Summer School

Connections between Probability and Geometric Functional Analysis

Jerusalem, 14-19 June, 2005

DEVIATION INEQUALITIES ON LARGEST EIGENVALUES

M. Ledoux

University of Toulouse, France

In these notes, we survey developments on the asymptotic behavior of the largest eigenvalues of random matrix and random growth models, and describe the corresponding known non-asymptotic exponential bounds. We then discuss some elementary and accessible tools from measure concentration and functional analysis to reach some of these quantitative inequalities at the correct small deviation rate of the fluctuation theorems. Results in this direction are rather fragmentary. For simplicity, we mostly restrict ourselves to Gaussian models.

INTRODUCTION

In the recent years, important developments took place in the analysis of the spectrum of large random matrices and of various random growth models. In particular, universality questions at the edge of the spectrum has been conjectured, and settled, for a number of apparently disconnected examples.

Let $X^N = (X_{ij}^N)_{1 \leq i,j \leq N}$ be a complex Hermitian matrix such that the entries on and above the diagonal are independent complex (real on the diagonal) centered Gaussian random variables with variance σ^2. Denote by $\lambda_1^N, \ldots, \lambda_N^N$ the real eigenvalues of X^N. Under the normalization $\sigma^2 = \frac{1}{4N}$ of the variance, the famous Wigner theorem indicates that the spectral measure $\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^N}$ converges to the semicircle law, supported on $(-1, +1)$. Furthermore, the largest eigenvalue $\lambda_{\max}^N$ converges almost surely to 1, the right-end point of the support of the semicircle law. As one main achievement in the recent developments of random matrix theory, it has been proved in the early nineties by P. Forrester [Fo1] and C. Tracy and H. Widom [T-W1] that the fluctuations of the largest eigenvalue are given by

$$N^{2/3}(\lambda_{\max}^N - 1) \to F$$

where F is the so-called Tracy-Widom distribution. A similar conclusion holds for real Gaussian matrices, and the result has been extended by A. Soshnikov [So1] to classes of real or complex matrices with independent entries under suitable moment assumptions.

In the striking contribution [B-D-J], J. Baik, P. Deift and K. Johansson proved in 1999 that the Tracy-Widom distribution governs the fluctuation of an apparently completely disconnected model, namely the length of the longest increasing subsequence in a random partition. Denote indeed by L_n the length of the longest increasing subsequence in a random permutation chosen uniformly in the symmetric group over n elements. Then, as shown in [B-D-J],

$$\frac{1}{2n^{1/6}} \big(L_n - 2\sqrt{n} \, \big) \to F$$

weakly, with F the Tracy-Widom distribution. (Note that the normalization is given by the third power of the mean order $2\sqrt{n}$, as it would be the case if we replace $\lambda_{\max}^N$ by $N\lambda_{\max}^N$ in the random matrix model.)

Since then, universality of the Tracy-Widom distribution is conjectured for a number of models, and has been settled recently for some specific ones, including corner growth models, last-passage times in directed percolation, exclusion processes, Plancherel measure, random Young tableaux... For example, let $w(i,j)$, $i,j \in \mathbb{N}$, be independent exponential or geometric random variables. For $M \geq N \geq 1$, set

$$W = W(M,N) = \max_{\pi} \sum_{(i,j)\in\pi} w(i,j)$$

where the maximum runs over all up/right paths π in $\mathbb{N}^2$ from $(1,1)$ to (M,N). The random growth function W may be interpreted as a directed last-passage time in percolation. K. Johansson [Joha1] showed that, for every $c \geq 1$, up to some normalization factor,

$$\frac{1}{N^{1/3}}\Big(W\big([cN], N\big) - \omega N\Big) \to F$$

weakly, where again F is the Tracy-Widom distribution (and ω the mean parameter).

These attractive results, and the numerous recent developments around them (cf. the review papers [Baik2], [Joha4], [T-W4]...) emphasize the unusual rate $(\text{mean})^{1/3}$ and the central role of the new type of distribution F in the fluctuations of largest eigenvalues and random growth models. The analysis of these models is actually made possible by a common determinantal point process structure and asymptotics of orthogonal polynomials for which sophisticated tools from combinatorics, complex analysis, integrable systems and probability theory have been developed. This determinantal structure is also the key to the study of the spacings between the eigenvalues, a topic of major interest in the recent developments of random matrix theory which led in particular to striking conjectures in connection with the Riemann zeta function (cf. [De], [Fo2], [Fy], [Kö], [Meh]...).

In these notes, we will be concerned with the simple question of non-asymptotic exponential deviation inequalities at the correct fluctuation rate in some of the preceding limit theorems. We only concentrate on the order of growth and do not discuss the limiting distributions. For example, in the preceding setting of the largest eigenvalue $\lambda_{\max}^N$ of random matrices, we would be interested to find, for fixed $N \geq 1$ and $\varepsilon > 0$, (upper-) estimates on

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1+\varepsilon\}\big) \quad \text{and} \quad \mathbb{P}\big(\{\lambda_{\max}^N \leq 1-\varepsilon\}\big)$$

which fit the weak convergence rate towards the Tracy-Widom distribution. In a sense, this purpose is similar to the Gaussian tail inequalities for sums of independent random variables in the context of the classical central limit theorem. Several results, usually concerned with large and moderate deviation asymptotics and convergence of moments, deal with this question in the literature. However, not all of them are easily accessible, and usually require a rather heavy analysis, connected with stationary phase asymptotics of contour integrals or non-classical analytical schemes of the theory of integrable systems such as Riemann-Hilbert asymptotic methods. In any case, the conclusions so far only deal with rather restricted classes of models. For example, in the random matrix context, only (complex) Gaussian entries allow at this point for satisfactory deviation inequalities at the appropriate rate. Directed percolation models have been answered only for geometric or exponential weights.

The aim of these notes is to provide a few elementary tools, some of them of functional analytic flavour, to reach some of these deviation inequalities. (We will only be concerned with upper bounds.) A first attempt in this direction deals with the modern tools of measure concentration. Measure concentration typically produces Gaussian bounds of the type

$$\mathbb{P}\Big(\Big\{\big|\lambda_{\max}^N - \mathbb{E}(\lambda_{\max}^N)\big| \geq r\Big\}\Big) \leq C\,\mathrm{e}^{-Nr^2/C}, \quad r \geq 0,$$

for some $C > 0$ independent of N. These inequalities are rather robust and hold for large families of distributions. While they describe the correct large deviations, they however do not reflect the small deviations at the rate $(\text{mean})^{1/3}$ of the Tracy-Widom theorem. Further functional tools (if any) would thus be necessary, and such a program was actually advertised by S. Szarek in [Da-S]. We present here a few arguments of possible usefulness to this task, relying on Markov operator ideas such as hypercontractivity and integration by parts. In particular, we try to avoid saddle point analysis on Laplace integrals for orthogonal polynomials which are at the root of the asymptotic results. We however still rely on determinantal and orthogonal polynomial representations of the random matrix models. Certainly, suitable bounds on orthogonal polynomials might supply for most of what is necessary to our purpose. Our first wish was actually to try a few abstract and (hopefully) general arguments to tackle some of these questions in the hope of extending some conclusions to more general models. The various conclusions from this particular viewpoint are however far from complete, not always optimal, and do not really extend to new examples of interest. It is the hope of the future research that new tools may answer in a more satisfactory way some of these questions.

The first chapter describes, in the particular example of the Gaussian Unitary Ensemble, the fundamental determinantal structure of the eigenvalue distribution and the orthogonal polynomial method which allow for the fluctuation and large deviation asymptotics of the top eigenvalues of random matrix and random growth models. In the second chapter, we present the known exponential deviation inequalities which may be drawn from the asymptotic theory and technology. Chapter 3 addresses the measure concentration tools in this setting, and discusses both their usefulness and limitations. In Chapter 4, the tool of hypercontractivity of Markov operators is introduced to the task of deviation and variance inequalities at the Tracy-Widom rate. The last chapter presents some moment recurrence equations which may be obtained from integration by parts for Markov operators, and discusses their interest in deviation inequalities both above and below the limiting expected mean.

These notes are only educational and do not present any new result. They moreover focus on a very particular aspect of random matrix theory, ignoring some main developments and achievements. In particular, references are far from exhaustive. Instead, we try to refer to some general references where more complete expositions and pointers to the literature may be found. We apologize for all the omissions and inaccuracies in this respect. In connection with these notes, let us thus mention, among others, the book [Meh] by M. L. Mehta which is a classical reference on the main random matrix ensembles from the mathematical physics point of view. It contains in particular numerous formulas on the eigenvalue densities, their correlation functions etc. The very recent third edition presents in addition some of the latest developments on the asymptotic behaviors of eigenvalues of random matrices. The monograph [De] by P. Deift discusses orthogonal polynomial ensembles and presents an introduction to the Riemann-Hilbert asymptotic method. P. Forrester [Fo] extensively describes the various mathematical physics models of random matrices and their relations to integrable systems. The survey paper by Z. D. Bai [Bai] offers a complete account on the spectral analysis of large dimensional random matrices for general classes of Wigner matrices by the moment method and the Stieltjes transform. The short reviews [Joha4], [T-W4], [Baik2] provide concise presentations of some main recent achievements. The lectures [Fy] by Y. Fyodorov are an introduction to the statistical properties of eigenvalues of large random Hermitian matrices, and treat in particular the paradigmatic example of the Gaussian Unitary Ensemble (much in the spirit of these notes). Finally, the recent nice and complete survey on orthogonal polynomial ensembles by W. König [Kö] achieves an accessible and inspiring account to some of these important developments. More references may be downloaded from the preceding ones.

Thanks are due to G. Schechtman and T. Szankowski for their invitation to this summer school, and to all the participants for their interest in these lectures.

1. ASYMPTOTIC BEHAVIORS

In this first chapter, we briefly present some basic facts about the asymptotic analysis of the largest eigenvalues of random matrix and random growth models. For simplicity, we restrict ourselves to some specific models (mostly the Hermite and Meixner Ensembles) for which complete descriptions are available. We follow the recent literature on the subject. In the particular example of the Gaussian Unitary Ensemble, we fully examine the basic determinantal structure of the correlation functions and the orthogonal polynomial method. We further discuss Coulomb gas and random growth functions, as well as large deviation asymptotics.

1.1. The largest eigenvalue of the Gaussian Unitary Ensemble

One main example of interest throughout these notes will be the so-called Gaussian Unitary Ensemble (GUE). This example is actually representative of a whole family of models. Consider, for each integer $N \geq 1$, $X = X^N = (X_{ij}^N)_{1 \leq i,j \leq N}$ a $N \times N$ selfadjoint centered Gaussian random matrix with variance σ^2. By this, we mean that X is a $N \times N$ Hermitian matrix such that the entries above the diagonal are independent complex (real on the diagonal) Gaussian random variables with mean zero and variance σ^2 (the real and imaginary parts are independent centered Gaussian variables with variance $\sigma^2/2$). Equivalently, the random matrix X is distributed according to the probability distribution

$$\mathbb{P}(dX) = \frac{1}{Z} \exp\big(- \mathrm{Tr}(X^2)/2\sigma^2\big) dX \tag{1.1}$$

on the space $\mathcal{H}_N \cong \mathbb{R}^{N^2}$ of $N \times N$ Hermitian matrices where

$$dX = \prod_{1 \leq i \leq N} dX_{ii} \prod_{1 \leq i < j \leq N} d\,\mathrm{Re}\,(X_{ij}) d\,\mathrm{Im}\,(X_{ij})$$

is Lebesgue measure on $\mathcal{H}_N$ and $Z = Z_N$ the normalizing constant. This probability measure is invariant under the action of the unitary group on $\mathcal{H}_N$ in the sense that UXU^* has the same law as X for each unitary element U of $\mathcal{H}_N$. The random matrix

X is then said to be element of the Gaussian Unitary Ensemble (GUE) ("ensemble" for probability distribution).

The real case is known as the Gaussian Orthogonal Ensemble (GOE) defined by a real symmetric random matrix $X = X^N = (X_{ij}^N)_{1 \leq i,j \leq N}$ such that the entries X_{ij}^N, $1 \leq i \leq j \leq N$, are independent centered real-valued Gaussian random variables with variance σ^2 ($2\sigma^2$ on the diagonal). Equivalently, the distribution of X on the space $\mathcal{S}_N$ of $N \times N$ symmetric matrices is given by

$$\mathbb{P}(dX) = \frac{1}{Z} \exp\big(-\operatorname{Tr}(X^2)/4\sigma^2 \big) dX \tag{1.2}$$

(where now dX is Lebesgue measure on $\mathcal{S}_N$). This distribution is invariant by the orthogonal group.

For such a symmetric or Hermitian random matrix $X = X^N$, denote by $\lambda_1^N, \ldots, \lambda_N^N$ its (real) eigenvalues.

It is a classical result due to E. Wigner [Wig] that, almost surely,

$$\frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i^N} \to \nu \tag{1.3}$$

in distribution as $\sigma^2 \sim \frac{1}{4N}$, $N \to \infty$, where ν is the semicircle law with density $\frac{2}{\pi}(1 - x^2)^{1/2}$ with respect to Lebesgue measure on $(-1, +1)$. This result has been extended, on the one hand, to large classes of both real (symmetric) and complex (Hermitian) random matrices with non-Gaussian independent (subject to the symmetry condition) entries, called Wigner matrices, under the variance normalization $\sigma^2 = \mathbb{E}(|X_{ij}|^2) \sim \frac{1}{4N}$, $i < j$. The basic techniques include moment methods, to show the convergence of

$$\frac{1}{N} \mathbb{E}\Big(\operatorname{Tr}\big((X^N)^p \big) \Big)$$

to the p-moment ($p \in \mathbb{N}$) of the semicircle law, or the Stieltjes transform (a kind of moment generating function) method. Another point of view on the Stieltjes transform is provided by the free probability calculus ([Vo], [V-D-N], [H-P], [Bi]...). In the particular example of the GUE, simple orthogonal polynomial properties may be used (see below). Actually, all these arguments first establish convergence of the mean spectral measure

$$\mu^N = \mathbb{E}\bigg(\frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i^N} \bigg). \tag{1.4}$$

This convergence has been improved to the almost sure statement (1.3) in [Ar]. We refer to the paper [Bai] by Z. D. Bai for a complete account on spectral distributions of large Wigner matrices not addressed here. On the other hand, Wigner's theorem has been extended to orthogonal or unitary invariant ensembles of the type (1.1) or (1.2) where X^2 is replaced (by the functional calculus) by $v(X)$ for some suitable function

$v : \mathbb{R} \to \mathbb{R}$. The main tool in this case is the Stieltjes transform, and the limiting spectral distribution (or equilibrium measure, cf. [De], [H-P], [S-T]...) then depends on (the potential) v. The quadratic potential is the only one leading to independent entries X_{ij}, $1 \leq i \leq j \leq N$, in the matrix X with law (1.1) or (1.2).

It is also well-known that in the GUE and GOE models, as well as in the more general setting of Wigner matrices (cf. [Bai]), under suitable moment hypotheses, the largest eigenvalue $\lambda_{\max}^N = \max_{1 \leq i \leq N} \lambda_i^N$ converges almost surely, as $\sigma^2 = \frac{1}{4N}$, to the right-end point of the support of the semicircle law, that is 1 in the normalization chosen here. (By symmetry, the smallest eigenvalue converges to -1. The result extends to the k-th extremal eigenvalues for every fixed k.) For the orthogonal and unitary invariant ensembles, the convergence is towards the right-end point of the compact support of the limiting spectral distribution (cf. [De]).

As one of the main recent achievements of the theory of random matrices, it has been shown by P. Forrester [Fo1] (in a mathematical physics language) and C. Tracy and H. Widom [T-W1] that the fluctuations of the largest eigenvalue $\lambda_{\max}^N$ of a GUE random matrix $X = X^N$ with $\sigma^2 = \frac{1}{4N}$ around its expected value 1 takes place at the rate $N^{2/3}$. More precisely,

$$N^{2/3}(\lambda_{\max}^N - 1) \to F_{\mathrm{GUE}} \tag{1.5}$$

weakly where F_{GUE} is the so-called (GUE) Tracy-Widom distribution. Note that the normalization $N^{2/3}$ may be somehow guessed from the Wigner theorem since, for $\varepsilon > 0$ small,

$$\mathrm{Card}\,\{1 \leq i \leq N; \lambda_i^N > 1 - \varepsilon\} \sim N\,\nu\big((1-\varepsilon, 1]\big) \sim N\varepsilon^{3/2}$$

so that for ε of the order of $N^{-2/3}$ the probability $\mathbb{P}(\{\lambda_{\max}^N \leq 1 - \varepsilon\})$ should be stabilized. The new distribution F_{GUE} occurs as a Fredholm determinant

$$F_{\mathrm{GUE}}(s) = \det\Big([\mathrm{Id} - K_{\mathrm{Ai}}]_{\mathrm{L}^2(s,\infty)}\Big), \quad s \in \mathbb{R}, \tag{1.6}$$

of the integral operator associated to the Airy kernel K_{Ai}, as a limit in this regime of the Hermite kernel using Plancherel-Rotach orthogonal polynomial asymptotics (see below). C. Tracy and H. Widom [T-W1] were actually able to provide an alternate description of this new distribution F_{GUE} in terms of some differential equation as

$$F_{\mathrm{GUE}}(s) = \exp\bigg(-\int_{2s}^{\infty}(x - 2s)u(x)^2 dx\bigg), \quad s \in \mathbb{R}, \tag{1.7}$$

where $u(x)$ is the solution of the Painlevé II equation $u'' = 2u^3 + xu$ with the asymptotics $u(x) \sim \frac{1}{2\sqrt{\pi}x^{1/4}}\, \mathrm{e}^{-\frac{2}{3}x^{3/2}}$ as $x \to \infty$. Similar conclusions hold for the Gaussian Orthogonal Ensemble (GOE) with a related limiting distribution F_{GOE} of the Tracy-Widom type [T-W2]. Random matrix theory is also concerned sometimes with quaternionic entries leading to the Gaussian Simplectic Ensemble (GSE), cf. [Meh], [T-W2].

