(N)}$$

where

$$\eta_k = (12 \cdots k) \in S_k.$$

(Here and below we employ cycle notation for permutations.) We have

$$\sum_{\mathbf{i}} EX_{\mathbf{i}}^{(N)} = E \text{ trace}((X^{(N)})^k),$$

where the sum on the left is extended over  $(N, k)$ -words  $\mathbf{i}$ .

#### The set partition associated to an $(N, k)$ -word

Given an  $(N, k)$ -word  $\mathbf{i}$ , consider the graph  $G_{\mathbf{i}}^K = (V_{\mathbf{i}}^K, E_{\mathbf{i}}^K)$ , where  $V_{\mathbf{i}}^K = \{1, \dots, k\}$  and

$$E_{\mathbf{i}}^K = \{\{\alpha, \beta\} \subset \{1, \dots, k\} \mid |\mathbf{i}(\alpha) - \mathbf{i}(\beta)| \leq K\}.$$

Here and below  $K$  is the constant figuring in Assumption 2.2(II). We define  $\pi_{\mathbf{i}}$  to be the set partition of  $k$  the parts of which are the equivalence classes for the relation “ $\alpha$  and  $\beta$  belong to the same connected component of  $G_{\mathbf{i}}^K$ ”. (Although the set partition  $\pi_{\mathbf{i}}$  depends on  $K$ , we suppress reference to  $K$  in the notation.) We have

$$|\mathbf{i}(\alpha) - \mathbf{i}(\beta)| \leq K \Rightarrow \pi_{\mathbf{i}}(\alpha) = \pi_{\mathbf{i}}(\beta), \quad |\mathbf{i}(\alpha) - \mathbf{i}(\beta)| > Kk \Rightarrow \pi_{\mathbf{i}}(\alpha) \neq \pi_{\mathbf{i}}(\beta),$$

for all  $\alpha, \beta \in \{1, \dots, k\}$ .

#### Distinguished $(N, k)$ -words

Let  $\mathbf{i}$  be an  $(N, k)$ -word. Consider the following conditions:

- (I)  $\pi_{\mathbf{i}} \in \mathscr{W}_k$ .
- (II) For all  $A \in \pi_{\mathbf{i}}$  we have  $\mathbf{i}(\min A) \in Q_{(k+1)K}^{(N)}$ .
- (III) For all distinct  $A, B \in \pi_{\mathbf{i}}$  we have  $|\mathbf{i}(\min A) - \mathbf{i}(\min B)| > 5kK$ .
- (IV)  $\max_{A \in \pi_{\mathbf{i}}} \max_{\alpha, \beta \in A} |\mathbf{i}(\alpha) - \mathbf{i}(\beta)| \leq Kk$ .
- (V)  $\min_{\alpha=1}^k |\mathbf{i}(\alpha) - \mathbf{i}(\eta_k(\alpha))| > 3Kk$ .
- (VI) For  $\alpha = 1, \dots, k$  we have  $\mathbf{i}(\alpha) \in Q_K^{(N)}$ .
- (VII) For all  $A \in \pi$  and  $\alpha, \beta \in A$ , the numbers  $\frac{\mathbf{i}(\alpha)}{N}$  and  $\frac{\mathbf{i}(\beta)}{N}$  belong to the same interval of the partition  $\mathscr{I}$ .

If conditions (I,II,III) hold we say that  $\mathbf{i}$  is *distinguished*, in which case  $\mathbf{i}$  automatically also satisfies conditions (IV,V,VI,VII).

### Strategy

We will show that the only nonnegligible contributions to

$$(4.1) \quad \lim_{N \rightarrow \infty} N^{-k/2-1} \sum_{\mathbf{i}: (N,k)\text{-word}} EX_{\mathbf{i}}^{(N)} = \lim_{N \rightarrow \infty} N^{-k/2-1} E \text{trace}((X^{(N)})^k)$$

come from distinguished  $(N,k)$ -words. Then we will evaluate  $EX_{\mathbf{i}}^{(N)}$  for distinguished  $\mathbf{i}$ . Finally, we will calculate the limit on the left with the summation restricted to distinguished  $\mathbf{i}$ .

### 4.2 Negligibility of nondistinguished $(N,k)$ -words

**Lemma 4.2.** *Let  $\pi$  be a set partition of  $k$ . There exists  $C_{\pi} > 0$  such that for every positive integer  $N$  the sum  $\sum_{\mathbf{i}} E|X_{\mathbf{i}}^{(N)}|$  extended over  $(N,k)$ -words  $\mathbf{i}$  such that  $\pi = \pi_{\mathbf{i}}$  does not exceed  $C_{\pi} N^{|\pi|}$ .*

*Proof.* By Assumption 2.2(Ib) and the Hölder inequality, it suffices simply to estimate the number of  $(N,k)$ -words such that  $\pi = \pi_{\mathbf{i}}$ . A crude estimate of the latter is  $(1 + 2Kk)^{k-|\pi|} N^{|\pi|}$ .  $\square$

**Lemma 4.3.** *Let  $\pi$  be a set partition of  $k$ . Let  $\mathbf{i}$  be an  $(N,k)$ -word such that  $\pi = \pi_{\mathbf{i}}$ . If  $|\pi| \geq k/2 + 1$  and  $EX_{\mathbf{i}}^{(N)} \neq 0$ , then  $\pi \in \mathcal{W}_k$  (and hence  $|\pi| = k/2 + 1$ ).*

*Proof.* Let  $G_{\pi} = (V_{\pi}, E_{\pi})$  be the graph associated to  $\pi$ . For each  $\alpha \in \{1, \dots, k\}$  put  $e(\alpha) = \{\pi(\alpha), \pi(\alpha + 1)\} \in E_{\pi}$ . Now fix  $e \in E_{\pi}$ . It is enough to show that  $e = e(\alpha)$  for at least two  $\alpha \in \{1, \dots, k\}$ , for then, since  $G_{\pi}$  is connected, we have

$$k/2 + 1 \leq |\pi| = |V_{\pi}| \leq |E_{\pi}| + 1 \leq k/2 + 1,$$

and hence  $\pi$  is a Wigner set partition. To derive a contradiction, suppose rather that  $e = e(\alpha)$  for unique  $\alpha \in \{1, \dots, k\}$ . For every  $\gamma \in \{1, \dots, k\}$  let  $i(\gamma) \leq j(\gamma)$  be the integers  $\mathbf{i}(\gamma)$  and  $\mathbf{i}(\eta_k(\gamma))$  rearranged. Then for every  $\beta \in \{1, \dots, k\} \setminus \{\alpha\}$  we have  $\max(|i(\alpha) - i(\beta)|, |j(\alpha) - j(\beta)|) > K$ , for otherwise  $e(\alpha) = e(\beta)$ . It follows by Assumption 2.2(IIb) that the random variable  $X_{\mathbf{i}(\alpha), \mathbf{i}(\eta_k(\alpha))}^{(N)}$  is independent of the rest of the random variables appearing in the product  $X_{\mathbf{i}}^{(N)}$ , and hence  $EX_{\mathbf{i}}^{(N)} = 0$  by Assumption 2.2(Ia). This contradiction proves the lemma.  $\square$

**Lemma 4.4.** *Let  $\pi$  be a Wigner set partition. There exists  $C'_{\pi} > 0$  such that for every positive integer  $N$  the sum  $\sum_{\mathbf{i}} E|X_{\mathbf{i}}^{(N)}|$  extended over  $(N,k)$ -words  $\mathbf{i}$  such that  $\pi = \pi_{\mathbf{i}}$  but  $\mathbf{i}$  is not distinguished does not exceed  $C'_{\pi} N^{k/2}$ .*

*Proof.* The proof is similar to that of Lemma 4.2. We omit the details.  $\square$

### 4.3 Evaluation of $EX_{\mathbf{i}}^{(N)}$ for distinguished $\mathbf{i}$

#### Definitions of $\tau_\pi$ and $\sigma_\pi$

Let  $\pi$  be a set partition of  $k$ . Enumerate  $\pi$  and its parts thus:

$$\begin{aligned}\pi &= \{I_1, \dots, I_{|\pi|}\}, \quad \min I_1 < \dots < \min I_{|\pi|}, \\ I_\alpha &= \{i_{\alpha 1} < \dots < i_{\alpha, |I_\alpha|}\} \text{ for } \alpha = 1, \dots, |\pi|.\end{aligned}$$

Put

$$\tau_\pi = (i_{11} \cdots i_{1, |I_1|}) \cdots (i_{|\pi|, 1} \cdots i_{|\pi|, |I_{|\pi|}|}) \in S_k, \quad \sigma_\pi = \eta_k^{-1} \tau_\pi \in S_k.$$

By construction

$$\pi(\tau_\pi(i)) = \pi(i)$$

for  $i = 1, \dots, k$ .

**Lemma 4.5.** *Let  $\pi \in \mathscr{W}_k$  be a Wigner set partition. Then the permutation  $\sigma_\pi$  is fixed-point-free and squares to the identity. Furthermore, for all distinct  $i, j \in \{1, \dots, k\}$ , we have  $\{\pi(i), \pi(i+1)\} = \{\pi(j), \pi(j+1)\} \Leftrightarrow j = \sigma_\pi(i)$ .*

*Proof.* Let  $G_\pi = (V_\pi, E_\pi)$  be the graph (tree) associated to  $\pi$ . For each  $e \in E_\pi$ , there are by Lemma 3.1(ii) exactly two indices  $i$  such that  $e = \{\pi(i), \pi(i+1)\}$ , and  $\sigma_\pi$  swaps them by Lemma 3.1(iii).  $\square$

**Lemma 4.6.** *Let  $\pi \in \mathscr{W}_k$  be a Wigner set partition. Put  $\sigma = \sigma_\pi$  and  $\tau = \tau_\pi$ . Then we have*

$$EX_{\mathbf{i}}^{(N)} = \prod_{\substack{\alpha \in \{1, \dots, k\} \\ \text{s.t. } \alpha \leq \sigma(\alpha)}} EX_{\mathbf{i}(\alpha), \mathbf{i}(\tau(\sigma(\alpha)))}^{(N)} X_{\mathbf{i}(\sigma(\alpha)), \mathbf{i}(\tau(\alpha))}^{(N)}$$

for every  $(N, k)$ -word  $\mathbf{i}$  such that  $\pi_{\mathbf{i}} = \pi$ .

*Proof.* Put  $\eta = \eta_k$ . By definition we have  $\eta\sigma = \tau$ . By the previous lemma we have  $\tau\sigma = \eta$ . Therefore after rearranging the product

$$X_{\mathbf{i}}^{(N)} = \prod_{\alpha \in \{1, \dots, k\}} X_{\mathbf{i}(\alpha), \mathbf{i}(\eta(\alpha))},$$

we have

$$X_{\mathbf{i}}^{(N)} = \prod_{\substack{\alpha \in \{1, \dots, k\} \\ \text{s.t. } \alpha \leq \sigma(\alpha)}} X_{\mathbf{i}(\alpha), \mathbf{i}(\tau(\sigma(\alpha)))}^{(N)} X_{\mathbf{i}(\sigma(\alpha)), \mathbf{i}(\tau(\alpha))}^{(N)}.$$

It suffices to prove enough independence to justify pushing the expectation under the product. For every  $\gamma \in \{1, \dots, k\}$  let  $i(\gamma) \leq j(\gamma)$  be the integers  $\mathbf{i}(\gamma)$  and  $\mathbf{i}(\eta(\gamma))$  rearranged. By definition of  $\pi_{\mathbf{i}}$  and the preceding lemma, for all distinct  $\alpha, \beta \in \{1, \dots, k\}$ , if  $\max(|i(\alpha) - i(\beta)|, |j(\alpha) - j(\beta)|) \leq K$ , then  $\beta = \sigma(\alpha)$ . By Assumption 2.2(IIb) we deduce the desired independence.  $\square$

### The difference operator associated to a set partition $\pi$

Let  $\pi$  be a set partition of  $k$ . Let  $\tau = \tau_\pi \in S_k$  be the canonically associated permutation. Let  $f : \{1, \dots, k\} \rightarrow \mathbb{Z}$  be a function. We define  $\partial_\pi f : \{1, \dots, k\} \rightarrow \mathbb{Z}$  by the formula  $(\partial_\pi f)(i) = f(i) - f(\tau(i))$  for  $i = 1, \dots, k$ . Given a function  $g : \{1, \dots, k\} \rightarrow \mathbb{Z}$ , the equation  $\partial_\pi f = g$  has a solution  $f : \{1, \dots, k\} \rightarrow \mathbb{Z}$  if and only if  $\sum_{\alpha \in A} g(\alpha) = 0$  for all parts  $A \in \pi$ , and  $f$  is unique up to the addition of a function  $\{1, \dots, k\} \rightarrow \mathbb{Z}$  constant on every part of  $\pi$ .

**Lemma 4.7.** *Let  $\pi \in \mathcal{W}_k$  be a Wigner set partition. Put  $\sigma = \sigma_\pi$ ,  $\tau = \tau_\pi$  and  $\partial = \partial_\pi$ . Fix  $\alpha \in \{1, \dots, k\}$  such that  $\alpha \leq \sigma(\alpha)$ . We have*

$$(4.2) \quad EX_{\mathbf{i}(\alpha), \mathbf{i}(\tau(\sigma(\alpha)))}^{(N)} X_{\mathbf{i}(\sigma(\alpha)), \mathbf{i}(\tau(\alpha))}^{(N)} = s_{(\partial \mathbf{i})(\alpha), (\partial \mathbf{i})(\sigma(\alpha))} \left( \frac{\mathbf{i}(\min \pi(\alpha))}{N}, \frac{\mathbf{i}(\min \pi(\sigma(\alpha)))}{N} \right)$$

for every distinguished  $(N, k)$ -word  $\mathbf{i}$  such that  $\pi_{\mathbf{i}} = \pi$ .

In particular, the expectation in question vanishes unless  $|\partial \mathbf{i}| \leq K$  by Assumption 2.2(IIa).