A few caracteristics of the distribution F_{GUE} are known. It is non-centered, with a mean around $-.879$, and its respective behaviors at $\pm\infty$ are given by

$$C^{-1}\,\mathrm{e}^{-Cs^3} \leq F_{\mathrm{GUE}}(-s) \leq C\,\mathrm{e}^{-s^3/C} \tag{1.8}$$

and

$$C^{-1}\,\mathrm{e}^{-Cs^{3/2}} \leq 1 - F_{\mathrm{GUE}}(s) \leq C\,\mathrm{e}^{-s^{3/2}/C} \tag{1.9}$$

for s large and C numerical (cf. e.g. [Au], [John], [L-M-R]...)

As already emphasized in the introduction, the Tracy-Widom distributions actually appeared recently in a number of apparently disconnected problems, from the length of the longest increasing subsequence in a random permutation, to corner growth models, last-passage times in oriented percolation, exclusion processes, Plancherel measure, random Young tableaux etc, cf. [Joha4], [T-W4], [Baik2], [Kö]... The Tracy-Widom distributions are conjectured to be the universal limiting laws for this type of models, with a common rate $(\mathrm{mean})^{1/3}$ (in contrast with the $(\mathrm{mean})^{1/2}$ rate of the classical central limit theorem).

The fluctuation result (1.5) has been extended by A. Soshnikov in the stricking contribution [So1] to Wigner matrices $X = X^N$ with real or complex non-Gaussian independent entries with variance $\sigma^2 = \mathbb{E}(|X_{ij}|^2) = \frac{1}{4N}$ and a Gaussian control of the moments $\mathbb{E}(|X_{ij}|^{2p}) \leq (Cp)^p$, $p \in \mathbb{N}$, $1 \leq i < j \leq N$. In particular, the assumptions cover the case of matrices $X = (X_{ij}/2\sqrt{N})_{1\leq i,j\leq N}$ where the X_{ij}'s, $i \leq j$, are independent symmetric Bernoulli variables. This is one extremely rare case so far for which universality of the Tracy-Widom distributions has been fully justified. Interestingly enough, one important aspect of Soshnikov's remarkable proof is that it is actually deduced from the GUE or GOE cases by a moment approximation argument (and not directly from the initial matrix distribution). In another direction, asymptotics of orthogonal polynomials have been deeply investigated to extend the GUE fluctuations to large classes of unitary invariant ensembles. Depending on the structure of the underlying orthogonal polynomials, the proofs can require rather deep arguments involving the steepest descent/stationary phase method for Riemann-Hilbert problems (cf. [De], [Ku], [Baik2], [B-K-ML-M]...). For a strategy based on $1/n$-expansion in unitary invariant random matrix ensembles avoiding the Riemann-Hilbert analysis, see [P-S]. Further developments are still in progress.

1.2 Determinantal representations

The analysis of the GUE, and more general unitary invariant ensembles, is made possible by the determinantal representation of the eigenvalue distribution as a Coulomb gas and the use of orthogonal polynomials. This determinantal point process representation is the key towards the asymptotics results on eigenvalues of large random matrices, both inside the bulk (spacing between the eigenvalues) and at the edge of the spectrum. We follow below the classical literature on the subject [Meh], [De], [Fo2], [P-L]... to which we refer for further details.

Keeping the GUE example, by unitary invariance of the ensemble (1.1) and the Jacobian change of variables formula, the distribution of the eigenvalues $\lambda_1^N \leq \cdots \leq \lambda_N^N$ of $X = X^N$ on the Weyl chamber $E = \{x \in \mathbb{R}^N; x_1 < \cdots < x_N\}$ may be shown to be given by

$$\frac{1}{Z}\,\Delta_N(x)^2 \prod_{i=1}^{N} d\mu(x_i/\sigma) \tag{1.10}$$

where

$$\Delta_N = \Delta_N(x) = \prod_{1\leq i<j\leq N} (x_j - x_i)$$

is the Vandermonde determinant, $d\mu(x) = \mathrm{e}^{-x^2/2}\frac{dx}{\sqrt{2\pi}}$ the standard normal distribution on $\mathbb{R}$ and $Z = Z_N$ the normalization factor. We actually extend the probability distribution (1.10) to the whole of $\mathbb{R}^N$ by symmetry under permutation of the coordinates, and thus speak, with some abuse, of the joint distribution of the eigenvalues $(\lambda_1^N, \ldots, \lambda_N^N)$ as a random vector in $\mathbb{R}^N$.

It is on the basis of the representation (1.10) that the so-called orthogonal polynomial method may be developed. Denote by P_ℓ, $\ell \in \mathbb{N}$, the normalized Hermite polynomials with respect to μ, which form an orthonormal basis of $\mathrm{L}^2(\mu)$. Since, for each ℓ, P_ℓ is a polynomial function of degree ℓ, elementary manipulations on rows or columns show that the Vandermonde determinant $\Delta_N(x)$ is equal, up to a constant depending on N, to

$$D_N = D_N(x) = \det\big(P_{\ell-1}(x_k)\big)_{1\leq k,\ell\leq N}\,.$$

The following lemma is then a useful tool in the study of the correlation functions. It is a simple consequence of the definition of the determinant together with Fubini's theorem.

Lemma 1.1. *On some measure space $(S, \mathcal{S}, m)$, let φ_i, ψ_j, $i, j = 1, \ldots, N$, be square integrable functions. Then*

$$\int_{S^N} \det\big(\varphi_i(x_j)\big)_{1\leq i,j\leq N} \det\big(\psi_i(x_j)\big)_{1\leq i,j\leq N} \prod_{k=1}^{N} dm(x_k) = N!\det\bigg(\int_S \varphi_i\psi_j dm\bigg)_{1\leq i,j\leq N}.$$

Replacing thus Δ_N by D_N in (1.10), a first consequence of Lemma 1.1 applied to

$\varphi_i = \psi_j = P_{\ell-1}$ and $dm = \mathbf{1}_{(-\infty,t/\sigma]}d\mu$ is that, for every $t \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}\big(\{\lambda^N_{\max} \le t\}\big) &= \frac{1}{Z'}\int_{\mathbb{R}^N} D_N(x)^2 \prod_{k=1}^N dm(x_k) \\ &= \det\Big(\langle P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2(]-\infty,t/\sigma],d\mu)}\Big)_{1\le k,\ell\le N} \\ &= \det\Big(\mathrm{Id} - \langle P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2((t/\sigma,\infty),d\mu)}\Big)_{1\le k,\ell\le N}\end{aligned} \tag{1.11}$$

where $Z' = \int_{\mathbb{R}^N} D_N(x)^2 \prod_{k=1}^N d\mu(x_k)$ and $\langle\cdot\,,\,\cdot\rangle_{\mathrm{L}^2(A,d\mu)}$ is the scalar product in the Hilbert space $\mathrm{L}^2(A,d\mu)$, $A \subset \mathbb{R}$.

On the basis of Lemma 1.1 and the orthogonality properties of the polynomials P_ℓ, the eigenvalue vector $(\lambda^N_1,\dots,\lambda^N_N)$ may be shown to have determinantal correlation functions in terms of the (Hermite) kernel

$$K_N(x,y) = \sum_{\ell=0}^{N-1} P_\ell(x)P_\ell(y), \quad x,y \in \mathbb{R}. \tag{1.12}$$

The following statement provides such a description.

Proposition 1.2. *For any bounded measurable function $f : \mathbb{R} \to \mathbb{R}$,*

$$\mathbb{E}\bigg(\prod_{i=1}^N \big[1 + f(\lambda^N_i)\big]\bigg) = \sum_{r=0}^N \frac{1}{r!}\int_{\mathbb{R}^r}\prod_{i=1}^r f(\sigma x_i)\det\big(K_N(x_i,x_j)\big)_{1\le i,j\le r} d\mu(x_1)\cdots d\mu(x_r).$$

Proof. Starting from the eigenvalue distribution (1.10), we have

$$\mathbb{E}\bigg(\prod_{i=1}^N \big[1 + f(\lambda^N_i)\big]\bigg) = \frac{1}{Z'}\int_{\mathbb{R}^N}\prod_{i=1}^N \big[1 + f(\sigma x_i)\big] D_N(x)^2 d\mu(x_1)\cdots d\mu(x_N)$$

where, as above, $Z' = \int_{\mathbb{R}^N} D_N(x)^2 d\mu(x_1)\cdots d\mu(x_N)$. By Lemma 1.1, $Z' = N!$ while similarly

$$\begin{aligned}\int_{\mathbb{R}^N}\prod_{i=1}^N \big[1 + f(\sigma x_i)\big] D_N(x)^2 d\mu(x_1)\cdots d\mu(x_N) \\ &= N!\det\Big(\langle (1+g)P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2(\mu)}\Big)_{1\le k,\ell\le N} \\ &= N!\det\Big(\mathrm{Id} + \langle P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2(g d\mu)}\Big)_{1\le k,\ell\le N}\end{aligned}$$

where we set $g(x) = f(\sigma x)$, $x \in \mathbb{R}$. Hence,

$$\mathbb{E}\bigg(\prod_{i=1}^{N}\big[1+f(\lambda_i^N)\big]\bigg) = \det\Big(\mathrm{Id} + \langle P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2(gd\mu)}\Big)_{1\leq k,\ell\leq N}.$$

Now, the latter is equal to

$$\sum_{r=0}^{N}\frac{1}{r!}\sum_{\ell_1,\ldots,\ell_r=1}^{N}\det\Big(\langle P_{\ell_i-1}, P_{\ell_j-1}\rangle_{\mathrm{L}^2(gd\mu)}\Big)_{1\leq i,j\leq r},$$

and thus, by Lemma 1.1 again, also to

$$\sum_{r=0}^{N}\frac{1}{r!}\sum_{\ell_1,\ldots,\ell_r=1}^{N}\frac{1}{r!}\int_{\mathbb{R}^r}\det\big(f(\sigma x_j)P_{\ell_i-1}(x_j)\big)_{1\leq i,j\leq r} \\ \times\det\big(P_{\ell_i-1}(x_j)\big)_{1\leq i,j\leq r}d\mu(x_1)\cdots d\mu(x_r).$$

By the Cauchy-Binet formula, this amounts to

$$\sum_{r=0}^{N}\frac{1}{r!}\int_{\mathbb{R}^r}\prod_{i=1}^{r}f(\sigma x_i)\det\big(K_N(x_i,x_j)\big)_{1\leq i,j\leq r}d\mu(x_1)\cdots d\mu(x_r)$$

which is the announced claim. □

What Proposition 1.2 (more precisely its immediate extension to the computation of $\mathbb{E}\big(\prod_{i=1}^{N}[1+f_i(\lambda_i^N)]\big)$ for bounded measurable functions $f_i : \mathbb{R}\to\mathbb{R}$, $i = 1,\ldots,N$) puts forward is the fact that the distribution of the eigenvalues, and its marginals, are completely determined by the kernel K_N of (1.12) In particular, replacing f by εf in Proposition 1.2 and letting $\varepsilon\to 0$, the mean spectral measure μ^N of (1.4) is given, for every bounded measurable function f, by

$$\mathbb{E}\bigg(\frac{1}{N}\sum_{i=1}^{N}f(\lambda_i^N)\bigg) = \int_{\mathbb{R}}f(\sigma x)\,\frac{1}{N}\sum_{\ell=0}^{N-1}P_\ell^2\,d\mu. \tag{1.13}$$

Choosing $f = -\mathbf{1}_{(t,\infty)}$ in Proposition 1.2 shows at the other end that the distribution of the largest eigenvalue $\lambda_{\max}^N$ may be expressed by

$$\begin{aligned}&\mathbb{P}\big(\{\lambda_{\max}^N\leq t\}\big)\\ &\quad=\sum_{r=0}^{N}\frac{(-1)^r}{r!}\int_{(t/\sigma,\infty)^r}\det\big(K_N(x_i,x_j)\big)_{1\leq i,j\leq r}d\mu(x_1)\cdots d\mu(x_r),\quad t\in\mathbb{R}.\end{aligned}\tag{1.14}$$

This identity emphasizes the distribution of the largest eigenvalue $\lambda_{\max}^N$ as the Fredholm determinant of the (finite rank) operator

$$\varphi\mapsto\int_{t/\sigma}^{\infty}\varphi(y)K_N(\cdot,y)d\mu(y)$$

with kernel K_N. That this expression be called a determinant is justified in particular by (1.11) (cf. e.g. [Du-S] or [G-G] for generalities on Fredholm determinants).

A classical formula due to Christoffel and Darboux (cf. [Sze]) indicates that

$$K_N(x,y) = \kappa_N \, \frac{P_N(x)P_{N-1}(y) - P_{N-1}(x)P_N(y)}{x-y}\,, \quad x, y \in \mathbb{R}. \tag{1.15}$$

(Note, see Chapter 3, that $P_N' = \sqrt{N}\,P_{N-1}$.) In the regime given by (1.5), set then $t = 1 + sN^{-2/3}$, while as usual $\sigma^2 = \frac{1}{4N}$. After a change of variables in (1.14),

$$\mathbb{P}\big(\{\lambda_{\max}^N \leq 1 + sN^{-2/3}\}\big) = \sum_{r=0}^{N} \frac{(-1)^r}{r!} \int_{(s,\infty)^r} \det\big(\widetilde{K}_N(x_i,x_j)\big)_{1\leq i,j\leq r}\, dx_1 \cdots dx_r$$

where

$$\begin{aligned}&\widetilde{K}_N(x,y)\\ &= K_N\Big(2\sqrt{N} + 2xN^{-1/6}, 2\sqrt{N} + 2yN^{-1/6}\Big)\sqrt{\frac{2}{\pi}}\,\frac{1}{N^{1/6}}\,\mathrm{e}^{-[\sqrt{N}+xN^{-1/6}]^2}\mathrm{e}^{-[\sqrt{N}+yN^{-1/6}]^2}.\end{aligned}$$

Now, in this regime, the kernel $\widetilde{K}_N(x,y)$ may be shown to converge to the Airy kernel

$$K_{\mathrm{Ai}}(x,y) = \frac{\mathrm{Ai}\,(x)\mathrm{Ai}'\,(y) - \mathrm{Ai}'\,(x)\mathrm{Ai}\,(y)}{x-y}\,, \quad x, y \in \mathbb{R},$$

through the appropriate asymptotics on the Hermite polynomials known as Plancherel-Rotach asymptotics (cf. [Sze], [Fo1]...). Here Ai is the special Airy function solution of $\mathrm{Ai}'' = x\mathrm{Ai}$ with the asymptotics $\mathrm{Ai}(x) \sim \frac{1}{2\sqrt{\pi}x^{1/4}}\,\mathrm{e}^{-\frac{2}{3}x^{3/2}}$ as $x \to \infty$. By further functional arguments, the convergence may be extended at the level of Fredholm determinants to show that, for every $s \in \mathbb{R}$,

$$\begin{aligned}\lim_{N\to\infty} &\mathbb{P}\big(\{\lambda_{\max}^N \leq 1 + sN^{-2/3}\}\big)\\ &= \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} \int_{(s,\infty)^r} \det\big(K_{\mathrm{Ai}}(x_i,x_j)\big)_{1\leq i,j\leq r}\, dx_1 \cdots dx_r\\ &= \det\Big([\mathrm{Id} - K_{\mathrm{Ai}}]_{\mathrm{L}^2(s,\infty)}\Big) = F_{\mathrm{GUE}}(s),\end{aligned}$$

justifying thus (1.5) (cf. [Fo1], [T-W1], [De]).

1.3 Coulomb gas and random growth functions

Probability measures on $\mathbb{R}^N$ of the type (1.10) may be considered in more generality. Given for example a (continuous or discrete) probability measure ρ on $\mathbb{R}^N$, and $\beta > 0$, let

$$dQ(x) = \frac{1}{Z} \left|\Delta_N(x)\right|^\beta d\rho(x) \tag{1.16}$$

where $Z = \int |\Delta_N|^\beta d\rho < \infty$ is the normalization constant. As we have seen, such probability distributions naturally occur as the joint law of the eigenvalues of matrix models. For example, in the GUE case (cf. (1.10)), $\beta = 2$ and ρ is a product Gaussian measure. (In the GOE and GSE cases, $\beta = 1$ and 4 respectively, cf. [Meh].) For more general (orthogonal, unitary or simplectic) ensembles induced by the probability law $Z^{-1}\exp(-\mathrm{Tr}\, v(X))dX$ on matrices, ρ is the product measure of the density $\mathrm{e}^{-v(x)}$.

One general idea is that among reasonable families of distributions ρ, for example product measures of identical factors, the asymptotic behavior of the probability laws (1.16) is governed by the Vandermonde determinant, and thus exhibits common features. It would be of interest to describe a few general facts about these laws. Distributions of the type (1.16) are called Coulomb gas in mathematical physics. The largest eigenvalue of the matrix models thus appears here as the rightmost point or charge $\max_{1\le i\le N} x_i$ under (1.16).

When $d\rho(x) = \prod_{i=1}^N d\mu(x_i)$ for some probability measure μ on $\mathbb{R}$ or $\mathbb{Z}$, and $\beta = 2$, the preceding Coulomb gas distributions may be analyzed through the orthogonal polynomials of the underlying probability measure μ (provided they exists) as in the example of the GUE discussed previously. In particular, the correlation functions admit determinantal representations. In this case, the probability measures (1.16) are thus sometimes called orthogonal polynomial ensembles. Accordingly, the joint law of the eigenvalues of the GUE is called the Hermite (orthogonal polynomial) Ensemble. In what follows, we only consider Coulomb gas of this sort given as orthogonal polynomial ensembles (cf. [De], [Kö]...).