*Proof.* Let  $E(\alpha)$  be the left side of (4.2). Assume at first that  $\mathbf{i}(\alpha) \leq \mathbf{i}(\tau(\sigma(\alpha)))$ . Then  $\mathbf{i}(\alpha) + 3Kk < \mathbf{i}(\tau(\sigma(\alpha)))$  by property (V) of a distinguished  $(N, k)$ -word, and hence  $\mathbf{i}(\tau(\alpha)) + Kk < \mathbf{i}(\sigma(\alpha))$  by property (IV) of a distinguished  $(N, k)$ -word. So we have

$$(4.3) \quad E(\alpha) = EX_{\mathbf{i}(\alpha), \mathbf{i}(\tau(\sigma(\alpha)))}^{(N)} \bar{X}_{\mathbf{i}(\tau(\alpha)), \mathbf{i}(\sigma(\alpha))}^{(N)} = s_{(\partial \mathbf{i})(\alpha), (\partial \mathbf{i})(\sigma(\alpha))} \left( \frac{\mathbf{i}(\alpha)}{N}, \frac{\mathbf{i}(\tau(\sigma(\alpha)))}{N} \right)$$

by Assumption 2.2(IIc) and property (VI) of a distinguished  $(N, k)$ -word. Then we deduce that the desired formula holds by Assumption 2.1(III) and property (VII) of a distinguished  $(N, k)$ -word. Assume next and finally that  $\mathbf{i}(\alpha) \geq \mathbf{i}(\tau(\sigma(\alpha)))$ . Reasoning as above we have

$$E(\alpha) = EX_{\mathbf{i}(\sigma(\alpha)), \mathbf{i}(\tau(\alpha))}^{(N)} \bar{X}_{\mathbf{i}(\tau(\sigma(\alpha))), \mathbf{i}(\alpha)}^{(N)} = s_{(\partial \mathbf{i})(\sigma(\alpha)), (\partial \mathbf{i})(\alpha)} \left( \frac{\mathbf{i}(\sigma(\alpha))}{N}, \frac{\mathbf{i}(\tau(\alpha))}{N} \right).$$

Now we apply the symmetry (2.1), and then continue to reason as above. We deduce the desired formula just as before.  $\square$

### 4.4 Contribution of $EX_{\mathbf{i}}^{(N)}$ to the limit for distinguished $\mathbf{i}$

**Lemma 4.8.** *Let  $\pi \in \mathcal{W}_k$  and an  $(N, k)$ -word  $\mathbf{i}$  be given such that the following hold:*

- For all  $A \in \pi$  we have  $\mathbf{i}(\min A) \in \mathcal{Q}_{(k+1)K}^{(N)}$ .
- For all distinct  $A, B \in \pi$  we have  $|\mathbf{i}(\min A) - \mathbf{i}(\min B)| > 5Kk$ .
- $|\partial \mathbf{i}| \leq K$ .

Then  $\pi_{\mathbf{i}} = \pi$  and (hence)  $\pi$  is distinguished.

No proof is needed, but this point bears emphasis as an important step in the proof of Proposition 4.1.

**Lemma 4.9.** Fix a Wigner set partition  $\pi \in \mathcal{W}_k$ . Then we have

$$(4.4) \quad \mathbb{E}M_\pi = \lim_{N \rightarrow \infty} N^{-k/2-1} \sum_{\mathbf{i}} EX_{\mathbf{i}}^{(N)},$$

where the sum is extended over distinguished  $(N, k)$ -words  $\mathbf{i}$  such that  $\pi_{\mathbf{i}} = \pi$ .

*Proof.* Let  $\partial = \partial_\pi$ ,  $\sigma = \sigma_\pi$ , and  $\tau = \tau_\pi$ . Let  $\{t_A\}$  (resp.,  $\{z_A\}$ ) be an i.i.d. family of random variables uniform in  $[0, 1]$  (resp.,  $S^1$ ), indexed by finite nonempty sets  $A$  of positive integers. We further suppose that the families  $\{t_A\}$  and  $\{z_A\}$  are defined on a common probability space and are independent. We denote expectations with respect to these variables by  $\mathbf{E}$ . We have by Lemma 4.5 and the definitions that

$$\begin{aligned} \mathbb{E}M_\pi &= \mathbf{E} \prod_{\substack{i \in \{1, \dots, k\} \\ \text{s.t. } i \leq \sigma(i)}} \left( \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} s_{mn}(t_{\pi(i)}, t_{\pi(\sigma(i))}) z_{\pi(i)}^m z_{\pi(\sigma(i))}^n \right) \\ &= \sum_{f: \{1, \dots, k\} \rightarrow \mathbb{Z}} \mathbf{E} \prod_{\substack{i \in \{1, \dots, k\} \\ \text{s.t. } i \leq \sigma(i)}} s_{f(i), f(\sigma(i))}(t_{\pi(i)}, t_{\pi(\sigma(i))}) \cdot \prod_{A \in \pi} z_A^{\sum_{i \in A} f(i)} \\ &= \sum_{\substack{f: \{1, \dots, k\} \rightarrow \mathbb{Z} \cap [-K, K] \\ \forall A \in \pi, \sum_{i \in A} f(i) = 0}} \mathbf{E} \prod_{\substack{i \in \{1, \dots, k\} \\ \text{s.t. } i \leq \sigma(i)}} s_{f(i), f(\sigma(i))}(t_{\pi(i)}, t_{\pi(\sigma(i))}). \end{aligned}$$

At the last step we integrate out the  $z$ 's and take into account Assumption 2.2(IIa). We then have

$$\mathbb{E}M_\pi = \sum_{\substack{f: \{1, \dots, k\} \rightarrow \mathbb{Z} \\ |\partial f| \leq K \\ \forall A \in \pi f(\min A) = f_0(A)}} \mathbf{E} \prod_{\substack{i \in \{1, \dots, k\} \\ \text{s.t. } i \leq \sigma(i)}} s_{(\partial f)(i), (\partial f)(\sigma(i))}(t_{\pi(i)}, t_{\pi(\sigma(i))}),$$

where  $f_0 : \pi \rightarrow \mathbb{Z}$  is any fixed function defined on the parts of  $\pi$ . After some straightforward bookkeeping which we omit, it follows by the preceding two lemmas that (4.4) holds with the summation on  $\mathbf{i}$  restricted to distinguished  $(N, k)$ -words such that  $\pi_{\mathbf{i}} = \pi$  and  $|\partial \mathbf{i}| \leq K$ . But then by Assumption 2.2(IIa), the limit does not change if we drop the restriction  $|\partial \mathbf{i}| \leq K$  (the further terms all vanish), whence the result.  $\square$

## 4.5 Completion of the proof of Proposition 4.1

The suite of lemmas proved in §4.2 shows that the limit on the left side of equation (4.1) does not change if we restrict attention to distinguished  $(N, k)$ -words. The last lemma above evaluates the limit on the left side of (4.1) with  $\mathbf{i}$  restricted to distinguished  $(N, k)$ -words, and gives the desired value. The proof of Proposition 4.1 is complete.  $\square$

### 5 Completion of the proof of Theorem 2.3

Fix a positive integer  $k$ . As in [AZ06, pf. of Thm. 3.2, Section 6, p. 305], Theorem 2.3 will follow as soon as we can prove that

$$(5.1) \quad \text{Var}(\langle L^{(N)}, x^k \rangle) \rightarrow_{N \rightarrow \infty} 0.$$

We will prove this by lightly modifying the arguments of §4.2. Now

$$(5.2) \quad \text{Var}(\langle L^{(N)}, x^k \rangle) = N^{-k-2} \sum_{(\mathbf{i}, \mathbf{j})} \left( E[X_{\mathbf{i}}^{(N)} X_{\mathbf{j}}^{(N)}] - E[X_{\mathbf{i}}^{(N)}] E[X_{\mathbf{j}}^{(N)}] \right)$$

where the sum is extended over pairs  $(\mathbf{i}, \mathbf{j})$  of  $(N, k)$ -words. By Assumption 2.2(Ib) and the Hölder inequality, it is (more than) enough to show that the number of pairs  $(\mathbf{i}, \mathbf{j})$  making a nonzero contribution to the sum on the right side of (5.2) is  $O(N^{k+1})$ . Now fix a pair  $(\mathbf{i}, \mathbf{j})$  of  $(N, k)$ -words indexing a nonzero term in the sum on the right side of (5.2). Let  $\mathbf{ij}$  be the  $(N, 2k)$ -word obtained by concatenating  $\mathbf{i}$  and  $\mathbf{j}$ , i. e.,

$$\mathbf{ij}(\alpha) = \begin{cases} \mathbf{i}(\alpha) & \text{if } \alpha \in \{1, \dots, k\}, \\ \mathbf{j}(\alpha - k) & \text{if } \alpha \in \{k+1, \dots, 2k\}. \end{cases}$$

Consider the set partition  $\pi = \pi_{\mathbf{ij}}$  defined in §4.1. We need also to consider a graph associated to  $\pi$ . The relevant graph is no longer  $G_\pi$ , but rather a slightly modified version  $\tilde{G}_\pi = (\tilde{V}_\pi, \tilde{E}_\pi)$ , where  $\tilde{V}_\pi = \pi$  and

$$\tilde{E}_\pi = \{ \{ \pi(1), \pi(2) \}, \dots, \{ \pi(k), \pi(1) \} \} \cup \{ \{ \pi(k+1), \pi(k+2) \}, \dots, \{ \pi(2k), \pi(k+1) \} \}.$$

By construction  $\tilde{G}_\pi$  comes equipped with two walks, namely

$$\pi(1), \dots, \pi(k), \pi(1) \quad \text{and} \quad \pi(k+1), \dots, \pi(2k), \pi(k+1).$$

Arguing as in the proof of Lemma 4.3, we find that nonvanishing of the term on the right side of (5.2) indexed by  $(\mathbf{i}, \mathbf{j})$  implies that the walks jointly visit every edge of  $\tilde{G}_\pi$  at least twice, and moreover there must exist some edge of  $\tilde{G}_\pi$  visited by both walks. Thus  $|\tilde{E}_\pi| \leq k$  and  $\tilde{G}_\pi$  is connected. It follows that  $|\pi| = |\tilde{V}_\pi| \leq k+1$ . Finally, arguing as in the proof of Lemma 4.2, we find that the number of nonzero terms on the right side of (5.2) is indeed  $O(N^{k+1})$ . The proof of Theorem 2.3 is complete.  $\square$

### 6 An algebraicity criterion

In this section, which is completely independent of the preceding ones, we develop a “soft” method for proving that a holomorphic function is algebraic under hypotheses commonly encountered in random matrix theory.

## 6.1 Formulation of an algebraicity criterion

### Notation

Let  $\{X_i\}_{i=1}^\infty$  be independent (algebraic) variables. Let  $\mathbb{C}[X_1, \dots, X_n]$  denote the ring of polynomials in  $X_1, \dots, X_n$  with coefficients in  $\mathbb{C}$ . We view  $\mathbb{C}[X_1, \dots, X_n]$  as a subring of  $\mathbb{C}[X_1, \dots, X_{n+1}]$ . Given  $F \in \mathbb{C}[X_1, \dots, X_n]$ , let  $F(0) \in \mathbb{C}$  be the result of setting  $X_1 = \dots = X_n = 0$ . Let  $\mathbb{C}(X_1, \dots, X_n)$  be the field of rational functions in the variables  $X_1, \dots, X_n$ , i. e., the ring consisting of fractions  $A/B$  where  $A, B \in \mathbb{C}[X_1, \dots, X_n]$  and  $B$  does not vanish identically. We say that  $F \in \mathbb{C}(X_1, \dots, X_n)$  is *defined at the origin* if  $F = A/B$  with  $A, B \in \mathbb{C}[X_1, \dots, X_n]$  such that  $B(0) \neq 0$ , in which case we put  $F(0) = A(0)/B(0)$ .

### DIRE families

Let  $\varphi_1, \dots, \varphi_N$  be a finite family of holomorphic functions each defined in a connected open neighborhood of the origin in  $\mathbb{C}^n$ ; the domains need not be the same. We will say that  $\varphi_1, \dots, \varphi_N$  are *defined implicitly by rational equations* (*DIRE* for short) if there exist  $\Phi_1, \dots, \Phi_N \in \mathbb{C}(X_1, \dots, X_{n+N})$  such that

- (I)  $\Phi_i$  is defined at the origin and  $\Phi_i(0) = 0$  for  $i = 1, \dots, N$ ,
- (II)  $(\det_{i,j=1}^N \frac{\partial \Phi_i}{\partial X_{j+n}})(0) \neq 0$ , and
- (III)  $\Phi_i(z_1, \dots, z_n, \varphi_1(z) - \varphi_1(0), \dots, \varphi_N(z) - \varphi_N(0)) = 0$  for  $i = 1, \dots, N$  and  $z = (z_1, \dots, z_n) \in \mathbb{C}^n$  sufficiently near the origin.

The relationship between this definition and the implicit function theorem for holomorphic functions [Ca63, Proposition 6.1] is close. Indeed, given  $\Phi_1, \dots, \Phi_N \in \mathbb{C}(X_1, \dots, X_{n+N})$  with properties (I,II) above, the implicit function theorem for holomorphic functions says that there exist holomorphic functions  $\varphi_1, \dots, \varphi_N$  each defined in a connected open neighborhood of the origin such that (III) holds, and the theorem further asserts uniqueness of these functions in the sense that if  $\psi_1, \dots, \psi_N$  are holomorphic functions each defined in a connected open neighborhood of the origin in  $\mathbb{C}^n$  and also satisfying (III), then for  $i = 1, \dots, N$  there exists a neighborhood of the origin in which  $\psi_i$  and  $\varphi_i$  differ by a constant.

In the sequel, for brevity, when we say “ $\varphi_1, \dots, \varphi_N$  is an  $n$ -variable DIRE family”, this is short for the assertion that “ $\varphi_1, \dots, \varphi_N$  are holomorphic functions each defined in some connected open neighborhood of the origin in  $\mathbb{C}^n$  which together have the property of being defined implicitly by rational equations”.

### Algebraic functions

A holomorphic function  $\varphi$  defined in a nonempty open subset  $D \subset \mathbb{C}^n$  is called an  *$n$ -variable algebraic function* if there exists a not-identically-vanishing polynomial  $F = F(X_1, \dots, X_{n+1}) \in \mathbb{C}[X_1, \dots, X_{n+1}]$  such that  $F(z_1, \dots, z_n, \varphi(z)) = 0$  for all  $z = (z_1, \dots, z_n) \in D$ . If  $D$  is connected and  $U \subset D$  is any nonempty open subset, then algebraicity of  $F$  in  $U$  implies algebraicity of  $F$  in  $D$  by the principle of analytic continuation.

The main result of Section 6 is the following.

**Theorem 6.1.** *Let  $\varphi_1, \dots, \varphi_N$  be an  $n$ -variable DIRE family. Then each  $\varphi_i$  is an  $n$ -variable algebraic function.*

The proof takes up the last several subsections of Section 6. Before turning to the proof we give key examples of DIRE families, and describe techniques for constructing new DIRE families from old.