Following the analysis of the GUE, fluctuations of the largest eigenvalue or rightmost charge of orthogonal polynomial ensembles toward the Tracy-Widom distribution may be developed on the basis of the common Airy asymptotics of orthogonal polynomials (at this regime). The principle of proof extends to kernels K_N properly convergent as $N \to \infty$ to the Airy kernel. When the orthogonal polynomials admit suitable integral representations, the asymptotic behaviors may generally be obtained from a saddle point analysis. For example, the ℓ-th Hermite polynomial may be described as

$$P_\ell(x) = \frac{(-1)^\ell}{\sqrt{\ell!}}\, \mathrm{e}^{x^2/2} \frac{d^\ell}{dx^\ell}\big(\mathrm{e}^{-x^2/2}\big),$$

and thus, after a standard Fourier identity,

$$P_\ell(x) = \frac{(-i)^\ell}{\sqrt{\ell!}}\, \mathrm{e}^{x^2/2} \int_{-\infty}^{+\infty} s^\ell \mathrm{e}^{isx - s^2/2} \, \frac{ds}{\sqrt{2\pi}}\, .$$

The asymptotic behavior as $\ell \to \infty$ may then be handled by the so-called saddle point method (or steepest descent, or stationary phase) of asymptotic evaluation of integrals of the form

$$\int_\Gamma \phi(z) \mathrm{e}^{t\psi(z)} dz$$

over a contour Γ in the complex plane as the parameter t is large (cf. [Fy] for a brief introduction). While these asymptotics are available for the classical orthogonal polynomials from suitable representation of their generating series, the study of more general weights can lead to rather delicate investigations. This might require deep arguments involving steepest descent methods of highly non-trivial Riemann-Hilbert analysis as developed by P. Deift and X. Zhou [D-Z]. We refer to the monograph [De] by P. Deift for an introduction to these methods and complete references up to 1999, including the important contribution [D-K-ML-V-Z]. A further introduction is the set of notes [Ku] including more recent developments and references. See also references in [Baik2]. Discrete orthogonal polynomial ensembles are deeply investigated in [B-K-ML-M]. When suitable contour integral representations of the kernels are available, the standard saddle point method is however enough to determine the expected asymptotics (see e.g. [Joha2] for an example of regularized Wigned matrices).

The orthogonal polynomial method cannot be developed however outside the (complex) case $\beta = 2$. Specific arguments have to be found. The real case for example uses Pfaffians and requires non-trivial modifications (cf. [Meh]). In particular, it is possible to relate the asymptotic behavior of the largest eigenvalues of the GOE to the one of the GUE through a generalized two-dimensional kernel, and thus to conclude to similar fluctuation results (cf. [T-W2], [Wid1]). In particular, the limiting GOE Tracy-Widom law takes the form

$$F_{\mathrm{GOE}}(s) = F_{\mathrm{GUE}}(s)^{1/2} \exp\left(-\frac{1}{2} \int_{2s}^{\infty} u(x) dx \right), \quad s \in \mathbb{R}.$$

Coulomb gas associated to the classical orthogonal polynomial ensembles are of particular interest. Among these ensembles, the Laguerre and (discrete) Meixner ensembles play a central role and exhibit some remarkable features. The Laguerre Ensemble represents the joint law of the eigenvalues of Wishart matrices. Let G be a complex $M \times N$, $M \geq N$, random matrix the entries of which are independent complex Gaussian random variables with mean zero and variance σ^2, and set $Y = Y^N = G^* G$. The law of Y defines a unitary invariant probability measure on $\mathcal{H}_N$, and the distribution of the eigenvalues is given by a Coulomb gas (1.16) with $\beta = 2$ and ρ (up to the scaling parameter σ) the product measure of the Gamma distribution $d\mu(x) = \Gamma(\gamma+1)^{-1} x^\gamma \mathrm{e}^{-x} dx$ on $(0, \infty)$ with $\gamma = M - N$. The Laguerre polynomials being the orthogonal polynomials for the Gamma law, the corresponding joint distribution of the eigenvalues is called the Laguerre Ensemble. Real Wishart matrices are defined similarly (with $\beta = 1$), and Wishart matrices with non-Gaussian entries may also be considered. The limiting spectral measure of Wishart matrices, as $\sigma^2 \sim \frac{1}{4N}$, and $M \sim cN$, $c \geq 1$, is described by the so-called Marchenko-Pastur distribution (or free Poisson law) [M-P] (cf. [Bai]).

The Meixner Ensemble is associated to a discrete weight. Let μ be the so-called negative binomial distribution on $\mathbb{N}$ with parameters $0 < q < 1$ and $\gamma > 0$ given by

$$\mu\big(\{x\}\big) = \frac{(\gamma)_x}{x!}\, q^x (1-q)^\gamma, \quad x \in \mathbb{N}, \tag{1.17}$$

where $(\gamma)_x = \gamma(\gamma+1)\cdots(\gamma+x-1)$, $x \geq 1$, $(\gamma)_0 = 1$. If $\gamma = 1$, μ is just the geometric distribution with parameter q. The orthogonal polynomials for μ are called the Meixner polynomials (cf. [Sze], [Ch], [K-S]).

As already mentioned above, asymptotics of the Laguerre and Meixner polynomials may then be used as for the GUE, but with increased technical difficulty, to show that the largest eigenvalue of (properly rescaled) Wishart matrices and the rightmost charge, that is the function $\max_{1 \leq i \leq N} x_i$ ('largest eigenvalue'), of the Meixner orthogonal polynomial Ensemble, fluctuate around their limiting value at the Tracy-Widom regime. This has been established by K. Johansson [Joha1] for the Meixner Ensemble. The Laguerre Ensemble appears as a limit as $q \to 1$, and has been investigated independently by I. Johnstone [John] who also carefully analyzes the real case along the lines of [T-W2]. In [So2], A. Soshnikov extends these conclusions to Wishart matrices with non-Gaussian entries following his previous contribution [So1] for Wigner matrices.

Further classical orthogonal polynomial ensembles may be considered. For example, fluctuations of the Jacobi Ensemble constructed over Jacobi polynomials and associated to Beta matrices are addressed in [Co].

The Laguerre and Meixner Ensembles actually share some specific Markovian type properties which make them play a central role in connection with various probabilistic models. In the remarkable contribution [Joha1], K. Johansson indeed showed that the Meixner orthogonal polynomial Ensemble entails an extremely rich mathematical structure connected with many different interpretations. In particular, its rightmost charge may be interpreted in terms of shape functions and last-passage times. Let $w(i,j)$, $i, j \in \mathbb{N}$, be independent geometric random variables with parameter q, $0 < q < 1$. For $M \geq N \geq 1$, set

$$W = W(M,N) = \max_{\pi} \sum_{(i,j)\in\pi} w(i,j) \tag{1.18}$$

where the maximum runs over all up/right paths π in $\mathbb{N}^2$ from $(1,1)$ to (M,N). An up/right path π from $(1,1)$ to (M,N) is a collection of sites $\{(i_k, j_k)\}_{1 \leq k \leq M+N-1}$ such that $(i_1, j_1) = (1,1)$, $(i_{M+N-1}, j_{M+N-1}) = (M,N)$ and $(i_{k+1}, j_{k+1}) - (i_k, j_k)$ is either $(1,0)$ or $(0,1)$. The random growth function W may be interpreted as a directed last-passage time in percolation. Using the Robinson-Schensted-Knuth correspondence between permutations and Young tableaux (cf. [Fu]), K. Johansson [Joha1] proved that, that for every $t \geq 0$,

$$\mathbb{P}\big(\{W \leq t\}\big) = Q\Big(\big\{\max_{1 \leq i \leq N} x_i \leq t + N - 1\big\}\Big) \tag{1.19}$$

where Q is the Meixner orthogonal polynomial Ensemble with parameters q and $\gamma = M - N + 1$. As described in [Joha1], this model is also closely related to the one-dimensional totally asymmetric exclusion process. It may also be interpreted as a randomly growing Young diagram or a zero-temperature directed polymer in a random environment (cf. also [Kö]).

Provided with this correspondence, the fluctuations of the rightmost charge of the Meixner Ensemble may be translated on the growth function $W(M, N)$. As indeed shown in [Joha1], for every $c \geq 1$, (some multiple of) the random variable

$$\frac{W\big([cN], N\big) - \omega N}{N^{1/3}},$$

where

$$\omega = \frac{(1 + \sqrt{qc}\,)^2}{1 - q} - 1,$$

converges weakly to the Tracy-Widom distribution F_{GUE}.

In the limit as $q \to 1$, the model covers the fluctuation of the largest eigenvalue of Wishart matrices, studied independently in [John]. Namely, if w is a geometric random variable with parameter $0 < q < 1$, as $q \to 1$, $(1 - q)w$ converges in distribution to a exponential random variable with parameter 1. If $W = W(M, N)$ is then understood as a maximum over up/right paths of independent such exponential random variables, the identity (1.19) then translates into

$$\mathbb{P}\big(\{W \leq t\}\big) = Q\Big(\big\{\max_{1 \leq i \leq N} x_i \leq t\big\}\Big), \quad t \geq 0, \tag{1.20}$$

where Q is now the Coulomb gas of the Laguerre Ensemble with parameter $\sigma = 1$. (It should be mentioned that no direct proof of (1.20) is so far available.) This example thus admits the double description as a largest eigenvalue of random matrices and a last-passage time.

The central role of the Meixner model covers further instances of interest. Among them are the Plancherel measure and the length of the longest increasing subsequence in a random permutation. (See [A-D] for a general presentation on the length of the longest increasing subsequence in a random permutation.) It was namely observed by K. Johansson [Joha3] (see also [B-O-O]) that, as $q = \frac{\theta}{N^2}$, $N \to \infty$, $\theta > 0$, the Meixner orthogonal polynomial Ensemble converges to the θ-Poissonization of the Plancherel measure on partitions. Since the Plancherel measure is the push-forward of the uniform distribution on the symmetric group S_n by the Robinson-Schensted-Knuth correspondence which maps a permutation $\sigma \in S_n$ to a pair of standard Young tableaux of the same shape, the length of the first row is equal to the length $L_n(\sigma)$ of the longest increasing subsequence in σ. As a consequence, in this regime,

$$\lim_{N \to \infty} \mathbb{P}\big(\big\{W(N, N) \leq t\big\}\big) = \mathbb{P}\big(\{L_{\mathcal{N}} \leq t\}\big), \quad t \geq 0,$$

where $\mathcal{N}$ is an independent Poisson random variable with parameter $\theta > 0$. The orthogonal polynomial approach may then be used to produce a new proof of the important Baik-Deift-Johansson theorem [B-D-J] on the fluctuations of L_n stating that

$$\frac{L_n - 2\sqrt{n}}{2n^{1/6}} \to F_{\mathrm{GUE}} \tag{1.21}$$

in distribution.

The Markovian properties of the specific geometric and exponential distributions make it thus possible to fully analyze the shape functions W and their asymptotic behaviors. (If the definition of up/righ paths is modified, a few more isolated cases have been studied [Baik1], [Joha1], [Joha3], [Se2], [T-W3]...) It would be a challenging question to establish the same fluctutation results, with the same $(\mathrm{mean})^{1/3}$ rate, for random growth functions $W(M,N)$ (1.18) constructed on more general families of distributions of the $w(i,j)$'s, such as for example Bernoulli variables. While superadditivity arguments show that $W([cN],N)/N$ is convergent almost surely as $N \to \infty$ under rather mild conditions, fluctuations around the (usually unknown) limit are almost completely open so far. Even the variance growth (see Chapter 4) has not yet been determined.

1.4 Large deviation asymptotics

In addition to the preceding fluctuation results for the largest eigenvalues or rightmost charges of orthogonal polynomial ensembles, some further large deviation theorems have been investigated during the past years. The analysis again relies of the determinantal structure of Coulomb gas together with a careful examination of the equilibrium measure from the logarithmic potential point of view [S-T].

For example, translated into the framework of the preceding random growth function $W(M,N)$ defined from geometric random variables, K. Johansson also proved in the contribution [Joha1] a large deviation theorem in the form of

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big(\Big\{W\big([cN],N]\big) \geq N(\omega+\varepsilon)\Big\}\Big) = -J(\varepsilon) \tag{1.22}$$

for each $\varepsilon > 0$, where J is an explicit function such that $J(x) > 0$ if $x > 0$ (see below). The result is actually due to T. Seppäläinen [Se3] in the simple exclusion process interpretation of the model. The large deviation principle on the left of the mean takes place at the speed N^2 and expresses

$$\lim_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\Big(\Big\{W\big([cN],N]\big) \leq N(\omega-\varepsilon)\Big\}\Big) = -I(\varepsilon) \tag{1.23}$$

for each $\varepsilon > 0$, where $I(x) > 0$ for $x > 0$. As we will see it below, the rate functions J and I of (1.22) and (1.23) actually partly reflect the $N^{2/3}$ rate of the fluctuation results.

In contrast with the fluctuation theorems of the preceding section which rely on specific orthogonal polynomial asymptotics, such large deviation principles hold for large classes of (both continuous or discrete) Coulomb gas (1.16)

$$dQ(x) = \frac{1}{Z} \left|\Delta_N(x)\right|^\beta \prod_{i=1}^N d\mu(x_i),$$

for arbitrary $\beta > 0$ and under mild hypotheses on μ (cf. [Joha1], [BA-D-G], [Fe]). They are closely related to the large deviation principles at the level of the spectral measures emphasized by D. Voiculescu [Vo] (as a microstate description) and G. Ben Arous and A. Guionnet [BA-G] (cf. [H-P]) (as a Sanov type theorem). These results examine the large deviation principles for the empirical measures $\frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ at the speed N^2 in the space of probability measures on $\mathbb{R}$. The rate function is minimized at the equilibrium measure, almost sure limit of the empirical measure (the semicircle law for example in case of the Hermite Ensemble).

The corresponding rate function of the large deviation principles for the right-most charges (largest eigenvalues) is then usually deduced from the one for the empirical measures. The speed of convergence is however different on the right and on the left of the mean. In the example of the largest eigenvalue $\lambda^N_{\max}$ of the GUE with $\sigma^2 = \frac{1}{4N}$, it is shown in [BA-D-G] that

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}\big(\{\lambda^N_{\max} \geq 1+\varepsilon\}\big) = -J_{\mathrm{GUE}}(\varepsilon) \tag{1.24}$$

where, for every $\varepsilon > 0$,

$$J_{\mathrm{GUE}}(\varepsilon) = 4\int_0^\varepsilon \sqrt{x(x+2)}\, dx. \tag{1.25}$$

Note that $J_{\mathrm{GUE}}(\varepsilon)$ is of the order of $\varepsilon^{3/2}$ for the small values of ε, in accordance with the Tracy-Widom theorem (1.5). Similarly, on the left of the mean,

$$\lim_{N\to\infty} \frac{1}{N^2} \log \mathbb{P}\big(\{\lambda^N_{\max} \leq 1-\varepsilon\}\big) = -I_{\mathrm{GUE}}(\varepsilon) \tag{1.26}$$

for some function I_{GUE} such that $I_{\mathrm{GUE}}(x) > 0$ for every $x > 0$ (cf. also [Joha1], [Fe]). The speed N^2 partly indicates that the largest eigenvalues tend to accumulate below the right-end point of the support of the spectrum. The Laguerre and Meixner examples will be discussed in the next chapter. It is expected that for large classes of potentials v in the driving measure $d\mu = \mathrm{e}^{-v}dx$ of the Coulomb gas Q, the corresponding rate function J on the right of the mean is such that $J(\varepsilon) \sim \varepsilon^{3/2}$ for small ε. Large deviations for the length of the longest increasing subsequence have been described in [Se1], [D-Ze].

2. KNOWN RESULTS ON NON-ASYMPTOTIC BOUNDS

The purpose of these notes is to describe some non-asymptotic exponential deviation inequalities on the largest eigenvalues or rightmost charges of random matrix and random growth models at the order $(\text{mean})^{1/3}$ of the fluctuation results. It actually turns out that several results are already available in the literature, motivated by convergence of moments in Tracy-Widom type theorems or moderate deviation principles interpolating between fluctuations and large deviations. We thus survey in this chapter some results developed to this aim, which however, as we will see it, usually require a rather heavy analysis and only concern some rather specific models. In particular, Wigner matrices or random growth functions with arbitrary weights do not seem to have been accessed by any method so far. We treat upper deviation inequalities both above and below the mean, and analyze their consequences to variance inequalities.

2.1. Upper tails on the right of the mean

As presented in the first chapter, the Tracy-Widom theorem on the behavior of the largest eigenvalue $\lambda_{\max}^N$ of the GUE with the scaling $\sigma^2 = \frac{1}{4N}$ expresses that

$$\lim_{N\to\infty} \mathbb{P}\big(\{\lambda_{\max}^N \leq 1 + sN^{-2/3}\}\big) = F_{\mathrm{GUE}}(s), \quad s \in \mathbb{R}. \tag{2.1}$$

In addition to this fluctuation result, the largest eigenvalue $\lambda_{\max}^N$ also satisfies the large deviation theorems of Section 1.4. In order to quantify these asymptotic results, one would be interested in finding (upper-) estimates, for each fixed $N \geq 1$ and $\varepsilon > 0$, on $\mathbb{P}(\{\lambda_{\max}^N \geq 1+\varepsilon\})$ and $\mathbb{P}(\{\lambda_{\max}^N \leq 1-\varepsilon\})$. Actually, from the discussion on the speed of convergence in the large deviation asymptotics (1.24) and (1.26) and the behaviors (1.8) and (1.9) of the Tracy-Widom distribution F_{GUE}, we typically expect that for some $C > 0$ and all $N \geq 1$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1+\varepsilon\}\big) \leq C\,\mathrm{e}^{-N\varepsilon^{3/2}/C} \quad \text{and} \quad \mathbb{P}\big(\{\lambda_{\max}^N \leq 1-\varepsilon\}\big) \leq C\,\mathrm{e}^{-N^2\varepsilon^3/C}$$

for $\varepsilon > 0$. The range of interest concerns particularly small $\varepsilon > 0$ to cover the values $\varepsilon = sN^{-2/3}$ in (2.1), justifying the terminology of small deviation inequalities. Bounds of this type may then be used towards convergence of moments and variance bounds, or moderation deviation results. (We do not address the question of lower estimates which does not seem to have been investigated in the literature.)

A first approach to such a project would be to carefully follow the proof of the Tracy-Widom theorem, and to control the various Fredholm determinants by appropriate (finite range) bounds on orthogonal polynomials. This is the route taken by G. Aubrun in [Au] which allowed him to state the following small deviation inequality for the largest eigenvalue $\lambda^N_{\max}$ of the GUE (with $\sigma^2 = \frac{1}{4N}$).

Proposition 2.1. *For some numerical constant $C > 0$, and all $N \geq 1$ and $\varepsilon > 0$,*

$$\mathbb{P}\big(\{\lambda^N_{\max} \geq 1 + \varepsilon\}\big) \leq C\,\mathrm{e}^{-N\varepsilon^{3/2}/C}.$$

As announced, when $\varepsilon = sN^{-2/3}$, the deviation inequality of Proposition 2.1 fits the fluctuation result (2.1). The bound is also in accordance with the tail behavior (1.9) of the Tracy-Widom distribution F_{GUE} at $+\infty$.