**Proposition 6.2.** *Fix a positive integer  $\ell$  and put  $L = 2\ell + 1$ . For  $z = (z_1, \dots, z_L) \in \mathbb{C}^L$  such that  $\sum_{i=1}^L |z_i| < 1$  and integers  $j = 1, \dots, L$ , consider the quantity*

$$(6.1) \quad \vartheta_j(z) = -\delta_{j,\ell+1} + \frac{1}{2\pi} \int_0^{2\pi} \frac{\exp(-\mathbf{i}(j-\ell-1)x) dx}{1 - \sum_{k=1}^L z_k \exp(\mathbf{i}(k-\ell-1)x)},$$

where  $\mathbf{i}^2 = -1$ , which depends holomorphically on  $z$  and vanishes for  $z = 0$ . Then the family  $\vartheta_1, \dots, \vartheta_L$  can be extended to an  $L$ -variable DIRE family  $\vartheta_1, \dots, \vartheta_N$ .

*Proof.* Fix  $z = (z_1, \dots, z_L) \in \mathbb{C}^L$  such that  $\sum_{i=1}^L |z_i| < 1$ . Let  $p = [p_{ij}]_{i,j \in \mathbb{Z}}$  be the unique matrix of complex numbers with rows and columns indexed by  $\mathbb{Z}$  with the following properties:

- $p_{ij} = 0$  for all  $i$  and  $j$  such that  $|i - j| > \ell$ .
- $p_{ij} = z_{j-i+\ell+1}$  for all  $i$  and  $j$  such that  $|i - j| \leq \ell$ .

Let  $|p|$  be the  $\mathbb{Z}$ -by- $\mathbb{Z}$  matrix with entries  $|p_{ij}|$ . We have the crude estimate

$$(6.2) \quad (|p|^t)_{ij} \leq \left( \sum_{i=1}^L |z_i| \right)^t$$

holding for all positive integers  $t$ .

Now supposing  $p$  were a Markov matrix (which of course it is not), we could view each entry  $p_{ij}$  as a transition probability

$$P(S_{t+1} = j \mid S_t = i) = p_{ij},$$

for a random walk  $\{S_t\}_{t=0}^\infty$  on  $\mathbb{Z}$  with step-length bounded by  $\ell$  and we would have, for any subset  $M \subset \mathbb{Z}^t$ , an equality

$$P((S_1, \dots, S_t) \in M \mid S_0 = i_0) = \sum_{(i_1, \dots, i_t) \in M} p_{i_0 i_1} \cdots p_{i_{t-1} i_t}.$$

Because it is a valuable aid to intuition, we will make the line above a definition. Notice that by (6.2) the sum on the right is absolutely convergent, and that the sum of the absolute values of the terms is  $\leq (\sum_{i=1}^L |z_i|)^t$ . We will be able to calculate using the usual rules of probability, provided we never invoke positivity  $p_{ij} \geq 0$  or the Markov property  $\sum_j p_{ij} = 1$ . We are interested in functions of  $z$  describable in the language of random walk because, as we explain presently, the functions  $\vartheta_j$  belong to this class, and moreover within this class we can easily find the ‘‘extra’’ functions needed to extend  $\vartheta_1, \dots, \vartheta_L$  to a DIRE family.

Let  $U, V, A, B, C$  be  $\ell$ -by- $\ell$  matrices with complex entries defined as follows:

$$\begin{aligned} U_{ij} &= \sum_{t=1}^{\infty} P(S_t = j, \min_{u=0}^t S_u > 0 | S_0 = i), \\ V_{ij} &= \sum_{t=1}^{\infty} P(S_t = j, \max_{u=0}^t S_u < \ell + 1 | S_0 = i), \\ A_{ij} &= P(S_1 = j + \ell | S_0 = i), \\ B_{ij} &= P(S_1 = j | S_0 = i), \\ C_{ij} &= P(S_1 = j - \ell | S_0 = i). \end{aligned}$$

Let  $W, D$  be  $L$ -by- $L$  matrices with complex entries defined as follows:

$$\begin{aligned} W_{ij} &= \sum_{t=1}^{\infty} P(S_t = j | S_0 = i), \\ D_{ij} &= P(S_1 = j | S_0 = i). \end{aligned}$$

By (6.2), all the series in question here converge absolutely and define functions of  $z$  holomorphic in the domain  $\sum_{i=1}^L |z_i| < 1$ . Notice that  $A, B, C, D$  depend linearly on  $z$ , and that all matrices  $A, B, C, D, U, V, W$  vanish for  $z = 0$ .

Consider the expansion of the integrand of (6.1) by geometric series and then integrate term by term. One finds in this way a series expression for  $\vartheta_j$  identical to the series expression defining some entry of the matrix  $W$ . In short, every  $\vartheta_j$  appears in  $W$ .

By breaking paths down according to visits to the sets  $\{-\ell, \dots, -1\}$ ,  $\{0\}$  and  $\{1, \dots, \ell\}$ , we obtain in the usual way recursions

$$(6.3) \quad \begin{aligned} 0 &= -U + B + A(1 + U)C(1 + U), \\ 0 &= -V + B + C(1 + V)A(1 + V), \\ 0 &= -W + D + \begin{bmatrix} C(1 + V)A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A(1 + U)C \end{bmatrix} (1 + W). \end{aligned}$$

Extend the given family  $\vartheta_1, \dots, \vartheta_n$  to an enumeration  $\vartheta_1, \dots, \vartheta_N$  of all entries of  $U, V$  and  $W$ . Rewrite the system of equations (6.3) as a system of  $N$  polynomial equations in  $z_1, \dots, z_L, \vartheta_1, \dots, \vartheta_N$ , in order to find polynomials

$$\Phi_i(X_1, \dots, X_{L+N}) \in \mathbb{C}[X_1, \dots, X_{L+N}] \text{ for } i = 1, \dots, N$$

such that

$$\Phi_i(0, \dots, 0, X_{L+1}, \dots, X_{L+N}) = -X_{i+L}, \quad \Phi_i(z_1, \dots, z_L, \vartheta_1, \dots, \vartheta_N) = 0.$$

Note that part (II) of the definition of DIRE family holds because  $\frac{\partial \Phi_i}{\partial X_{j+L}}(0) = -\delta_{ij}$ . Thus  $\vartheta_1, \dots, \vartheta_N$  is indeed a DIRE family.  $\square$

## 6.2 Natural operations on DIRE families

We write down some lemmas which will be helpful in applying the notion of DIRE family. The first three are trivial but deserve being stated for the sake of emphasis. The last is the trick decisive for the application of Theorem 6.1 to the proof of Theorem 2.4.