This line of reasoning can certainly be pushed similarly for the orthogonal polynomial ensembles for which a Tracy-Widom theorem holds, and for which asymptotics of orthogonal polynomials together with the corresponding bounds are available. This issue is seemingly not clearly addressed in the literature. Results of this type seem to be discussed in particular in [G-T-W]. As already emphasized, this might however require a quite deep analysis, including steepest descent arguments of Riemann-Hilbert type (cf. [De]). An attempt relying on measure concentration and weak convergence to the equilibrium measure is undertaken in [Bl] to yield asymptotic deviation inequalities of the correct order for some families of unitary invariant ensembles.

Another direction to deviation inequalities on the right of the mean may be developed in the context of last-passage times, relying on superaddivity and large deviation asymptotics. Let us consider for example the random growth function $W(M,N)$ of Chapter 1, last-passage time in directed percolation for geometric random variables with parameter $0 < q < 1$. As we have seen it, up to some multiplicative factor,

$$\lim_{N\to\infty} \mathbb{P}\Big(\big\{W\big([cN], N\big) \leq \omega N + sN^{1/3}\big\}\Big) = F_{\mathrm{GUE}}(s), \quad s \in \mathbb{R}. \tag{2.2}$$

where we recall that

$$\omega = \frac{(1+\sqrt{qc}\,)^2}{1-q} - 1.$$

Fix $N \geq 1$ and $c \geq 1$, and set $W = W([cN], N)$. As for the largest eigenvalue $\lambda^N_{\max}$ of the GUE, to quantify (2.2), we may ask for exponential bounds on the probabilities

$$\mathbb{P}\big(\{W \geq N(\omega+\varepsilon)\}\big) \quad \text{and} \quad \mathbb{P}\big(\{W \leq N(\omega-\varepsilon)\}\big)$$

for $\varepsilon > 0$. One may also ask for example for bounds on the variance of W, which are expected to be of the order of $N^{2/3}$ by (2.2).

As observed by K. Johansson in [Joha1], inequalities on $\mathbb{P}(\{W \geq N(\omega+\varepsilon)\})$ may be obtained from the large deviation asymptotics (1.22) together with a superadditivity argument (compare [Se1]). It is indeed immediate to see that $W(M,N)$ is superadditive in the sense that

$$W(M,N) + W\big([M+1,2M],[N+1,2N]\big) \leq W(2M,2N)$$

where $W([M+1,2M],[N+1,2N])$ is understood as the supremum over all up/right paths from $(M+1,N+1)$ to $(2M,2N)$. Since $W([M+1,2M],[N+1,2N])$ is independent with the same distribution as $W(M,N)$, it follows that for every $t \geq 0$,

$$\mathbb{P}\big(\big(\{W(M,N) \geq t\}\big)^2 \leq \mathbb{P}\big(\big(\{W(2M,2N) \geq 2t\}\big).$$

Iterating, for every integer $k \geq 1$,

$$\mathbb{P}\big(\big(\{W(M,N) \geq t\}\big)^{2^k} \leq \mathbb{P}\big(\big(\{W(2^kM,2^kN) \geq 2^kt\}\big).$$

Together with the large deviation property (1.22), as $k \to \infty$, for every fixed $N \geq 1$ and $\varepsilon > 0$,

$$\mathbb{P}\big(\{W \geq N(\omega+\varepsilon)\}\big) \leq \mathrm{e}^{-NJ(\varepsilon)}. \tag{2.3}$$

Now the function $J(\varepsilon)$ is explicitely known. It has however a rather intricate description, based itself on the knowledge of the equilibrium measure of the Meixner Ensemble. Precisely, as shown in [Joha1],

$$J(\varepsilon) = J_{\mathrm{MEIX}}(\varepsilon) = \frac{1}{1-q}\int_1^x (x-y)\Big[\frac{c-q}{y+B} + \frac{1-qc}{y+D}\Big]\frac{dy}{\sqrt{y^2-1}}$$

where

$$x = 1 + \frac{(1-q)\varepsilon}{2\sqrt{qc}}, \quad B = \frac{c+q}{2\sqrt{qc}}, \quad D = \frac{1+qc}{2\sqrt{qc}}.$$

One may nevertheless check that

$$J(\varepsilon) \geq C^{-1}\min(\varepsilon, \varepsilon^{3/2}), \quad \varepsilon > 0,$$

where $C > 0$ only depends on c and q. As a consequence, we may state the following exponential deviation inequality. Note that this conclusion requires both the delicate large deviation theorem (1.22) for the Meixner Coulomb gas together with the deep combinatorial description (1.19) (in order to make use of superaddivity of the growth function $W(M,N)$).

Proposition 2.2. *For some constant $C > 0$ only depending on the parameter $0 < q < 1$ of the underlying geometric distribution and $c \geq 1$, and all $N \geq 1$ and $\varepsilon > 0$,*

$$\mathbb{P}\big(\{W \geq N(\omega+\varepsilon)\}\big) \leq C\,\mathrm{e}^{-N\min(\varepsilon,\varepsilon^{3/2})/C}.$$

Note that in addition to the small deviation inequality at the Tracy-Widom rate, Proposition 2.2 also emphasizes the order $\mathrm{e}^{-N\varepsilon/C}$ for the large values of ε due to the precise knowledge of the rate function J_{MEIX}. We will come back to this observation in the context of the GUE below.

The explicit knowledge of the rate function J_{MEIX} actually allows one to make use of the non-asymptotic inequality (2.3) for several related models. For example as we already saw it, if w is geometric with parameter $0 < q < 1$, then as $q \to 1$, $(1-q)w$ converges in distribution to an exponential random variable with parameter 1. In this limit, (2.3) turns into

$$\mathbb{P}\big(\{W \geq N(\omega + \varepsilon)\}\big) \leq \mathrm{e}^{-NJ_{\mathrm{LAG}}(\varepsilon)} \tag{2.4}$$

where now W is the supremum (1.18) over up/right paths of independent exponential random variables with parameter 1, $\omega = \big(1+\sqrt{c}\big)^2$ and J_{LAG} is the Laguerre rate function

$$J_{\mathrm{LAG}}(\varepsilon) = \int_1^x (x-y)\,\frac{(1+c)y + 2\sqrt{c}}{(y+B)^2}\,\frac{dy}{\sqrt{y^2-1}}$$

with

$$x = 1 + \frac{\varepsilon}{2\sqrt{c}}\,, \quad B = \frac{1+c}{2\sqrt{c}}\,.$$

One may check similarly that $J_{\mathrm{LAG}}(\varepsilon) \geq C^{-1}\min(\varepsilon, \varepsilon^{3/2})$ so that the bound (2.4) thus provides an analogue of Propositions 2.1 and 2.2 for the Laguerre Ensemble for both the interpretation in terms of the last-passage time W or the largest eigenvalue of Wishart matrices.

Now one may further go from Wishart matrices to random matrices from the GUE, and actually recover in this way Proposition 2.1. Namely, recalling the Wishart matrix $Y = Y^N = GG^*$ with variance σ^2, as $M \to \infty$,

$$\sigma^{-1}\sqrt{M}\Big(\frac{Y}{M} - \sigma^2\,\mathrm{Id}\Big) \to X$$

in distribution where X follows the GUE law (with variance σ^2). In particular,

$$\sigma^{-1}\sqrt{M}\Big(\frac{1}{M}\,\lambda^N_{\max}(Y) - \sigma^2\Big) \to \lambda^N_{\max}(X). \tag{2.5}$$

Now, after the scaling $\sigma^2 = \frac{1}{4N}$, (2.4) indicates that for $M = [cN]$, $c \geq 1$,

$$\mathbb{P}\big(\{\lambda^N_{\max}(Y) \geq \tfrac{\omega+\varepsilon}{4}\}\big) \leq \mathrm{e}^{-NJ_{\mathrm{LAG}}(\varepsilon)}$$

for every $N \geq 1$ and $\varepsilon > 0$. Change then ε into $2\sqrt{c}\,\varepsilon$ and take the limit (2.5) as $c \to \infty$. Denoting as usual by $\lambda^N_{\max}$ the largest eigenvalue of the GUE with variance $\sigma^2 = \frac{1}{4N}$, it follows that

$$\mathbb{P}\big(\{\lambda^N_{\max} \geq 1 + \varepsilon\}\big) \leq \mathrm{e}^{-NJ_{\mathrm{GUE}}(\varepsilon)} \tag{2.6}$$

where $J_{\rm GUE}(\varepsilon)$, $\varepsilon > 0$, was given in (1.25) as

$$J_{\rm GUE}(\varepsilon) = 4 \int_0^\varepsilon \sqrt{x(x+2)}\, dx.$$

Since $J_{\rm GUE}(\varepsilon) \geq C^{-1} \max(\varepsilon^2, \varepsilon^{3/2})$, $\varepsilon > 0$, we thus recover in this way Proposition 2.1, and actually more precisely

$$\mathbb{P}\big(\{\lambda^N_{\max} \geq 1 + \varepsilon\}\big) \leq C\, {\rm e}^{-N \max(\varepsilon^2, \varepsilon^{3/2})/C} \tag{2.7}$$

for every $\varepsilon > 0$. As in Proposition 2.2, this exponential deviation inequality emphasizes both the small deviations of order $\varepsilon^{3/2}$ in accordance with the Tracy-Widom theorem and the large deviations of the order ε^2.

It may actually be shown directly that (2.6) follows from the large deviation principle (1.24) as a consequence of superadditivity, however on some related representation. It has been proved namely, via various arguments, that the largest eigenvalue $\lambda^N_{\max}$ from the GUE with $\sigma^2 = 1$ has the same distribution as

$$\sup \sum_{i=1}^N \big(B^i_{t_i} - B^i_{t_{i-1}}\big) \tag{2.8}$$

where $B^1, \ldots, B^N$ are independent standard Brownian motions and the supremum runs over all $0 = t_0 \leq t_1 \leq \cdots \leq t_{N-1} \leq 1$. Proofs in [Bar], [G-T-W] are based on the Robinson-Schensted-Knuth correspondence, while in [OC-Y] advantage is taken from non-colliding Brownian motions and queuing theory, and include generalizations related to the classical Pitman theorem (see [OC]). The representation (2.8) may be thought of as a kind of continuous version of directed last-passage percolation for Brownian paths. On the basis of this identification, it is not difficult to adapt the superadditivity argument developed for the random growth function (1.18) to deduce (2.6) from the large deviation bound (1.24). In any case, the price to pay to reach Propositions 2.1 and 2.2 is rather expensive.

It should be pointed out that outside these specific models, non-asymptotic small deviation inequalities at the Tracy-Widom rate are so far open. Universality conjectures would expect similar deviation inequalities for general Wigner matrices or directed last passage times W with general independent weights. Soshnikov's proof [So1] only allows for asymptotic inequalities (cf. Section 5.2). Similarly, only the choice of geometric and exponential random variables w_{ij} gives rise so far to statements such as Proposition 2.2.

The central role of the Meixner model shows, by appropriate scalings and the explicit expression of the rate function $J_{\rm MEIX}$, that the tail (2.3) actually covers further instances of interest. As discussed in Section 1.3, one such instance is the length of the longest increasing subsequence in a random permutation and the Baik-Deift-Johansson theorem (1.21). Namely, in the regime $q = \frac{\theta}{N^2}$, $N \to \infty$, the deviation inequality (2.3) may indeed be used to show that, for every $n \geq 1$ and every $\varepsilon > 0$,

$$\mathbb{P}\big(\big\{L_n \geq 2\sqrt{n}\,(1+\varepsilon)\big\}\big) \leq C \exp\bigg(-\frac{1}{C}\sqrt{n}\,\min\big(\varepsilon^{3/2},\varepsilon\big)\bigg) \tag{2.9}$$

where $C > 0$ is numerical, in accordance thus with (1.21) and the large deviation theorem of [D-Z]. The previous bound also matches the upper tail moderate deviation theorem of [L-M].

2.2 Upper tails on the left of the mean

We next turn to the probability that the largest eigenvalue or rightmost charge is less than or equal to the right-end point of the spectral measure. As already mentioned, the intuition, together with the large deviation asymptotics (1.23) and (1.26), suggests that it is much smaller than the probability that the largest eigenvalue exceeds the right-end point. Let us consider again the Meixner model in terms of the directed last-passage time function W of (1.18) with geometric random variables. We thus look for the probability that $W = W([cN], N)$ is less than or equal to $N(\omega - \varepsilon)$ for each $\varepsilon > 0$ and fixed $N \geq 1$. Things are here much more delicate. Seemingly, only a few results are available, relying furthermore on delicate and quite difficult to access methods and arguments. The following result has been put forward in [B-D-ML-M-Z] by refined Riemann-Hilbert steepest descent methods in order to investigate convergence of moments and moderate deviations. Some related estimates are developed in [B-D-R] in the context of random Young tableaux, and in [L-M-R] for the length of the longest increasing subsequence.

Proposition 2.3. *For some constant $C > 0$ only depending on the parameter $0 < q < 1$ of the underlying geometric distribution and $c \geq 1$, and all $N \geq 1$ and $0 < \varepsilon \leq \omega$,*

$$\mathbb{P}\big(\{W \leq N(\omega - \varepsilon)\}\big) \leq C\,\mathrm{e}^{-N^2\varepsilon^3/C}.$$

(Actually, the statement in [B-D-ML-M-Z] seems to concern only large values of N.)

As for Proposition 2.1, the preceding inequality matches the behavior at $-\infty$ of the Tracy-Widom distribution F_{GUE} given by (1.8).

After [B-D-R] and [B-D-ML-M-Z], H. Widom [Wid2] noticed a somewhat less precise estimate, replacing $N^2\varepsilon^3$ by its square root, using a more simple trace bound, however still requiring steepest descent. We will come back to this observation in Chapter 5.

It is plausible that the behavior of the constant C in Proposition 2.3 allows for limits to the Laguerre and Hermite Ensembles as in the preceding section. This is however not completely obvious from the analysis in [B-D-ML-M-Z]. On the other hand, there is no doubt that a similar Riemann-Hilbert analysis may be performed anagously for these examples, and that the statements corresponding to Proposition 2.3 hold true. We may for example guess the following for the largest eigenvalue $\lambda^N_{\max}$ of the GUE with $\sigma^2 = \frac{1}{4N}$.

Proposition 2.4. *For some numerical constant $C > 0$, and all $N \geq 1$ and $0 < \varepsilon \leq 1$,*

$$\mathbb{P}\big(\{\lambda^N_{\max} \leq 1 - \varepsilon\}\big) \leq C\,\mathrm{e}^{-N^2\varepsilon^3/C}.$$

As already mentioned, similar estimates have been obtained in [L-M-R] in the proof of the lower tail moderate deviations for longest increasing subsequences, where, based on the investigation [B-D-J], the following speed of convergence is established: there exists a numerical constant $C > 0$ such that for every $n \geq 1$ and every $0 < \varepsilon \leq 1$,

$$\mathbb{P}\big(\big\{L_n \leq 2\sqrt{n}\,(1-\varepsilon)\big\}\big) \leq C\,\mathrm{e}^{-n\,\varepsilon^3/C}. \tag{2.10}$$

2.3 Variance inequalities

The non-asymptotic deviation inequalities of Sections 2.1 and 2.2 allow for convergence of moments towards the Tracy-Widom distribution [B-D-ML-M-Z], [Wid2]. In particular, they may easily be combined to reach variance bounds. For example, the next statement on the growth function $W = W([cN], N)$ follows from Propositions 2.2 and 2.3.

Corollary 2.5. *For some constant $C > 0$ (only depending on q and c), and every $N \geq 1$,*

$$\mathrm{var}\,(W) = \mathbb{E}\Big(\big[W - \mathbb{E}(W)\big]^2\Big) \leq CN^{2/3}.$$

Proof. Fix $N \geq 1$. We may write

$$\begin{aligned} N^{-2}\,\mathrm{var}\,(W) &\leq N^{-2}\,\mathbb{E}\big([W-\omega]^2\big) \\ &\leq \int_0^\infty \mathbb{P}\big(\big\{W \geq N(\omega + t)\big\}\big)dt^2 \\ &\quad + \int_0^\omega \mathbb{P}\big(\big\{W \leq N(\omega - t)\big\}\big)dt^2. \end{aligned}$$

By Proposition 2.2,

$$\int_0^\infty \mathbb{P}\big(\big\{W \geq N(\omega+t)\big\}\big)dt^2 \leq C\int_0^\infty \mathrm{e}^{-N\min(t,t^{3/2})/C}dt^2 \leq CN^{-4/3}$$

where C, here and below, may vary from line to line. On the other hand, by Proposition 2.3,

$$\int_0^\omega \mathbb{P}\big(\{W \le N(\omega - t)\}\big) d(t^2) \le C \int_0^\omega \mathrm{e}^{-N^2 t^3/C} dt^2 \le CN^{-4/3}.$$

The proposition is established. □

It is worthwhile mentioning that a weaker bound in Proposition 2.3, with $N^2\varepsilon^3$ replaced by $N\varepsilon^{3/2}$ as proved in [Wid2], is sufficient for the proof of Corollary 2.5.

Taking Proposition 2.4 for granted, we get similarly for the largest eigenvalue $\lambda_{\max}^N$ of the GUE with $\sigma^2 = \frac{1}{4N}$ the following variance bound.

Corollary 2.6. *For some numerical constant $C > 0$, and all $N \ge 1$,*

$$\mathrm{var}\,(\lambda_{\max}^N) \le CN^{-4/3}.$$

As the proof of Corollary 2.5 shows, we actually have that for some $C > 0$ and all $N \ge 1$,

$$\mathbb{E}\big(|\lambda_{\max}^N - 1|^2\big) \le CN^{-4/3}.$$

The same arguments furthermore leads to

$$\sup_N \mathbb{E}\Big(\big|N^{2/3}(\lambda_{\max}^N - 1)\big|^p\Big) < \infty$$

for any $p > 0$ which allow for convergence of moments in the Tracy-Widom theorem. In particular,

$$\big|\mathbb{E}(\lambda_{\max}^N) - 1\big| \le CN^{-2/3}. \tag{2.11}$$

We will observe in Chapter 5 below that moment recurrence equations may be used to see that actually

$$\mathbb{E}(\lambda_{\max}^N) \le 1 - \frac{1}{CN^{2/3}} \tag{2.12}$$

for some $C > 0$ and all $N \ge 1$. In particular thus, (2.11) means that

$$1 - \frac{1}{C'N^{2/3}} \le \mathbb{E}(\lambda_{\max}^N) \le 1 - \frac{1}{CN^{2/3}}$$

for some $C, C' > 0$ and all $N \ge 1$.