**Lemma 6.3.** *Let  $\varphi_1, \dots, \varphi_N$  and  $\psi_1, \dots, \psi_M$  be  $n$ -variable DIRE families. Then the concatenation  $\varphi_1, \dots, \varphi_N, \psi_1, \dots, \psi_M$  is an  $n$ -variable DIRE family.*

**Lemma 6.4.** *Let  $\varphi_1, \dots, \varphi_N$  be an  $n$ -variable DIRE family. Let  $\varphi_{N+1}$  be a  $\mathbb{C}$ -linear combination of  $\varphi_1, \dots, \varphi_N$ . Then the extended family  $\varphi_1, \dots, \varphi_{N+1}$  is an  $n$ -variable DIRE family.*

**Lemma 6.5.** *Let  $\varphi_1, \dots, \varphi_N$  be an  $n$ -variable DIRE family. Let  $A$  be an  $n$  by  $m$  matrix with complex entries. Identify  $\mathbb{C}^n$  and  $\mathbb{C}^m$  with spaces of column vectors. Then there exists an  $m$ -variable DIRE family  $\psi_1, \dots, \psi_N$  such that for  $i = 1, \dots, N$  we have  $\psi_i(z) = \varphi_i(Az)$  for all  $z \in \mathbb{C}^m$  sufficiently near the origin.*

**Lemma 6.6.** *Let  $\psi_1, \dots, \psi_n$  be a family of holomorphic functions defined in an open disk centered at the origin in  $\mathbb{C}$ . Let  $\varphi_1, \dots, \varphi_N$  be an  $n_0$ -variable DIRE family, where  $N \geq n \geq n_0$ . Assume that*

$$\psi_i(z) = z\varphi_i(z\psi_1(z), \dots, z\psi_n(z))$$

for  $i = 1, \dots, n$  and  $z \in \mathbb{C}$  sufficiently near the origin. Then  $\psi_1, \dots, \psi_n$  can be extended to a 1-variable DIRE family.

*Proof.* By the preceding lemma we may assume without loss of generality that  $n = n_0$ . For  $i = n+1, \dots, n+N$  the formula

$$\psi_i(z) = \varphi_{i-n}(z\psi_1(z), \dots, z\psi_n(z)) - \varphi_{i-n}(0)$$

defines a holomorphic function  $\psi_i$  in some open neighborhood of the origin in  $\mathbb{C}$ . We will prove that the extended family  $\psi_1, \dots, \psi_{N+n}$  is a 1-variable DIRE family. Note that all the functions  $\psi_i$  vanish at the origin. Let  $\Phi_1, \dots, \Phi_{n+N} \in \mathbb{C}(X_1, \dots, X_{N+n})$  be with respect to  $\varphi_1, \dots, \varphi_N$  as called for by the definition of an  $n$ -variable DIRE family. For  $i = 1, \dots, N+n$  define  $\Psi_i \in \mathbb{C}(X_1, \dots, X_{N+n+1})$  by the formula

$$\begin{aligned} & \Psi_i(X_1, \dots, X_{n+N+1}) \\ = & \begin{cases} X_{i+1} - X_1(X_{i+n+1} + \varphi_i(0)) & \text{if } 1 \leq i \leq n, \\ \Phi_{i-n}(X_1 X_2, \dots, X_1 X_{n+1}, X_{n+2}, \dots, X_{n+N+1}) & \text{if } n+1 \leq i \leq n+N. \end{cases} \end{aligned}$$

Then:

- (I)  $\Psi_i$  is defined at the origin and  $\Psi_i(0) = 0$  for  $i = 1, \dots, N+n$ ,
- (II)  $(\det_{i,j=1}^{N+n} \frac{\partial \Psi_i}{\partial X_{j+1}})(0) \neq 0$ , and
- (III)  $\Psi_i(z, \psi_1(z), \dots, \psi_{N+n}(z)) = 0$  for  $i = 1, \dots, N+n$  and  $z \in \mathbb{C}$  sufficiently near the origin.

In other words,  $\Psi_1, \dots, \Psi_{n+N}$  are with respect to  $\psi_1, \dots, \psi_{n+N}$  as called for by the definition of 1-variable DIRE family.  $\square$

We next formulate a purely algebraic result and explain how to deduce Theorem 6.1 from it.

**Theorem 6.7.** *Let  $n$  and  $N$  be positive integers. Let*

$$F_1, \dots, F_N \in \mathbb{C}[X_1, \dots, X_{N+n}]$$

*be given with the following properties:*

$$(6.4) \quad F_i(0) = 0 \text{ for } i = 1, \dots, N.$$

$$(6.5) \quad \left( \det_{i,j=1}^N \frac{\partial F_i}{\partial X_{j+n}} \right) (0) \neq 0.$$

*Let  $I \subset \mathbb{C}[X_1, \dots, X_{N+n}]$  be the ideal generated by  $F_1, \dots, F_N$ . Then there exist polynomials  $G \in \mathbb{C}[X_1, \dots, X_{n+1}]$  and  $H \in \mathbb{C}[X_1, \dots, X_{n+N}]$  such that  $G \neq 0$ ,  $H(0) \neq 0$  and  $GH \in I$ .*

The proof of Theorem 6.7 will be given in §6.5 after we review in §6.4 the needed material from commutative algebra.

**Remarks.** (i) The important point here is that  $G$  is a polynomial involving only the variables  $X_1, \dots, X_{n+1}$ ; the variables  $X_{n+2}, \dots, X_{n+N}$  are uninvolved. (ii) The proof of the theorem is a routine application of the theory of noetherian local rings, and in particular is purely algebraic. The proof will show that the theorem holds with  $\mathbb{C}$  replaced by any algebraically closed field.

### 6.3 Deduction of Theorem 6.1 from Theorem 6.7

By symmetry it is enough to show that  $\varphi_1$  is algebraic. Let  $\Phi_1, \dots, \Phi_N \in \mathbb{C}(X_1, \dots, X_{n+N})$  be as required to exist by the definition of  $n$ -variable DIRE family with respect to  $\varphi_1, \dots, \varphi_N$ . Write

$$\Phi_i = F_i/D_i \quad (F_i, D_i \in \mathbb{C}[X_1, \dots, X_{n+N}], D_i(0) \neq 0).$$

Without loss of generality we may simply assume that  $D_i = 1$  and hence  $\Phi_i = F_i$ . Then conditions (I,II) are precisely the hypotheses (6.4,6.5) of Theorem 6.7. Let  $G$  and  $H$  be as provided by Theorem 6.7. Then  $H(z_1, \dots, z_n, \varphi_1(z), \dots, \varphi_N(z))$  is non-vanishing for  $z = (z_1, \dots, z_n) \in \mathbb{C}^n$  sufficiently near the origin, hence  $G(z_1, \dots, z_n, \varphi_1(z))$  vanishes for  $z \in \mathbb{C}^n$  sufficiently near the origin, and hence  $\varphi_1$  is indeed algebraic.  $\square$

### 6.4 Review of dimension theory of noetherian local rings

The material reviewed here is developed in detail in the texts [AM69] and [Mat80].

#### (i) The setting

In our review *rings* are always commutative with unit. A ring  $R$  is *noetherian* if every ideal is finitely generated. All rings to be considered below are assumed to be noetherian. A *local ring* is a ring possessing a unique maximal ideal. For the rest of the discussion we fix a noetherian local ring  $R$  with maximal ideal  $M$  and

denote the residue field  $R/M$  by  $k$ . We urge the reader to keep the following key example of a triple  $(R, M, k)$  in mind:

$$(6.6) \quad \begin{cases} R &= \{F \in \mathbb{C}(X_1, \dots, X_d) \mid F \text{ is defined at the origin}\}, \\ M &= \{F \in R \mid F(0) = 0\}, \text{ and} \\ k &= \mathbb{C}. \end{cases}$$

### (ii) Dimension of a noetherian local ring

For each integer  $n$  the quotient  $M^n/M^{n+1}$  is a finite-dimensional vector space over  $k$ . (Here  $M^n$  stands for the ideal generated by all  $n$ -fold products of elements of  $M$ , and by convention  $M^0 = R$ .) Consider the nonnegative-integer-valued function

$$\chi(n) = \sum_{i=0}^{n-1} \dim_k M^i/M^{i+1}$$

of nonnegative integers  $n$ . There exists a unique polynomial  $F(t)$  in a variable  $t$  with rational coefficients such that  $\chi(n) = F(n)$  for all  $n \gg 0$ . (See [Mat80, 12.C] or [AM69, Cor. 11.5].) The polynomial  $F(t)$  is called the *Hilbert-Samuel* polynomial of  $R$ . The degree in  $t$  of  $F(t)$  is by definition the *dimension* of  $R$ , and denoted  $d(R)$ . In example (6.6) we have  $d(R) = d$ .