3. CONCENTRATION INEQUALITIES

In this chapter, we present a few classical measure concentration tools that may be used in the investigation of exponential deviation inequalities on the largest eigenvalues and growth functions. These tools are of general interest, apply in rather large settings and provide useful informations at the level of large deviation bounds for both extremal eigenvalues and spectral distributions. While measure concentration yields the appropriate (Gaussian) large deviation bounds, it however does not produce the correct small deviation rate $(\mathrm{mean})^{1/3}$ of the asymptotic theorems presented in Chapter 1. For simplicity, we mostly detail below the relevant inequalities in the Gaussian case. We successively present concentration inequalities for largest eigenvalues and random growth functions, spectral measures, as well as Coulomb gas.

3.1 Concentration inequalities for largest eigenvalues

Let μ denote the standard Gaussian measure on $\mathbb{R}^n$ with density $(2\pi)^{-n/2}\mathrm{e}^{-|x|^2/2}$ with respect to Lebesgue measure. One basic concentration property (cf. [Le1]) indicates that for every Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}$ with $\|F\|_{\mathrm{Lip}} \leq 1$, and every $r \geq 0$,

$$\mu\big(\big\{F \geq \textstyle\int F d\mu + r\big\}\big) \leq \mathrm{e}^{-r^2/2}. \tag{3.1}$$

Together with the same inequality for $-F$, for every $r \geq 0$,

$$\mu\big(\big\{|F - \textstyle\int F d\mu| \geq r\big\}\big) \leq 2\,\mathrm{e}^{-r^2/2}. \tag{3.2}$$

The same inequalities hold for a median of F instead of the mean. Independence upon the dimension of the underlying state space is one crucial aspect of these properties.

We may apply for example these inequalities to the largest eigenvalue $\lambda^N_{\max}$ of the GUE. Namely, by the variational characterization,

$$\lambda^N_{\max} = \sup_{|u|=1} uX^N u^*, \tag{3.3}$$

so that $\lambda^N_{\max}$ is easily seen to be a 1-Lipschitz map of the N^2 independent real and imaginary entries X_{ii}, $1 \le i \le N$, $\mathrm{Re}\,(X_{ij})/\sqrt{2}$, $\mathrm{Im}\,(X_{ij})/\sqrt{2}$, $1 \le i < j \le N$, of X^N. Together with the scaling of the variance $\sigma^2 = \frac{1}{4N}$ we thus get the following concentration inequality on $\lambda^N_{\max}$.

Proposition 3.1 *For all $N \ge 1$ and $r \ge 0$,*

$$\mathbb{P}\Big(\Big\{\big|\lambda^N_{\max} - \mathbb{E}(\lambda^N_{\max})\big| \ge r\Big\}\Big) \le 2\,\mathrm{e}^{-2Nr^2}, \quad r \ge 0.$$

As a consequence, note that $\mathrm{var}\,(\lambda^N_{\max}) \le C\,N^{-1}$ that should be compared with Corollary 2.6. Actually, while Proposition 3.1 describes the Gaussian decay of $\lambda^N_{\max}$ for the large values of r, it does not catch the $r^{3/2}$ rate of the small deviation inequality (2.7). It actually seems that viewing the largest eigenvalue as one particular example of Lipschitz function of the entries of the matrix does not reflect enough the structure of the model. This comment more or less applies to all the results of this chapter deduced from the concentration principle.

A similar inequality holds for the GOE, and actually for more general families of Gaussian matrices. Before however going on with further applications of the general principle of measure concentration, a few words are necessary at the level of the centerings. The inequalities emphasized in the preceding chapter indeed discuss exponential deviation inequalities from the limiting expected value (for example 1 for the scaled largest eigenvalue $\lambda^N_{\max}$ of the GUE) while the concentration principle typically produces tail inequalities around some mean (or median) value of the given functional (such as $\mathbb{E}(\lambda^N_{\max})$). A comparison thus requires proper control over $\mathbb{E}(\lambda^N_{\max})$ or similar average values. In the example of the GUE, (2.11) is of course enough to this task, but to make the concentration inequalities relevant by themselves, one needs independent estimates. A few remarks in this regard may be developed.

Keep again the GUE example. We may ask whether $\mathbb{E}(\lambda^N_{\max})$, or a median of $\lambda^N_{\max}$, are smaller than 1, or at least suitably controlled. As emphasized in [Da-S], Gaussian comparison principles are of some help to this task. Consider the real-valued Gaussian process

$$G_u = uX^Nu^* = \sum_{i,j=1}^N X_{ij}u_i\overline{u_j}, \quad |u| = 1,$$

where $u = (u_1, \ldots, u_N) \in \mathbb{C}^N$. It is immediate to check that for every $u, v \in \mathbb{C}^N$,

$$\mathbb{E}\big(|G_u - G_v|^2\big) = \sigma^2 \sum_{i,j=1}^N |u_i\overline{u_j} - v_i\overline{v_j}|^2.$$

Hence, if we define the Gaussian process indexed by $u \in \mathbb{C}^N$, $|u| = 1$,

$$H_u = \sum_{i=1}^N g_i\,\mathrm{Re}\,(u_i) + \sum_{j=1}^N h_j\,\mathrm{Im}\,(u_j)$$

where $g_1, \dots, g_N$, $h_1, \dots, h_N$ are independent standard Gaussian variables, then, for every u, v such that $|u| = |v| = 1$,

$$\mathbb{E}\big(|G_u - G_v|^2\big) \le 2\sigma^2 \mathbb{E}\big(|H_u - H_v|^2\big).$$

By the Slepian-Fernique lemma (cf. [L-T]),

$$\mathbb{E}\big(\sup_{|u|=1} G_u\big) \le \sqrt{2}\,\sigma\,\mathbb{E}\big(\sup_{|u|=1} H_u\big) \le 2\sqrt{2}\,\sigma\,\mathbb{E}\bigg(\bigg[\sum_{i=1}^N g_i^2\bigg]^{1/2}\bigg).$$

When $\sigma^2 = \frac{1}{4N}$, we thus get that

$$\mathbb{E}\big(\lambda_{\max}^N\big) \le \sqrt{2}. \tag{3.4}$$

Together with the one-sided version of the inequality of Proposition 3.1, for every $r \ge 0$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \ge \sqrt{2} + r\}\big) \le \mathrm{e}^{-2Nr^2}$$

that thus agrees with (2.7) for r large.

It is worthwhile mentioning that in the real GOE case, the comparison theorem may be sharpened into

$$\mathbb{E}\big(\lambda_{\max}^N\big) \le 4\sigma^2 \mathbb{E}\bigg(\bigg[\sum_{i=1}^N g_i^2\bigg]^{1/2}\bigg) < 1 \tag{3.5}$$

(cf. [Da-S]). In particular therefore

$$\mathbb{P}\big(\{\lambda_{\max}^N \ge 1 + r\}\big) \le \mathrm{e}^{-Nr^2}$$

for every $r \ge 0$, which is more directly comparable to the Tracy-Widom theorem. However (3.5) is not sharp enough to reach (2.12).

Bounds such as (3.4) or (3.5) extend to the class of sub-Gaussian distributions (cf. [L-T], [Ta3]) including thus random matrices with symmetric Bernoulli entries. They may then be combined as above with Proposition 3.3 below.

On the basis of the supremum representation (3.3) of the largest eigenlue or the very definition (1.18) of last passage time in oriented percolation, one may actually wonder whether bounds on the supremum of Gaussian or more general processes $(Z_t)_{t \in T}$ may be useful in this type of investigation. Numerous developements took place in the last decades (cf. [L-T] and the recent monograph [Ta3]) in the analysis of bounds on $\mathbb{E}(\sup_{t \in T} Z_t)$ and $\mathbb{P}(\{\sup_{t \in T} Z_t \ge r\})$, $r \ge 0$, with rather sophisticated chaining arguments involving metric entropy or majorizing measures. For real symmetric matrices X, the task would be for example to investigate processes given by

$$Z_u = uX^N u^* = \sum_{i,j=1}^N X_{ij} u_i u_j, \quad |u| = 1,$$

where $u = (u_1, \ldots, u_N) \in \mathbb{R}^N$ and X_{ij}, $1 \leq i \leq j \leq N$, are independent centered either Gaussian or Bernoulli variables, and to study the size of the unit sphere $|u| = 1$ under the L^2-metric

$$\mathbb{E}\big(|Z_u - Z_v|^2\big) = \sum_{i,j=1}^{N} |u_i u_j - v_i v_j|^2, \quad |u| = |v| = 1.$$

While these tools provide general, typically Gaussian, bounds, the unusual and more refined rates from random matrix theory do not seem to have been accessed so far from this point of view. It might be a worthwhile project to investigate this question in more detail.

We now come back to the application of measure concentration to general families of random matrices and random growth functions. This is actually the main interest in the theory. The concentration inequality of Proposition 3.1 indeed applies to large families of both real and complex random matrices, the entries of which form a random vector with a dimension free concentration property. For notational simplicity, we only deal below with real matrices but up to numerical factors, all the results hold similarly in the complex case. That is, we are looking for measures μ on $\mathbb{R}^n$, representing the joint law of the entries of a given matrix, which satisfy, as (3.1) or (3.2) for Gaussian measures, the dimension free concentration inequality

$$\mu\big(\big\{|F - \textstyle\int F d\mu| \geq r\big\}\big) \leq C\, \mathrm{e}^{-r^2/C}, \quad r \geq 0 \tag{3.6}$$

for some $C > 0$ independent of n and every 1-Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}$. The mean may be replaced by a median of F. Actually, other tails than Gaussian may be considered, and we refer to [Le1] for a general account on the concentration of measure phenomenon and examples satisfying it. Now, it is immediate (cf. e.g. (3.3)) that the singular values (resp. eigenvalues) of a $N \times N$ matrix X (resp. symmetric matrix) are Lipschitz functions of the vector of the N^2 (resp. $N(N+1)/2$) entries of X. One thus immediately concludes to concentration inequalities of the type of Proposition 3.1 for singular values or eigenvalues of matrices the joint law of the entries satisfying a concentration inequality (3.6). This observation already yields various concentration inequalities for singular values and eigenvalues of families of Gaussian matrices. Another simple example of interest consists of matrices with independent uniform entries (which may be realized as a contraction of Gaussian variables). The following proposition summarizes this conclusion. Note that if $X = (X_{ij})_{1 \leq i,j \leq N}$ is a real symmetric $N \times N$ random matrix, then its eigenvalues are 1-Lipschitz functions of the entries X_{ii}, $1 \leq i \leq N$, $\sqrt{2}\, X_{ij}$, $1 \leq i < j \leq N$ (justifying in particular the normalization of the variances in the GOE). For simplicity, we do not distinguish below between the diagonal and non-diagonal entries, and simply use that the eigenvalues are Lipschitz with a Lipschitz coefficient less than or equal to $\sqrt{2}$ with respect to the vector X_{ij}, $1 \leq i \leq j \leq N$.

Proposition 3.2. *Let $X = (X_{ij})_{1 \le i,j \le N}$ be a real symmetric $N \times N$ random matrix and $Y = (Y_{ij})_{1 \le i,j \le N}$ be a real $N \times N$ random matrix. Assume that the distributions of the random vectors X_{ij}, $1 \le i \le j \le N$, and Y_{ij}, $1 \le i,j \le N$, in respectively $\mathbb{R}^{N(N+1)/2}$ and $\mathbb{R}^{N^2}$ satisfy the dimension free concentration property (3.6). Then, if τ is any eigenvalue of X, respectively singular value of Y, for every $r \ge 0$,*

$$\mathbb{P}\big(\big\{|\tau - \mathbb{E}(\tau)| \ge r\big\}\big) \le C\,\mathrm{e}^{-r^2/2C}, \quad \text{resp.} \quad C\,\mathrm{e}^{-r^2/C}.$$

We next discuss two examples of distributions satisfying concentration inequalities of the type (3.6) and illustrate there application to matrix models.

A first class of interest consists of measures satisfying a logarithmic Sobolev inequality which form a natural extension of the Gaussian example. A probability measure μ on $\mathbb{R}$ or $\mathbb{R}^n$ is said to satisfy a logarithmic Sobolev inequality if for some constant $C > 0$

$$\int_{\mathbb{R}^n} f^2 \log f^2 d\mu \le 2C \int_{\mathbb{R}^n} |\nabla f|^2 d\mu \tag{3.7}$$

for every smooth enough function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int f^2 d\mu = 1$. The prototype example is the standard Gaussian measure on $\mathbb{R}^n$ which satifies (3.7) with $C = 1$. Another example consists of probability measures on $\mathbb{R}^n$ of the type $d\mu(x) = \mathrm{e}^{-V(x)}dx$ where $V - c\,\frac{|x|^2}{2}$ is convex for some $c > 0$ which satisfy (3.7) for $C = \frac{1}{c}$. An important aspect of the logarithmic Sobolev inequality is its stability by product that yields dimension free constants. That is, if $\mu_1, \ldots, \mu_n$ are probability measures on $\mathbb{R}$ satisfying the logarithmic Sobolev inequality (3.7) with the same constant C, then the product measure $\mu_1 \otimes \cdots \otimes \mu_n$ also satisfies it (on $\mathbb{R}^n$) with the same constant. The application of logarithmic Sobolev inequalities to measure concentration is developed by the so-called Herbst argument that indicates that if μ satisfies (3.7), then for any 1-Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}$ and any $\lambda \in \mathbb{R}$,

$$\int \mathrm{e}^{\lambda F} d\mu \le \mathrm{e}^{\lambda \int F d\mu + C\lambda^2/2}.$$

In particular, by a simple use of Markov's exponential inequality (for both F and $-F$), for any $r \ge 0$,

$$\mu\big(\big\{|F - \textstyle\int F d\mu| \ge r\big\}\big) \le 2\,\mathrm{e}^{-r^2/2C},$$

so that the dimension free concentration property (3.6) holds. We refer to [Le1] for a complete discussion on logarithmic Sobolev inequalities and measure concentration. Related Poincaré inequalities, in connection with variance bounds and exponential concentration, may be considered similarly in this context and in the applications below.

As a consequence of this discussion, if $X = (X_{ij})_{1 \le i,j \le N}$ is a real symmetric $N \times N$ random matrix and $Y = (Y_{ij})_{1 \le i,j \le N}$ a real $N \times N$ random matrix such that the entries X_{ij}, $1 \leq i \leq j \leq N$ and Y_{ij}, $1 \leq i,j \leq N$ define random vectors in respectively $\mathbb{R}^{N(N+1)/2}$ and $\mathbb{R}^{N^2}$ the law of which satisfy the logarithmic Sobolev inequality (3.7), then the conclusion of Proposition 3.2 holds. By the product property of logarithmic Sobolev inequalities, this is in particular the case if the variables X_{ij} and Y_{ij} are independent and satisfy (3.7) with a common constant C (for example, they have a common distribution $\mathrm{e}^{-v}dx$ where $v'' \geq c = \frac{1}{C} > 0$). In particular thus, if $\lambda^N_{\max}$ denotes the largest eigenvalue of X,

$$\mathbb{P}\Big(\Big\{\big|\lambda^N_{\max} - \mathbb{E}(\lambda^N_{\max})\big| \geq r\Big\}\Big) \leq 2\,\mathrm{e}^{-r^2/4C}, \quad r \geq 0$$

for every $r \geq 0$.

Another family of interest are product measures. The application of measure concentration to this class however requires an additional convexity assumption on the functionals. Indeed, if μ is a product measure on $\mathbb{R}^n$ with compactly supported factors, a fundamental result of M. Talagrand [Ta2] shows that (3.2) holds for every Lipschitz convex function. More precisely, assume that $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ where each μ_i is supported on $[a,b]$. Then, for every 1-Lipschitz convex function $F : \mathbb{R}^n \to \mathbb{R}$,

$$\mu\big(\{|F - m| \geq r\}\big) \leq 4\,\mathrm{e}^{-r^2/4(b-a)^2} \tag{3.8}$$

where m is a median of F for μ. (Classical arguments, cf. [Le1], allow for the replacement of m by the mean of F up to numerical constants.) Since, by the variational characterization, the largest eigenvalue $\lambda^N_{\max}$ of symmetric (or Hermitian) matrices is clearly a convex function of the entries, such a statement may immediately be applied to yield concentration inequalities similar to Propositions 3.1 and 3.2.

Proposition 3.3. *Let X be a real symmetric $N \times N$ matrix such that the entries X_{ij}, $1 \leq i \leq j \leq N$, are independent random variables with $|X_{ij}| \leq 1$. Denote by $\lambda^N_{\max}$ the largest eigenvalue of X. Then, for any $r \geq 0$,*

$$\mathbb{P}\Big(\big\{|\lambda^N_{\max} - M| \geq r\big\}\Big) \leq 4\,\mathrm{e}^{-r^2/32}$$

where M is a median of $\lambda^N_{\max}$.

Up to some numerical constants, the median may be replaced by the mean. A similar result is expected for all the eigenvalues. A partial result in [A-K-V] yields a bound of the order of $4\,\mathrm{e}^{-r^2/32\min(k,N-k+1)^2}$ on the k-th largest eigenvalue which, for k far from 1 or N is much bigger than the corresponding one in the Gaussian case for example. The analogous question for singular values (in particular the smallest one) in this context seems also to be open. Further inequalities on eigenvalues and norms following this principle, together with additional material, are discussed in [Mec] and [G-P].

We refer to [Ta2], [Le1] for further examples of distributions with the concentration property.

Similar measure concentration tools may be developed at the level of random growth functions. Consider for example an array $(w_{ij})_{1\le i\le M, 1\le j\le N}$ of real-valued random variables and let, as in the preceding sections,

$$W = \max_{\pi} \sum_{(i,j)\in\pi} w_{ij}$$

where the sup runs over all up/right paths from $(1,1)$ to (M,N). It is clear that $F(x) = \sup \sum_{(i,j)\in\pi} x_{ij}$ is a Lipschitz map of the MN coordinates $(x_{ij})_{1\le i\le M, 1\le j\le N}$ with Lipschitz constant $\sqrt{M+N-1}$. The following statement is thus an immediate consequence of the basic concentration principle. It applies thus in particular to independent Gaussian variables, or more general distributions satisfying a logarithmic Sobolev inequality. Since F is clearly a convex function of the coordinates, the result also applies to independent random variables with compact supports (such as for example Bernoulli variables).