### (iii) Regular local rings

We say that  $R$  is *regular* if  $d(R) = \dim_k M/M^2$ , in which case necessarily  $\chi(n) = \binom{n+d(R)}{d(R)}$  for all  $n$ , and  $R$  is an integral domain. (See [Mat80, (17.E) Thm. 35 and (17.F) Thm. 36] or [AM69, Thm. 11.22 and Lemma 11.23]. In example (6.6) this can be verified directly, noting that  $\dim M/M^2 = d$  in that case.) Suppose for the rest of this paragraph that  $R$  is regular of dimension  $d$ . Elements  $f_1, \dots, f_d \in M$  forming a basis over  $k$  for the quotient  $M/M^2$  are said to form a *regular system of parameters* for  $R$ . By a standard argument employing Nakayama's lemma (for the latter see [Mat80, (1.M) Lemma] or [AM69, Prop. 2.6]) any regular system of parameters for  $R$  necessarily generates the maximal ideal  $M$ . In example (6.6) the variables  $X_1, \dots, X_d$  form a regular system of parameters.

### (iv) Cutting down regular local rings

Again suppose that  $R$  is regular of dimension  $d$ . Given a regular system of parameters  $f_1, \dots, f_d$  in  $R$ , and also given  $i = 0, \dots, d$ , the ideal  $(f_1, \dots, f_i) \subset R$  generated by  $f_1, \dots, f_i$  is prime and the quotient  $R/(f_1, \dots, f_i)$  is a regular local ring in which the images of  $f_{i+1}, \dots, f_d$  form a regular system of parameters. (See [Mat80, (17.F) Thm. 36].) One should think of this fact as an algebraic version of the implicit function theorem.

**(v) Algebraic independence of regular systems of parameters**

Again assume that  $R$  is regular of dimension  $d$ , and further assume that  $R$  contains a field  $k_0$ . Then every regular system of parameters  $f_1, \dots, f_d$  in  $R$  is algebraically independent over  $k_0$ , i. e., for every polynomial  $F(X_1, \dots, X_d)$  in independent variables  $X_1, \dots, X_d$  with coefficients in  $k_0$ , if  $F(f_1, \dots, f_d) = 0$ , then  $F = 0$ . (See [AM69, Cor. 11.21] or [Mat80, (20.D) App. 1].) In example (6.6) we may take  $k_0 = \mathbb{C}$ .

**(vi) Relation of dimension to transcendence degree**

Assume now that  $(R, M, k)$  is of the form of example (6.6). (More generally we may assume that  $(R, M, k)$  is the evident analogue of example (6.6) with  $\mathbb{C}$  replaced by any algebraically closed field.) Let  $P$  be any prime ideal of  $R$ . The quotient  $R/P$  is again a noetherian local ring (but maybe not regular). (The ring  $R/P$  admits interpretation as the local ring at a point, perhaps singular, of an algebraic variety in  $k^d$ .) Let  $e$  be the *transcendence degree* of  $R/P$  over  $k$ , i. e., the supremum of the set of integers  $m \geq 0$  such that there exist  $m$  elements of  $R/P$  algebraically independent over  $k$ . Then we have  $e = d(R/P)$ . (See [AM69, Thm. 11.25].) One has this equality whether or not  $R/P$  is regular.

**6.5 Proof of Theorem 6.7**

We are ready to move rapidly through the proof. We will flag the relevant paragraphs above at each step. Consider the ring  $R \subset \mathbb{C}(X_1, \dots, X_{n+N})$  consisting of all fractions  $A/B$  where  $A, B \in \mathbb{C}[X_1, \dots, X_{n+N}]$  and  $B(0) \neq 0$ . Then  $R$ , see paragraph (iii), is a regular local ring of dimension  $n + N$ . Hypotheses (6.4, 6.5) imply that  $X_1, \dots, X_n, F_1, \dots, F_N$  form a regular system of parameters for  $R$ . Let  $P$  be the prime ideal of  $R$  generated by  $F_1, \dots, F_N$  and let  $x_1, \dots, x_{n+N}$  be the images in the quotient  $R/P$  of  $X_1, \dots, X_{n+N}$ , respectively. By paragraph (iv), the quotient  $R/P$  is a regular local ring of dimension  $n$  for which  $x_1, \dots, x_n$  forms a regular system of parameters. By paragraph (v), necessarily  $x_1, \dots, x_n$  are algebraically independent over  $\mathbb{C}$  and furthermore, see paragraph (vi), no set of elements of  $R/P$  algebraically independent over  $\mathbb{C}$  can have cardinality exceeding  $n$  (we emphasize that equality to 0 is taken here in  $R/P$ , not  $R$ ). So there exists  $0 \neq G = G(X_1, \dots, X_{n+1}) \in \mathbb{C}[X_1, \dots, X_{n+1}]$  such that  $G(x_1, \dots, x_{n+1}) = 0$  and hence (equivalently)  $G \in \mathbb{C}[X_1, \dots, X_{n+1}] \cap P$ . Now every element of  $P$  can be written  $A/B$  where  $A \in I$  and  $B \in \mathbb{C}[X_1, \dots, X_{n+N}]$  is such that  $B(0) \neq 0$ . In particular we may write  $G = A/B$  in such fashion. Taking  $H = B$ , we have  $GH = A \in I$ , as desired.  $\square$

**Remark** The argument above goes through with  $\mathbb{C}$  replaced by any algebraically closed field, and thus Theorem 6.7 actually holds at that greater level of generality.

## 7 Proof of Theorem 2.4

Let  $\ell$  be a large positive integer. Let  $\{Z_m\}_{m=1}^M$  be an enumeration of all complex-valued functions  $f$  on color space of the form

$$f(c) = \mathbf{1}_I(x)\xi^j/\sqrt{\text{length of } I}; \quad (c = (x, \xi) \in C = [0, 1] \times S^1, \quad I \in \mathcal{I}, \quad j = -\ell, \dots, \ell).$$

Note that  $\{Z_m\}_{m=1}^M$  is an orthonormal system in  $L^2(C)$ . We suppose  $\ell$  is chosen large enough so that we have an expansion

$$s(c, c') = \sum_{i=1}^M \sum_{j=1}^M \rho_{ij} Z_i(c) \bar{Z}_j(c')$$

for some complex constants  $\rho_{ij}$ . Also write

$$1 = \sum_{m=1}^M \rho_m Z_m \quad \left( \rho_m = \int \bar{Z}_m(c) P(dc) \right).$$

With  $\Psi(c, \lambda)$  as defined in the color equations (3.9), put

$$w_m(z) = \int \Psi(c, 1/z) \bar{Z}_m(c) P(dc) \quad (m = 1, \dots, M), \quad w_{M+1}(z) = S(1/z).$$

The functions  $w_m(z)$  are defined and holomorphic for  $|z|$  small and positive, and moreover are  $O(|z|)$ , and hence extend to holomorphic functions in a small disk about the origin which vanish at the origin. Consider the functions

$$F_m(z) = \sum_{j=1}^M \rho_{mj} \int \frac{\bar{Z}_j(c) P(dc)}{1 - \sum_{m=1}^M z_m Z_m(c)} \quad (m = 1, \dots, M),$$

$$F_{M+1}(z) = \sum_{j=1}^M \rho_j \int \frac{\bar{Z}_j(c) P(dc)}{1 - \sum_{m=1}^M z_m Z_m(c)}$$

defined and holomorphic for  $z = (z_1, \dots, z_M) \in \mathbb{C}^M$  sufficiently near the origin. The family  $F_1, \dots, F_{M+1}$  may be extended to a  $M$ -variable DIRE family by Proposition 6.2 along with Lemmas 6.3, 6.4, and 6.5. Now the general color equations can be rewritten in the form

$$(7.1) \quad w_m(z) = z F_m(z w_1(z), \dots, z w_M(z)) \quad \text{for } m = 1, \dots, K+1 \text{ and } |z| \text{ small.}$$

By Lemma 6.6 the family  $w_1, \dots, w_M$  may be extended to a 1-variable DIRE family, and hence each function  $w_m$  is algebraic by Theorem 6.1. Finally,  $S(\lambda) = w_{M+1}(1/\lambda)$  is algebraic.  $\square$

## Appendix

**Acknowledgment.** Ofer Zeitouni was partially supported by NSF grant number DMS-0503775.

### Bibliography

- [Ahl79] Ahlfors, L. *Complex analysis. An introduction to the theory of analytic functions of one complex variable*. Third edition. International Series in Pure and Applied Mathematics, McGraw-Hill Book Co., New York, 1978.
- [AM69] Atiyah, M. F.; Macdonald, I. G., *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.
- [Ao84] Aomoto, K. Spectral theory on a free group and algebraic curves. *J. Fac. Sci. Univ. Tokyo Sect. IA, Math.* **31** (1984), 297–318.
- [AZ06] Anderson, G.; Zeitouni, O. A CLT for a band matrix model. *Probability Theory and Related Fields* **134** (2006), 283–338.
- [Ba99] Bai, Z. D. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica* **9** (1999), 611–677.
- [BDM96] Boutet de Monvel, A.; Khorunzhy, A.; Vasilchuck, V. Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Proc. Rel. Fields* **2** (1996), 607–636.
- [BK99] Boutet de Monvel, A.; Khorunzhy, A. On the norm and eigenvalue distribution of large random matrices. *Annals Probab.* **27** (1999), 913–944.
- [Ca63] Cartan, H. *Elementary theory of analytic functions of one or several complex variables*. Translated from the French. Reprint of the 1973 edition. Dover Publications, Inc., New York, 1995.
- [Ch06] Chatterjee, S. A generalization of the Lindeberg theorem. *Annals Probab.* **34** (2006), 2061–2076.
- [CLO07] Cox, D.; Little, J.; O’Shea, D. *Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra*. Third edition. Undergraduate Texts in Mathematics. Springer, New York, 2007.
- [GB07] Garofalidis, S.; Bellissard, J. Algebraic  $G$ -functions associated to matrices over a group-ring, preprint, [arXiv:0708.4234v3\[math.CO\]](https://arxiv.org/abs/0708.4234v3).
- [Gi90] Girko, V. L. *Theory of random determinants*. Translated from the Russian. Mathematics and its Applications (Soviet Series), **45**. Kluwer Academic Publishers Group, Dordrecht, 1990.
- [Gi01] Girko, V. L. *Theory of stochastic canonical equations*. Vol. I. Mathematics and its Applications, **535**. Kluwer Academic Publishers, Dordrecht, 2001.
- [GoT05] Götze, F.; Tikhomirov, A. N. Limit theorems for spectra of random matrices with martingale structure. *Stein’s method and applications*, 181–193, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., **5**, Singapore Univ. Press, Singapore, 2005.
- [Gu02] Guionnet, A. Large deviation upper bounds and central limit theorems for band matrices. *Ann. Inst. H. Poincaré Probab. Statist.* **38** (2002), 341–384.
- [HLN05] Hachem, W.; Loubaton, P.; Najim, J. The empirical eigenvalue distribution of a Gram matrix: from independence to stationarity. *Markov Proc. Rel. Fields* **11** (2005), 629–648.
- [HP00] Hiai, F.; Petz, D. *The semicircle law, free random variables and entropy*. Mathematical Surveys and Monographs, **77**. American Mathematical Society, Providence, RI, 2000.
- [KKP96] Khorunzhy, A. M.; Khoruzhenko, B. A.; Pastur, L. A. Asymptotic properties of large random matrices with independent entries. *J. Math. Phys.* **37** (1996), 5033–5060.
- [MPK92] Molchanov, S. A.; Pastur, L. A.; Khorunzhii, A. M.; Distribution of the eigenvalues of random band matrices in the limit of their infinite order. *Theoret. and Math. Phys.* **90** (1992), 108–118.
- [PM67] Pastur, L. A.; Martchenko, V. A. The distribution of eigenvalues in certain sets of random matrices. *Math. USSR-Sbornik* **1** (1967), 457–483.
- [Mat80] Matsumura, H. *Commutative algebra*. Second edition. Math. Lecture Note Series **56**. Benjamin/Cummings Publishing Co., Inc., Reading, Mass., 1980.

- [RE06] Rao, N. R.; Edelman, A. The polynomial method for random matrices. Preprint, [Arxiv:math.PR/0601389v1](https://arxiv.org/abs/math/0601389v1).
- [SSB07] Schenker, J. H.; Schulz-Baldes, H. Semicircle law and freeness for random matrices with symmetries or correlations, *Int. Math. Res. Not.* (2007), 2007:rnm047-36.
- [Sh96] Shlyakhtenko, D. Random Gaussian band matrices and freeness with amalgamation. *Int. Math. Res. Notices* **20** (1996), 1013–1025.
- [SiC95] Silverstein, J. W.; Choi, S. Analysis of the limiting spectral distribution of large dimensional random matrices. *J. Mult. Anal.* **54** (1995), 295–306.
- [Sp98] Speicher, R. Combinatorial theory of the free product with amalgamation and operator-valued free probability theory. *Mem. Amer. Math. Soc.* **132** (1998).
- [St99] Stanley, R. P. *Enumerative combinatorics*. Vol. 2. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge Studies in Advanced Mathematics, **62**. Cambridge University Press, Cambridge, 1999.
- [TV04] Tulino, A. M.; Verdú, S. *Random matrix theory and wireless communication*. Now Publishers, Hanover, Massachusetts, 2004.
- [Voi91] Voiculescu, D. Limit laws for random matrices and free products. *Invent. Math* **104** (1991), 201–220.
- [Voi92] Voiculescu, D. Operations on certain noncommutative operator-valued random variables. *Astérisque* **232** (1995), 243–275.
- [Voi00] Voiculescu, D. Lectures on free probability theory. Lectures on probability theory and statistics (Saint-Flour, 1998), *Lecture Notes in Math.* **1738**, Springer (2000), 279–349.
- [Wig55] Wigner, E. P. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. Math.* **62** (1955), 548–564.
- [Wis28] Wishart, J. The generalized product moment distribution in samples from a Normal multivariate population. *Biometrika* **20A** (1928), 32–52.

Received Month 200X.