Proposition 3.4. *Let $(w_{ij})_{1\le i\le M, 1\le j\le N}$ be a set of real-valued random variables such that the distribution on $\mathbb{R}^{MN}$ satisfies the concentration property (3.6) for all Lipschitz convex functions. Then, for any $r \ge 0$,*

$$\mathbb{P}\Big(\big\{\big|W - \mathbb{E}(W)\big| \ge r\big\}\Big) \le C\,\mathrm{e}^{-r^2/C(M+N-1)}.$$

While again of interest, and of rather wide applicability, this exponential bound however does not describe the expected rate drawn from the Meixner model as examined in the previous sections. In particular, the variance growth drawn from Proposition 3.4 with $M = N$ only yields $\mathrm{var}\,(W) \le C'\,N$ (where $C' > 0$ only depends on C) while it is expected to be of the order of $N^{2/3}$ (cf. Corollary 2.5). Similar comments apply to the concentration inequalities for the length of the longest increasing subsequence investigated in [Ta2] which do not match the Baik-Deift-Johansson theorem (1.21). Indeed, building on the general principle underlying (3.8), M. Talagrand got for example that

$$\mathbb{P}\big(\big\{|L_n - m| \ge r\big\}\big) \le 4\,\mathrm{e}^{-r^2/8m}$$

for $0 \le r \le m$.

3.2 Concentration inequalities for spectral distributions

The general concentration principles do not yield the correct small deviation rate at the level of the largest eigenvalues. They however apply to large classes of Lipschitz functions. In particular, as investigated by A. Guionnet and O. Zeitouni [G-Z], applications to functionals of the spectral measure yield sharp exponential bounds in accordance with the large deviation asymptotics for empirical measures (cf. Section 1.4). For example, if $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz, it is not difficult to check that $F = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} f(\lambda_i^N)$ is a Lipschitz function of the (real and imaginary) entries of X^N. Moreover, if f is convex on the real line, then F is convex on the space of matrices (Klein's lemma). Therefore, the general concentration principle may be applied to functions of the spectral measure. For example, if X is a GUE random matrix with variance $\sigma^2 = \frac{1}{4N}$, and if $f : \mathbb{R} \to \mathbb{R}$ is 1-Lipschitz, as a consequence of (3.2), for any $r \geq 0$,

$$\mathbb{P}\bigg(\bigg\{\bigg|\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i^N) - \int f d\mu^N\bigg| \geq r\bigg\}\bigg) \leq 2\,\mathrm{e}^{-2N^2r^2} \tag{3.9}$$

(where we recall that μ^N is the mean spectral measure (1.4)). Inequality (3.9) is in accordance with the N^2 speed of the large deviation principles for spectral measures. With the additional assumption of convexity on f, similar inequalities hold for real or complex matrices the entries of which are independent with bounded support. The various examples of distributions with the measure concentration property discussed for example in the previous sections may thus be developed similarly at the level of the spectral measures, and Propositions 3.2 and 3.3 have immediate counterparts for the Lipschitz functions F as above. We may for example state the following.

Proposition 3.5. *Let $X = (X_{ij})_{1 \leq i,j \leq N}$ be a real symmetric $N \times N$ random matrix. Assume that the distribution of the random vector X_{ij}, $1 \leq i \leq j \leq N$, in $\mathbb{R}^{N(N+1)/2}$ satisfy the dimension free concentration property (3.6) for all Lipschitz (resp. Lipschitz and convex) functions Then, for any 1-Lipschitz (resp. 1-Lipschitz and convex) function $f : \mathbb{R} \to \mathbb{R}$,*

$$\mathbb{P}\bigg(\bigg\{\bigg|\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i^N) - \int f d\mu^N\bigg| \geq r\bigg\}\bigg) \leq C\,\mathrm{e}^{-Nr^2/2C}$$

for all $r \geq 0$.

Extended inequalities have been investigated along these lines in [G-Z] to which we refer for further applications to various families of random matrices.

Interestingly enough, these concentration inequalities may be used to improve the Wigner theorem from the statement on the mean spectral measure to the almost sure conclusion. For example, in the context of the GUE, as a consequence of (3.9),

$$\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i^N) - \int f d\mu^N \to 0$$

almost surely for every Lipschitz function $f : \mathbb{R} \to \mathbb{R}$. Assuming that $\mu^N \to \nu$, the semicircle law, it easily follows after a density argument that, almost surely,

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^N} \to \nu$$

weakly as probability measures on $\mathbb{R}$.

3.3 Concentration inequalities for Coulomb gas

The GUE model shares both the structure of a Wigner matrix with independent entries and the one of a unitary invariant ensemble. As a unitary ensemble, we have seen in Chapter 1 how the joint eigenvalue distribution may be represented as a Coulomb gas (1.16). Under suitable convexity assumption on the underlying potential, Coulomb gas actually also share concentration properties which follow from general convexity principles.

Let indeed, as in (1.16),

$$dQ(x) = \frac{1}{Z} \left|\Delta_N(x)\right|^\beta d\rho(x)$$

where ρ is a probability measure on $\mathbb{R}^N$ and $Z = Z_N = \int |\Delta_N|^\beta d\rho < \infty$ the normalization constant. For particular values of $\beta > 0$ and suitable distributions ρ, Q thus represents the eigenvalue distribution of some random matrix model. We consider probability measures ρ given by $d\rho = \mathrm{e}^{-V} dx$ for some symmetric (invariant by permutation of the coordinates) potential $V : \mathbb{R}^N \to \mathbb{R}$. Typically, in the context of eigenvalues of random matrix models, $V(x) = \sum_{i=1}^N v(x_i)$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, where $v : \mathbb{R} \to \mathbb{R}$ is the underlying potential of the matrix distribution $\exp(-\mathrm{Tr}\, v(X))dX$. Assume now that $V(x) - c\,\frac{|x|^2}{2}$ is convex for some $c > 0$. For example, if

$$V(x) = \frac{|x|^2}{2} = \frac{1}{2} \sum_{i=1}^N x_i^2,$$

we would deal with the joint eigenvalue distribution of the GUE. By exchangeability, we may describe equivalently the measure Q by

$$dQ(x) = \frac{N!}{Z}\, \Delta_N(x)^\beta\, \mathbf{1}_E\, d\rho(x) \tag{3.10}$$

where $E = \{x \in \mathbb{R}^N; x_1 < \cdots < x_N\}$. Now, $\log \Delta_N^\beta$ is concave on the convex set E, so that the probability measure Q of (3.10) enters the general setting of probability measures with density e^{-U}, U strictly convex, on a convex set in $\mathbb{R}^N$. The general theory of the Prékopa-Leindler and transportation cost inequalities as presented in [Le1] then shows that Q satisfies a Gaussian like concentration inequality for Lipschitz functions. The next statement describes the result.

Proposition 3.6. *Let Q be defined by (3.10) with $d\rho = \mathrm{e}^{-V} dx$ where V is symmetric and such that $V(x) - c\,\frac{|x|^2}{2}$ is convex for some $c > 0$. Then, for any 1-Lipschitz function $F : \mathbb{R}^N \to \mathbb{R}$ and any $r \geq 0$,*

$$Q\Big(\big\{|F - \textstyle\int F dQ| \geq r\big\}\Big) \leq 2\, \mathrm{e}^{-r^2/2c}.$$

Applied to the particular Lipschitz function given by $\max_{1 \leq i \leq N} x_i$, we recover Proposition 3.1 for the GUE, which thus applies to more general orthogonal of unitary ensembles with a strictly convex potential. Proposition 3.6 may also be used to cover the concentration inequalities for Lipschitz functions of the spectral measure of the preceding section. However, again, the Tracy-Widom rate does not seem to follow from this description. It actually appears that in distributions $dQ(x) = \frac{1}{Z} \left|\Delta_N(x)\right|^\beta d\rho(x)$, the important factor is the Vandermonde determinant Δ_N and not the underlying probability measure ρ, while in the concentration approach, we rather focus on ρ.

4. HYPERCONTRACTIVE METHODS

We presented in Chapter 2 the known asymptotic exponential deviation inequalities on largest eigenvalues and last-passage times. As described there, these actually follow from quite refined methods and results. The aim of this chapter and the next one is to suggest some more accessible tools to reach some of these bounds (or parts of them) at the correct small deviation order. The tools developed in this chapter are of functional analytic flavour, with a particular emphasis on hypercontractive methods. They however still rely on the orthogonal polynomial representation.

In the first part, we present an elementary approach, relying on the hypercontractivity property of the Hermite semigroup, to the small deviation inequality of Proposition 2.1 for the largest eigenvalue of the GUE. We then investigate, following the recent contribution [B-K-S] by I. Benjamini, G. Kalai and O. Schramm, variance bounds for directed last-passage percolation with the same tool of hypercontractivity.

4.1 Upper tails on the right of the mean

We first briefly describe the semigroup tools we will be using. Consider the Hermite of Ornstein-Uhlenbeck operator

$$\mathcal{L}f = \Delta f - x \cdot \nabla f$$

acting on smooth functions $f : \mathbb{R}^n \to \mathbb{R}$. It satisfies the integration by parts formula

$$\int f(-\mathcal{L}g) d\mu = \int \nabla f \cdot \nabla g d\mu \tag{4.1}$$

for smooth functions f, g on $\mathbb{R}^n$ with respect to the standard Gaussian measure μ on $\mathbb{R}^n$. The associated semigroup $\mathcal{P}_t = \mathrm{e}^{t\mathcal{L}}$, $t \geq 0$, solution of the heat equation $\frac{\partial}{\partial t} = \mathcal{L}$, is, in this case, explicitely described by the integral representation

$$\mathcal{P}_t f(x) = \int_{\mathbb{R}^n} f\big(\mathrm{e}^{-t}x + (1 - \mathrm{e}^{-2t})^{1/2} y\big) d\mu(y), \quad t \geq 0,\ x \in \mathbb{R}^n. \tag{4.2}$$

Note that $\mathcal{P}_0 f = f$ and $\mathcal{P}_t f \to \int f d\mu$ (for suitable f's).

To illustrate $\mathcal{L}$ and $\mathcal{P}_t$ in the one-dimensional case, recall the generating function of the (normalized) Hermite polynomials P_ℓ, $\ell \in \mathbb{N}$, on the real line is given by

$$\mathrm{e}^{\lambda x - \lambda^2/2} = \sum_{\ell=0}^{\infty} \frac{\lambda^\ell}{\sqrt{\ell!}}\, P_\ell(x), \quad \lambda, x \in \mathbb{R}.$$

Since for every $t \geq 0$ and $\lambda \in \mathbb{R}$,

$$\mathcal{P}_t\big(\mathrm{e}^{\lambda x - \lambda^2/2}\big) = \mathrm{e}^{(\lambda \mathrm{e}^{-t})x - (\lambda \mathrm{e}^{-t})^2/2},$$

it follows that $\mathcal{P}_t(P_\ell) = \mathrm{e}^{-\ell t} P_\ell$, $\ell \in \mathbb{N}$. Hence the Hermite polynomials are the eigenfunctions of $\mathcal{L}$, with eigenvalues $-\ell$, $\ell \in \mathbb{N}$ ($\mathcal{L}$ is sometimes called the number operator).

The central tool in this section is the celebrated hypercontractivity property of the Hermite semigroup first put forward by E. Nelson [Ne] in quantum field theory. It expresses that, for any function f (in L^p),

$$\|\mathcal{P}_t f\|_q \leq \|f\|_p \tag{4.3}$$

for every $1 < p < q < \infty$ and $t > 0$ such that $\mathrm{e}^{2t} \geq \frac{q-1}{p-1}$ (cf. [Ba]). L^p-norms are understood here with respect to the Gaussian measure μ.

For comparison, it might be worthwhile mentioning that hypercontractivity has been shown by L. Gross [Gr] to be equivalent to the logarithmic Sobolev inequality (3.7) (with $C = 1$) for the standard normal distribution μ, in actually the general setting of Markov operators (cf. [Bak]).

We now make use of hypercontractivity to reach small deviation inequalities for the largest eigenvalues of the GUE. We follow the note [Le2]. Recall thus X from the GUE with $\sigma^2 = \frac{1}{4N}$, with eigenvalues $\lambda_1^N, \ldots, \lambda_N^N$. The starting point is the representation (1.13) of the spectral measure μ^N in terms of the Hermite polynomials and the simple union bound

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq t\}\big) \leq N \mu^N\big([\,t, \infty)\big), \quad t \in \mathbb{R},\ N \geq 1. \tag{4.4}$$

Let $N \geq 1$. As a consequence, for every $\varepsilon > 0$ (recall $\sigma^2 = \frac{1}{4N}$),

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1 + \varepsilon\}\big) \leq \int_{2\sqrt{N}(1+\varepsilon)}^{\infty} \sum_{\ell=0}^{N-1} P_\ell^2 \, d\mu \tag{4.5}$$

(where μ is here the standard Gaussian measure on $\mathbb{R}$). Now, by Hölder's inequality, for every $r > 1$, and every $\ell = 0, \ldots, N-1$,

$$\begin{aligned}\int_{2\sqrt{N}(1+\varepsilon)}^{\infty} P_\ell^2 \, d\mu &\leq \mu\big([2\sqrt{N}(1+\varepsilon), \infty)\big)^{1-(1/r)} \, \|P_\ell\|_{2r}^2 \\ &\leq \mathrm{e}^{-2N(1+\varepsilon)^2(1-\frac{1}{r})} \, \|P_\ell\|_{2r}^2\end{aligned}$$

where we used the standard bound on the tail of the Gaussian measure μ. Since as we have seen $\mathcal{P}_t(P_\ell) = \mathrm{e}^{-\ell t} P_\ell$, it follows from the hypercontractivity property (4.3) that for every $r > 1$ and $\ell \geq 0$,

$$\|P_\ell\|_{2r} \leq (2r-1)^{\ell/2}.$$

Hence,

$$\begin{aligned} \int_{2\sqrt{N}(1+\varepsilon)}^{\infty} \sum_{\ell=0}^{N-1} P_\ell^2 d\mu &\leq \mathrm{e}^{-2N(1+\varepsilon)^2(1-\frac{1}{r})} \sum_{\ell=0}^{N-1} (2r-1)^\ell \\ &\leq \frac{1}{2(r-1)} \, \mathrm{e}^{-2N(1+\varepsilon)^2(1-\frac{1}{r}) + N\log(2r-1)}. \end{aligned}$$

Optimizing in $r \to 1$ then shows, after a Taylor expansion of $\log(2r-1)$ at the third order, that for some numerical constant $C > 0$ and all $0 < \varepsilon \leq 1$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1 + \varepsilon\}\big) \leq C\,\varepsilon^{-1/2}\,\mathrm{e}^{-N\varepsilon^{3/2}/C} \tag{4.6}$$

for $C > 0$ numerical. Up to some polynomial factor, we thus recover the content of Proposition 2.1. (The argument is easily extended to also include the large deviation behavior of the order of ε^2 (cf. (2.7) and Section 2.3).

The same strategy may be developed similarly for orthogonal polynomial ensembles which may be diagonalized by an hypercontractive operator [Le2]. This is the case for example of the Laguerre operator, so that this approach yields exponential deviation inequalities for the largest eigenvalue of Wishart matrices or last-passage times for exponential random variables. The class of interest seems however to be restricted to the classical examples of Hermite, Laguerre and Jacobi polynomials [Ma]. Even the application of the method to discrete orthogonal polynomial ensembles does not seem to be clear.

4.2 Variance bounds

This section is devoted to the question of the variance growth of last-passage time functions for more general distributions than geometric and exponential. We follow here a recent contribution by I. Benjamini, G. Kalai and O. Schramm [B-K-S] who proved sub-linear growth by means of hypercontractive tools. In connection with the growth models discussed in the preceding sections, we only investigate here the directed percolation model. Furthermore, for simplicity again, we restrict ourselves in the exposition of this result to a Gaussian setting. It actually holds for a variety of examples discussed at the end of the section.

Consider thus, as in Section 3.1, an array $(w_{ij})_{1 \leq i \leq M, 1 \leq j \leq N}$ of independent standard Gaussian random variables and let

$$W = \max_\pi \sum_{(i,j)\in\pi} w_{ij}$$

where the maximum runs over all up/right paths from $(1,1)$ to (M,N). Assume furthermore for simplicity that $M=N$. We saw from the general concentration bounds in Chapter 3 that $\mathrm{var}\,(W) \leq CN$ (while it is expected to be of the order of $N^{2/3}$ by Corollary 2.5). We provide here, following [B-K-S], a slight, but significant improvement.

For a suitably integrable function $f : \mathbb{R}^{N^2} \to \mathbb{R}$, denote by

$$\mathrm{var}_\mu(f) = \int f^2 d\mu - \left(\int f d\mu \right)^2$$

its variance with respect to the standard Gaussian measure μ on $\mathbb{R}^{N^2}$. When f is smooth enough, the heat equation for the Ornstein-Uhlenbeck semigroup $(\mathcal{P})_{t\geq 0}$ with generator $\mathcal{L}$ on $\mathbb{R}^{N^2}$ allows one to write

$$\begin{aligned}
\mathrm{var}_\mu(f) &= -\int_0^\infty dt \, \frac{d}{dt} \int (\mathcal{P}_t f)^2 d\mu \\
&= 2\int_0^\infty dt \int \mathcal{P}_t f(-\mathcal{L}\mathcal{P}_t f) d\mu \\
&= 2\sum_{i,j=1}^N \int_0^\infty dt \int (\partial_{ij}\mathcal{P}_t f)^2 d\mu.
\end{aligned}$$

From the integral representation (4.2) of the Ornstein-Uhlenbeck semigroup,

$$|\partial_{ij}\mathcal{P}_t f| \leq \mathrm{e}^{-t}\,\mathcal{P}_t\big(|\partial_{ij} f|\big), \quad i,j = 1,\ldots,N, \ t \geq 0. \tag{4.7}$$

Hence, together with hypercontractivity (4.3), for every $t \geq 0$ and every $i,j = 1,\ldots,N$,

$$\int (\partial_{ij}\mathcal{P}_t f)^2 d\mu \leq e^{-2t} \int \big[\mathcal{P}_t\big(|\partial_{ij} f|\big)\big]^2 d\mu \leq e^{-2t} \|\partial_{ij} f\|^2_{1+e^{-2t}} .$$

Setting $u = e^{-2t}$, and $v = u+1$,

$$\mathrm{var}_\mu(f) \leq \sum_{i,j=1}^N \int_0^1 \|\partial_{ij} f\|^2_{1+u} du = \sum_{i,j=1}^N \int_1^2 \|\partial_{ij} f\|^2_v \, dv. \tag{4.8}$$

By a simple upper-bound on the right-hand side, this inequality may also be written as

$$\mathrm{Var}_\mu(f) \leq 4\sum_{i,j=1}^N \frac{\|\partial_{ij} f\|_2^2}{1+\log\big(\|\partial_{ij} f\|_2^2 / \|\partial_{ij} f\|_1^2\big)} . \tag{4.9}$$

Inequality (4.9) is actually due to M. Talagrand [Ta1] on the discrete cube, and was investigated in [B-H] in the Gaussian case as a dual version of the logarithmic Sobolev (3.7).

Recall now the (Lipschitz) function

$$F(x) = \max_{\pi} \sum_{(i,j)\in\pi} x_{ij}, \quad x = (x_{ij})_{0\le i,j\le N} \in \mathbb{R}^{N^2}.$$

The idea is to apply (4.8) or (4.9) to this function F to gain a factor $\log N$ in the variance by showing that $\|\partial_{ij}F\|_1$ is small enough, at least on sufficiently many coordinates (i,j). It is not immediate how this may be achieved*. To overcome this difficulty, the authors of [B-K-S] had to resort to an averaging lemma in the context of non-oriented percolation for Bernoulli variables (see also [B-R] for more general distributions). It is a natural hope that the same argument may be adapted to the present context to yield a proof of the following conjecture.

Conjecture 4.1. *Let W be the directed last-passage time of an array of independent standard Gaussian random variables on the square from $(1,1)$ to (N,N), $N\ge 2$. Then*

$$\mathrm{var}\,(W) \le \frac{CN}{\log N}$$

where $C>0$ is numerical.

Provided the argument develops suitably, the proof should then apply more generally to examples where both the hypercontractive bound and the commutation property (4.7) may be applied. One instance would be the example of uniform random variables. A further example is the case of exponential variables, for which however the much stronger Corollary 2.5 is available.

---

* The published proof of this result in the GAFA Seminar Lecture Notes is erroneous since it assumes that all the sets A_π are equally distributed. It is thus necessary to more carefully follow the argument in [B-K-S]. We are grateful to R. Rossignol for pointing out this mistake to us.

5. MOMENT METHODS

In this chapter, we take a somewhat different route from the one of Chapter 4 and concentrate on moments of the spectral distribution. Moment methods and combinatorial arguments are at the roots of the study of random matrix models, and for example are typically used in proofs of Wigner's theorem (cf. Section 1.1). The combinatorial part has been significantly improved by A. Soshnikov in [So1] to reach fluctuation results. Here, we again make advantage of the orthogonal polynomial structure to derive recurrence equations for moments, which may be shown of interest in non-asymptotic deviation inequalities. The strategy is based on integration by parts for the underlying Markov operator of the orthogonal polynomial ensemble.

In the first paragraph, we derive in this way the moment equations of the GUE model using simple integration by parts arguments for the Hermite operator. We then emphasize their usefulness in non-asymptotic deviation inequalities on the largest eigenvalues. Below the mean, we follow an argument by H. Widom relying on a simple trace inequality.

5.1 Moment recurrence equations

Let μ denote again the standard Gaussian measure on $\mathbb{R}$. By integration by parts, for every smooth function f on $\mathbb{R}$,

$$\int x f d\mu = \int f' d\mu. \tag{5.1}$$

(This formula is actually a particular case of the integration by parts formula (4.1) applied to $g = P_1 = x$ the first eigenvector of the one-dimensional Ornstein-Uhlenbeck operator $\mathcal{L}f = f'' - xf'$ with eigenvalue 1.)

By (1.13), moments of the mean spectral measure (1.4) amounts to moments of orthogonal polynomial measures. In this direction, we examine first a reduced case. Let

$$a_p = a_p^N = \int x^{2p} P_N^2 d\mu, \quad p \in \mathbb{N},$$

where we recall that the Hermite polynomials P_ℓ, $\ell \in \mathbb{N}$, are normalized in $\mathrm{L}^2(\mu)$. (The odd moments are zero by symmetry.) By (5.1),

$$a_p = \int x\, x^{2p-1} P_N^2 d\mu = (2p-1)a_{p-1} + 2\int x^{2p-1} P_N P_N' d\mu. \tag{5.2}$$

Repeating the same step,

$$\begin{aligned} a_p - (4p-3)a_{p-1} &+ (2p-1)(2p-3)a_{p-2} \\ &= 2\int x^{2p-2} {P_N'}^2 d\mu + 2\int x^{2p-2} P_N P_N'' d\mu. \end{aligned} \tag{5.3}$$

Now, P_N is an eigenfunction of $-\mathcal{L}$ with eigenvalue N. Thus, by the integration by parts formula (4.1) for $\mathcal{L}$,

$$N a_p = \int x^{2p} P_N(-\mathcal{L}P_N) d\mu = 2p\int x^{2p-1} P_N P_N' d\mu + \int x^{2p} {P_N'}^2 d\mu.$$

Together with (5.2),

$$\int x^{2p} {P_N'}^2 d\mu = (N-p)a_p + p(2p-1)a_{p-1}. \tag{5.4}$$

In the same way, on the basis of (5.2),

$$\begin{aligned} N\big[a_p - (2p-1)a_{p-1}\big] &= 2\int x^{2p-1}(-\mathcal{L}P_N)P_N' d\mu \\ &= 2(2p-1)\int x^{2p-2} {P_N'}^2 d\mu + 2\int x^{2p-1} P_N' P_N'' d\mu. \end{aligned}$$

Now, since $P_N' = \sqrt{N}\, P_{N-1}$ (which may be checked from the generating function of the Hermite polynomials) is eigenfunction of $-\mathcal{L}$ with eigenvalue $N-1$, we also have that

$$\begin{aligned} (N-1)\big[a_p - (2p-1)a_{p-1}\big] &= 2\int x^{2p-1} P_N(-\mathcal{L}P_N') d\mu \\ &= 2(2p-1)\int x^{2p-2} P_N P_N'' d\mu + 2\int x^{2p-1} P_N' P_N'' d\mu. \end{aligned}$$

Substracting to the latter,

$$a_p - (2p-1)a_{p-1} = 2(2p-1)\int x^{2p-2} {P_N'}^2 d\mu - 2(2p-1)\int x^{2p-2} P_N P_N'' d\mu,$$

so that, by (5.4),

$$\begin{aligned} 2(2p-1)\int x^{2p-2} P_N P_N'' d\mu = -a_p &+ (2p-1)(2N-2p+1)a_{p-1} \\ &+ (2p-1)(2p-2)(2p-3)a_{p-2}. \end{aligned} \tag{5.5}$$

Plugging (5.5) and (5.4) into (5.3) finally shows the recurrence equations

$$a_p = (4N+2)\,\frac{2p-1}{2p}\,a_{p-1} + \frac{(2p-1)(2p-2)(2p-3)}{2p}\,a_{p-2} \tag{5.6}$$

$(a_0 = 1,\ a_1 = 2N+1)$.

We now make use of the following elementary lemma that appears as a version in this context of the classical Christoffel-Darboux formula (1.15).

Lemma 5.1. *For every integer $k \geq 1$, and every $N \geq 1$,*

$$k \int x^{k-1} \sum_{\ell=0}^{N-1} P_\ell^2 d\mu = \sqrt{N} \int x^k P_N P_{N-1} d\mu.$$

Proof. Let $\mathcal{A}$ be the first order operator $\mathcal{A}f = f' - xf$ acting on smooth functions f on the real line $\mathbb{R}$. The integration by parts formula for $\mathcal{A}$ (analogous to (4.1)) indicates that for smooth functions f and g,

$$\int g(-\mathcal{A}f)d\mu = \int g' f d\mu.$$

Since $-\mathcal{L}P_N = NP_N$ and $P_N' = \sqrt{N}P_{N-1}$ for every $N \geq 1$, the recurrence relation for the (normalized) Hermite polynomials P_N, takes the form

$$xP_N = \sqrt{N+1}\,P_{N+1} + \sqrt{N}\,P_{N-1}.$$

Hence,

$$\begin{aligned} \mathcal{A}(P_N^2) &= P_N\big[2P_N' - xP_N\big] \\ &= \sqrt{N}\,P_N P_{N-1} - \sqrt{N+1}\,P_{N+1}P_N. \end{aligned}$$

Therefore,

$$(-\mathcal{A})\bigg(\sum_{\ell=0}^{N-1} P_\ell^2\bigg) = \sqrt{N}\,P_N P_{N-1}$$

from which the conclusion follows from the integration by parts formula for $\mathcal{A}$. □

Recall now the GUE random matrix $X = X^N$ with $\sigma^2 = \frac{1}{4N}$ and $N \geq 1$ fixed. Set, for every integer p,

$$\begin{aligned} b_p = b_p^N = \frac{1}{N}\,\mathbb{E}\big(\mathrm{Tr}\,(X^{2p})\big) &= \mathbb{E}\bigg(\frac{1}{N}\sum_{i=1}^N \big(\lambda_i^N\big)^{2p}\bigg) \\ &= \int_{\mathbb{R}} \bigg(\frac{x}{2\sqrt{N}}\bigg)^{2p} \frac{1}{N}\sum_{\ell=0}^{N-1} P_\ell^2 d\mu \end{aligned}$$

where we used (1.13). (The odd moments are zero by symmetry.) By Lemma 5.1,

$$(2p-1)\int x^{2p-2}\sum_{\ell=0}^{N-1}P_\ell^2 d\mu = \int x^{2p-1}P_N P_N' d\mu$$

so that, by (5.2),

$$2^{2p-1}N^p(2p-1)b_{p-1} = a_p - (2p-1)a_{p-1}$$

for every $p \geq 1$. As a consequence of (5.6), we may then deduce the following recurrence equations on the moments of X.

Proposition 5.2. *For every integer $p \geq 2$,*

$$b_p = \frac{2p-1}{2p+2}\,b_{p-1} + \frac{2p-1}{2p+2}\cdot\frac{2p-3}{2p}\cdot\frac{p(p-1)}{4N^2}\,b_{p-2}$$

($b_0 = 1$, $b_1 = \frac{1}{4}$.)

This recurrence equation, reminiscent of the three-step recurrence equation for orthogonal polynomials, was first put forward in an algebraic context by J. Harer and D. Zagier [H-Z] (to determine the Euler characteristics of moduli spaces of curves). It is also discussed in the book by M. L. Mehta [Meh]. The proof above is essentially due to U. Haagerup and S. Thorbjørnsen [H-T]. Similar recurrence identities may be established, with the same strategy, for the Laguerre and Jacobi orthogonal polynomials, and thus the corresponding moments of Wishart and Beta matrices [H-T], [Le3].

It should be pointed out that the equation

$$\chi_p = \frac{2p-1}{2p+2}\,\chi_{p-1} = \frac{(2p)!}{2^{2p}p!(p+1)!} \tag{5.7}$$

is the recurrence relation of the (even) moments of the semicircle law (the so-called Catalan numbers, the number of non-crossing pair partitions of $\{1, 2, \ldots, 2p\}$). In particular, Proposition 5.2 may then be used to produce a quick proof of the Wigner theorem, showing namely that $b_p^N \to \chi_p$ for every p. Moreover, for every fixed p and every $N \geq 1$,

$$\chi_p \leq b_p^N \leq \chi_p + \frac{C_p}{N^2}$$

where $C_p > 0$ only depends on p.

5.2 Upper tails on the right of the mean

We next make use of the recurrence equations of Proposition 5.2 to recover the sharp exponential bounds on the probability that the largest eigenvalues of the GUE matrix exceeds its limiting value discussed by other means in the preceding chapters.

We start again from (4.5). Together with Markov's inequality, for every $N \geq 1$, $\varepsilon > 0$ and $p \geq 0$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1+\varepsilon\}\big) \leq (1+\varepsilon)^{-2p} N b_p$$

where we recall that $b_p = b_p^N$ are the $2p$-moments of μ^N (or $X = X^N$). Now, by induction on the recurrence formula of Proposition 5.2 for b_p, it follows that, for every $p \geq 2$,

$$b_p \leq \bigg(1 + \frac{p(p-1)}{4N^2}\bigg)^p \chi_p. \tag{5.8}$$

By Stirling's formula,

$$\chi_p \leq \frac{C}{p^{3/2}}\,, \quad p \geq 1. \tag{5.9}$$

Hence, for $0 < \varepsilon \leq 1$ and some numerical constant $C > 0$ possibly changing from line to line below,

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1+\varepsilon\}\big) \leq CNp^{-3/2}\, \mathrm{e}^{-\varepsilon p + p^3/4N^2}.$$

Therefore, optimizing in $p \sim \sqrt{\varepsilon}\, N$, $0 < \varepsilon \leq 1$, we recover the sharp small deviation inequality

$$\mathbb{P}\big(\{\lambda_{\max}^N \geq 1+\varepsilon\}\big) \leq C\, \mathrm{e}^{-N\varepsilon^{3/2}/C},$$

$N \geq 1$, $0 < \varepsilon \leq 1$, $C > 0$ numerical, of Proposition 2.1. When $\varepsilon \geq 1$, the optimization is modified to recover the large deviation rate of the order of $N\varepsilon^2$. With respect to the hypercontractive approach of Chapter 4, no further polynomial factors have to be added.

As observed by S. Szarek in [Sza], the moment recurrence equation of Proposition 5.2 and the preceding argument may be used to reach the sharp upper bound (2.12) on $\mathbb{E}(\lambda_{\max}^N)$, and even

$$\mathbb{E}\Big(\max_{1 \leq i \leq N} |\lambda_i^N|\Big) \leq 1 - \frac{1}{CN^{2/3}} \tag{5.10}$$

for some numerical $C > 0$ and all $N \geq 1$. Indeed, for every $p \geq 1$,

$$\mathbb{E}\Big(\max_{1 \leq i \leq N} |\lambda_i^N|\Big) \leq \big(Nb_p\big)^{1/2p},$$

and thus, by (5.8),

$$\mathbb{E}\Big(\max_{1 \leq i \leq N} |\lambda_i^N|\Big) \leq \big(N\chi_p\, \mathrm{e}^{p^3/4N^2}\big)^{1/2p}. \tag{5.11}$$

If $t > 0$ and $N \geq 1$ are such that $p = [tN^{2/3}] \geq 3$, then, together with (5.9),

$$\mathbb{E}\Big(\max_{1 \leq i \leq N} |\lambda_i^N|\Big) \leq \bigg[\Big(\frac{C^{2/3}}{t}\Big)^{3/2t} \mathrm{e}^{t^2/4}\bigg]^{1/2N^{2/3}}.$$

The constant $C > 0$ in (5.9) may be taken to be $\pi^{-1/2}$ so that taking for example $t = C\sqrt{\mathrm{e}}$ shows that the bracket in the preceding inequality is strictly less than 1. Therefore (5.10) holds except for some few values of N which may be checked directly on the basis of (5.11).

As a consequence of (5.8), for every $t > 0$,

$$\sup_{N \geq 1} N b^N_{[tN^{2/3}]} \leq C t^{-3/2} \mathrm{e}^{Ct^3} \tag{5.12}$$

for the moments of the GUE. One important step in Soshnikov's extension [So1] of the Tracy-Widom theorem to more general (real or complex) Wigner matrices amounts to establish that $\limsup_{N\to\infty} N b^N_{[N^{2/3}]} < \infty$, actually

$$\limsup_{N\to\infty} N b^N_{[tN^{2/3}]} \leq C t^{-3/2} \mathrm{e}^{Ct^3}$$

for every $t > 0$. This is accomplished through delicate combinatorial arguments on moments. It is however an open question so far whether its non-asymptotic version (5.12) also holds for these families of random matrices, which would then yield the expected deviation inequalities for every fixed $N \geq 1$. It would be of particular interest to study the case of Bernoulli entries.

Very recently, O. Khorunzhiy [Kh] developed Gaussian integration by parts methods at the level of traces, together with triangle recurrence schemes, to reach moment bounds on both the GUE and the GOE. The estimates provide exact expressions for the $1/N$-corrections of the moments, but do not allow yet for the sharp deviation inequalities on largest eigenvalues. This first step outside the orthogonal polynomial method might however be promising.

This strategy relying on integration by parts for Markov generators and moment equations may be used similarly for some other classical orthogonal polynomial ensembles of the continuous variables. For example, the Laguerre and Jacobi ensembles are studied along these lines in [Le3] to yield deviation inequalities at the Tracy-Widom rate of the largest eigenvalue of Wishart and Beta matrices. Discrete examples may also be considered, although not necessarily through recurrence equations, but rather the explicit expression for moments (this is actually also possible in the continuous variable). For example, integration by parts with respect to the negative binomial distribution (1.17) with parameters q and γ reads

$$\int x f d\mu = \frac{\gamma q}{1-q} \int f(x+1) d\mu$$

for any, say, polynomial function f on $\mathbb{N}$. It allows one to express the factorial moment (of order p) of the mean spectral measure of the Meixner Ensemble as

$$\begin{aligned}\int x(x-1)\cdots(x-p+1)\frac{1}{N}\sum_{\ell=0}^{N-1}P_\ell^2 d\mu \\ = \Big(\frac{q}{1-q}\Big)^p \sum_{i=0}^{p} q^{-i}\binom{p}{i}^2 \frac{1}{N}\sum_{\ell=i}^{N-1}\frac{(\gamma+\ell)_{p-i}\,\ell\,!}{(\ell-i)!}\end{aligned}\tag{5.13}$$

where P_ℓ are the associated normalized Meixner polynomials. Therefore, if Q is the Coulomb gas of the Meixner Ensemble, by the union bound inequality (4.4) and the representation (1.13) of the spectral measure, for every $t \geq 0$,

$$\begin{aligned}Q\Big(\big\{\max_{1\leq i\leq N} x_i \geq t\big\}\Big) &\leq \int_t^\infty \sum_{\ell=0}^{N-1}P_\ell^2 d\mu \\ &\leq \frac{(t-p)!}{t!}\Big(\frac{q}{1-q}\Big)^p \sum_{i=0}^{p} q^{-i}\binom{p}{i}^2 \sum_{\ell=i}^{N-1}\frac{(\gamma+\ell)_{p-i}\,\ell\,!}{(\ell-i)!}\,.\end{aligned}$$

Stirling's formula may then be used to control the right-hand side of the latter, and to derive exponential deviation inequalities on the rightmost charge. Together with Johansson's combinatorial formula (1.19), the conclusion of Proposition 2.2 on the random growth functions W may be recovered in this way, avoiding superadditivity and large deviation arguments. In the limit from the Meixner Ensemble to the length of the longest increasing subsequence L_n, it also covers the tail inequality (2.9) (cf. [Le4]). However, the key of the analysis still relies on the orthogonal polynomial representation.

5.3 Upper tails on the left of the mean

We next turn, with the tool of moment identities, to the probability that the largest eigenvalue of the GUE is less than or equal to 1. As discussed in Chapter 2, bounds on this probability turn out to be much more delicate. We present here a simple inequality, in the context of the GUE, taken from the note [Wid2] by H. Widom.

We start from the determinantal description (1.11)

$$\mathbb{P}\big(\{\lambda_{\max}^N \leq t\}\big) = \det\big(\mathrm{Id} - K\big)$$

where, for each $t \in \mathbb{R}$, $K = K_t$ is the symmetric $N \times N$ matrix

$$\big(\langle P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2((t/\sigma,\infty),d\mu)}\big)_{1\leq k,\ell\leq N}.$$

Since for any unit vector $u = (u_1, \ldots, u_N) \in \mathbb{R}^N$,

$$0 \leq \sum_{k,\ell=1}^{N} u_k u_\ell \langle P_{\ell-1}, P_{k-1}\rangle_{\mathrm{L}^2((t/\sigma,\infty),d\mu)} \leq 1,$$

the eigenvalues $\rho_1, \ldots, \rho_N$ of K are all non-negative and less than or equal to 1. Hence,

$$\mathbb{P}\big(\{\lambda_{\max}^N \le t\}\big) = \prod_{i=1}^N (1-\rho_i) \le \mathrm{e}^{-\Sigma_{i=1}^N \rho_i}.$$

Now, by (1.13),

$$\sum_{i=1}^N \rho_i = \sum_{\ell=1}^N \langle P_{\ell-1}^2 \rangle_{\mathrm{L}^2((t/\sigma,\infty),d\mu)} = N\mu^N\big((t,\infty)\big).$$

Therefore, for every $t \in \mathbb{R}$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \le t\}\big) \le \exp\Big(-N\mu^N\big((t,\infty)\big)\Big). \tag{5.14}$$

Note that (5.14) would be the inequality that one would deduce if the λ_i^N's were independent. The latter are however strongly correlated so that (5.14) already misses a big deal of the interactions between the eigenvalues.

Let us now apply (5.14) to $t = 1-\varepsilon$, $0 < \varepsilon \le 1$. By Wigner's theorem, $\mu^N((1-\varepsilon,\infty) \to \nu((1-\varepsilon,1))$ where we recall that ν is the semicircle law (on $(-1,+1)$). It is easy to evaluate

$$\nu\big((1-\varepsilon,1)\big) \ge C^{-1}\varepsilon^{3/2}, \quad 0 < \varepsilon \le 1,$$

where $C > 0$ is numerical. We expect that

$$\mu^N\big((1-\varepsilon,\infty)\big) \ge C^{-1}\varepsilon^{3/2}, \quad 0 < \varepsilon \le 1, \tag{5.15}$$

at least for every $N \ge 1$ such that $CN^{-2/3} \le \varepsilon \le 1$. To this task, we could invoke a recent result of F. Götze and A. Tikhomirov [G-T] on the rate of convergence of the spectral measure of the GUE to the semicircle which implies that

$$\Big|\mu^N\big((1-\varepsilon,\infty)\big) - \nu\big((1-\varepsilon,1)\big)\Big| \le \frac{C}{N}$$

for some $C > 0$ and all $N \ge 1$, and thus (5.15). While the proof of [G-T] requires quite a bit of analysis, we provide here an independent elementary argument to reach (5.15) using the moment equations.

Fix $N \ge 1$ and $0 < \varepsilon \le 1$. For every $p \ge 1$,

$$b_{2p} = \int_{\mathbb{R}} x^{4p} d\mu^N(x) \le (1-\varepsilon)^{2p}\, b_p + 2\int_{1-\varepsilon}^{\infty} x^{4p} d\mu^N(x).$$

From the recurrence equations put forward in Proposition 5.2, for every p,

$$b_{2p} \ge \chi_{2p}$$

while (cf. (5.8))

$$b_p \leq \left(1 + \frac{p(p-1)}{4N^2}\right)^p \chi_p \leq \mathrm{e}^{p^3/4N^2}\, \chi_p.$$

By the Cauchy-Schwarz inequality,

$$\int_{1-\varepsilon}^{\infty} x^{4p} d\mu^N(x) \leq \mu^N\big((1-\varepsilon,\infty)\big)^{1/2} b_{4p}^{1/2}.$$

Hence,

$$\mu^N\big((1-\varepsilon,\infty)\big) \geq 4^{-1}\, \mathrm{e}^{-16p^3/N^2}\, \chi_{4p}^{-1} \Big[\chi_{2p} - \mathrm{e}^{p^3/4N^2}\, \chi_p\Big]^2.$$

Choose then $p = [\varepsilon^{-1}]$ and assume that $N^{-2/3} \leq \varepsilon \leq 1$. Then

$$\mu^N\big((1-\varepsilon,\infty)\big) \geq \mathrm{e}^{-18}\, \chi_{4p}^{-1} \Big[\chi_{2p} - \mathrm{e}^{1/4}\, \chi_p\Big]^2.$$

Since by Stirling's formula $\chi_p \sim \pi^{-1/2} p^{-3/2}$ as $p \to \infty$, uniform bounds show that for some constant $C > 0$,

$$\mu^N\big((1-\varepsilon,\infty)\big) \geq C^{-1} \varepsilon^{3/2}.$$

Together with (5.14), we thus conclude that for some $C > 0$, every $N \geq 1$ and every ε such that $N^{-2/3} \leq \varepsilon \leq 1$,

$$\mathbb{P}\big(\{\lambda_{\max}^N \leq 1-\varepsilon\}\big) \leq C\, \mathrm{e}^{-N\varepsilon^{3/2}/C}. \tag{5.16}$$

Increasing if necessary C, the inequality easily extends to all $0 < \varepsilon \leq 1$. The deviation inequality (5.16) is weaker than the one of Proposition 2.4 and does not reflect the N^2 rate of the large deviation asymptotics. Its proof is however quite accessible, and gives a firm basis to $N^{-4/3}$ growth rate of Corollary 2.6.

The preceding argument may be extended to more general orthogonal polynomial ensembles provided the corresponding version of (5.15) can be established. In case for example of the Meixner Ensemble, the explicit expression (5.13) for the factorial moments of the mean spectral measure might be useful to this task.

REFERENCES

[A-D] D. Aldous, P. Diaconis. Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem. Bull. Amer. Math. Soc. 36, 413–432 (1999).

[A-K-V] N. Alon, M. Krivelevich, V. Vu. On the concentration of eigenvalues of random symmetric matrices. Israel J. Math. 131, 259–267 (2002).

[Ar] L. Arnold. On the asymptotic distribution of the eigenvalues of random matrices. J. Math. Anal. Appl. 20, 262–268 (1967).

[Au] G. Aubrun. An inequality about the largest eigenvalue of a random matrix (2003). Séminaire de Probabilités XXXVIII. Lecture Notes in Math. 1857, 320–337 (2005). Springer.

[Bai] Z. D. Bai. Methodologies in the spectral analysis of large dimensional random matrices: a review. Statist. Sinica 9, 611–661 (1999).

[Baik1] J. Baik. Random vicious walks and random matrices. Comm. Pure Appl. Math. 53, 1385–1410 (2000).

[Baik2] J. Baik. Limit distribution of last passage percolation models (2003).

[B-D-K] J. Baik, P. A. Deift, K. Johansson. On the distribution of the length of the longest increasing subsequence in a random permutation. J. Amer. Math. Soc. 12, 1119–1178 (1999).

[B-D-ML-M-Z] J. Baik, P. A. Deift, K. McLaughlin, P. Miller, X. Zhou. Optimal tail estimates for directed last passage site percolation with geometric random variables. Adv. Theor. Math. Phys. 5, 1207–1250 (2001).

[B-D-R] J. Baik, P. A. Deift, E. Rains. A Fredholm determinant identity and the convergence of moments for random Young tableaux. Comm. Math. Phys. 223, 627–672 (2001).

[B-K-ML-M] J. Baik, T. Kriecherbauer, K. McLaughlin, P. Miller. Uniform Asymptotics for Polynomials Orthogonal With Respect to a General Class of Discrete Weights and Universality Results for Associated Ensembles. Announcements of results: Int. Math. Res. Not. 15, 821–858 (2003). Full version: Annals of Math. Studies, to appear.

[Bak] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. Ecole d'Eté de Probabilités de St-Flour. Lecture Notes in Math. 1581, 1–114 (1994). Springer.

[Bar] Y. Baryshnikov. GUES and queues. Probab. Theor. Ral. Fields 119, 256–274 (2001).

[BA-G] G. Ben Arous, A. Guionnet. Large deviations for Wigner's law and Voiculescu's noncommutative entropy. Probab. Theor. Rel. Fields 108, 517–542 (1997).

[BA-D-G] G. Ben Arous, A. Dembo, A. Guionnet. Aging of spherical spin glasses. Probab. Theor. Relat. Fields 120, 1–67 (2001).

[B-R] M. Benaïm, R. Rossignol. In preparation (2006).

[B-K-S] I. Benjamini, G. Kalai, O. Schramm. First passage percolation has sublinear distance variance. Ann. Probab. 31, 1970–1978 (2003).

[Bi] P. Biane. Free probability for probabilists. Quantum Probability Communications, vol. XI, 55–71 (2003). World Sci. Publishing.

[B-I] P. Bleher, A. Its. Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. Ann. Math. 150, 185–266 (1999).

[Bl] G. Blower. Concentration of measure for large eigenvalues of unitary ensembles, and Airy operators (2004).

[B-H] S. Bobkov, C. Houdré. A converse Gaussian Poincaré-type inequality for convex functions. Statist. Probab. Letters 44, 281–290 (1999).

[B-O-O] A. Borodin, A. Okounkov, G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. J. Amer. Math. Soc. 13, 481–515 (2000).

[Ch] T. S. Chihara. An introduction to orthogonal polynomials. Gordon & Breach (1978).

[Co] B. Collins. Product of random projections, Jacobi ensembles and universality problems arising from free probability. Probab. Theory Relat. Fields 133, 315–344 (2005).

[Da-S] K. Davidson, S. Szarek. Local operator theory, random matrices and Banach spaces. Handbook of the Geometry of Banach Spaces, vol. I, 317–366 (2001). North Holland.

[De] P. A. Deift. Orthogonal polynomials and random matrices: a Riemann-Hilbert approach. CIMS Lecture Notes 3. Courant Institute of Mathematical Sciences (1999).

[D-Z] P. A. Deift, X. Zhou. A steepest descent method for oscillatory Riemann-Hilbert problems; asymptotics for the MKdV equation. Ann. Math. 137, 295–368 (1993).

[D-K-ML-V-Z] P. A. Deift, T. Kriecherbauer, K. McLaughlin, S. Venakides, X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. Comm. Pure Appl. Math. 52, 1335–1425 (1999).

[D-Ze] J.-D. Deuschel, O. Zeitouni. On increasing subsequences of I.I.D. samples. Combin. Probab. Comput. 8, 247–263 (1999).

[Du-S] N. Dunford, J. T. Schwartz. Linear Operators. Part II: Spectral Theory. Wiley (1988).

[Fe] D. Feral. On large deviations for the spectral measure of discrete Coulomb gas (2004).

[Fo1] P. J. Forrester. The spectrum edge of random matrix ensembles. Nuclear Phys. B 402, 709–728 (1993).

[Fo2] P. J. Forrester. Log-gases and random matrices (2002).

[Fu] W. Fulton. Young tableaux. With applications to representation theory and geometry. London Math. Soc. Student Texts, 35. Cambridge University Press (1997).

[Fy] Y. V. Fyodorov. Introduction to the random matrix theory: Gaussian Unitary Ensemble and beyond. Recent perspectives in random matrix theory and number theory, London Math. Soc. Lecture Note Ser., 322, 31–78 (2005). Cambridge Univ. Press.

[G-G] I. Gohberg, S. Goldberg. Basic operator theory. Birkhaüser (1981).

[G-T] F. Götze, A. Tikhomirov. The rate of convergence for spectra of GUE and LUE matrix ensembles. Cent. Eur. J. Math. 3, 666–704 (2005).

[G-T-W] J. Gravner, G. Tracy, H. Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. J. Statist. Phys. 102, 1085–1132 (2001).

[Gr] L. Gross. Logarithmic Sobolev inequalities. Amer. J. Math. 97, 1061–1083 (1975).

[G-P] O. Guédon, G. Paouris. Concentration of mass on the Schatten classes (2005).

[G-Z] A. Guionnet, O. Zeitouni. Concentration of the spectral measure for large matrices. Electron. Comm. Probab. 5, 119-136 (2000).

[H-T] U. Haagerup, S. Thorbjørnsen.Random matrices with complex Gaussian entries. Expo. Math. 21, 293–337 (2003).

[H-Z] J. Harer, D. Zagier. The Euler characteristic of the moduli space of curves. Invent. math. 85, 457–485 (1986).

[H-P] F. Hiai, D. Petz. The semicircle law, free random variables and entropy. Mathematical Surveys and Monographs 77. AMS (2000).

[Joha1] K. Johansson. Shape fluctuations and random matrices. Comm. Math. Phys. 209, 437–476 (2000).

[Joha2] K. Johansson. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. Comm. Math. Phys. 215, 683–705 (2001).

[Joha3] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. Ann. Math. 259–296 (2001).

[Joha4] K. Johansson. Toeplitz determinants, random growth and determinantal processes. Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002), 53–62. Higher Ed. Press, Beijing (2002).

[John] I. M. Johnstone. On the distribution of the largest principal component. Ann. Statist. 29, 295–327 (2001).

[Kh] O. Khorunzhyi. Estimates for moments of random matrices with Gaussian elements (2005).

[K-S] R. Koekoek, R. F. Swarttouw. The Askey scheme of hypergeometric orthogonal polynomials and its q-analogue (1998). http://aw.twi.tudelft.nl/~koekoek/askey/index.html.

[Kö] W. König. Orthogonal polynomial ensembles in probability theory. Probability Surveys 2, 385–447 (2005).

[Ku] A. Kuijlaars. Riemann-Hilbert analysis for orthogonal polynomials. Orthogonal polynomials and special functions (Leuven, 2002). Lecture Notes in Math. 1817, 167–210 (2003). Springer.

[Le1] M. Ledoux. The concentration of measure phenomenon. Mathematical Surveys and Monographs 89. AMS (2001).

[Le2] M. Ledoux. A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices. Séminaire de Probabilités XXXVII. Lecture Notes in Math. 1832, 360–369 (2003). Springer.

[Le3] M. Ledoux. Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The continuous case. Electron. J. Probab. 9, 177–208 (2004).

[Le4] M. Ledoux. Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The discrete case. Electron. J. Probab. 10, 1116-1146 (2005).

[L-T] M. Ledoux, M. Talagrand. Probability in Banach spaces (Isoperimetry and processes). Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer (1991).

[L-M] M. Löwe, F. Merkl. Moderate deviations for longest increasing subsequences: the upper tail. Comm. Pure Appl. Math. 54, 1488–1520 (2001).

[L-M-R] M. Löwe, F. Merkl, S. Rolles. Moderate deviations for longest increasing subsequences: the lower tail. J. Theoret. Probab. 15, 1031–1047 (2002).

[M-P] V. Marchenko, L. Pastur. The distribution of eigenvalues in certain sets of random matrices. Math. Sb. 72, 507–536 (1967).

[Ma] O. Mazet. Classification des semi-groupes de diffusion sur $\mathbb{R}$ associés à une famille de polynômes orthogonaux. Séminaire de Probabilités XXXI. Lecture Notes in Math. 1655, 40–53 (1997). Springer.

[Mec] M. Meckes. Concentration of norms and eigenvalues of random matrices. J. Funct. Anal. 211, 508–524 (2004).

[Meh] M. L. Mehta. Random matrices. Academic Press (1991). Third edition Elsevier (2004).

[Ne] E. Nelson. The free Markov field. J. Funct. Anal. 12, 211–227 (1973).

[OC] N. O'Connell. Random matrices, non-colliding processes and queues. Séminaire de Probabilités XXXVI. Lecture Notes in Math. 1801, 165-182 (2003). Springer.

[OC-Y] N. O'Connell, M. Yor. A representation for non-colliding random walks. Electron. Commun. Probab. 7, 1–12 (2000).

[P-L] L. Pastur, A. Lejay. Matrices aléatoires: statistique asymptotique des valeurs propres. Séminaire de Probabilités XXXVI. Lecture Notes in Math. 1801, 135–164 (2003). Springer.

[P-S] L. Pastur, M. Sherbina. On the edge unviversality of the local eigenvalue statistics of matrix models (2003).

[S-T] E. B. Saff, V. Totik. Logarithmic potentials with external fields. Springer (1997).

[Se1] T. Seppäläinen. Large deviations for increasing sequences on the plane. Probab. Theory Relat. Fields 112, 221–244 (1998).

[Se2] T. Seppäläinen. Exact limiting shape for a simplified model of first-passage percolation on the plane. Ann. Probab. 26, 1232-1250 (1998).

[Se3] T. Seppäläinen. Coupling the totally asymmetric simple exclusion process with a moving interface. Markov Proc. Relat. Fields 4, 593–628 (1998).

[So1] A. Soshnikov. Universality at the edge of the spectrum in Wigner random matrices. Comm. Math. Phys. 207, 697–733 (1999).

[So2] A. Soshnikov. Determinantal random point fields. Uspekhi Mat. Nauk 55, 107–160 (2000).

[Sza] S. Szarek. Volume of separable states is super-doubly-exponentially small in the number of qubits. Phys. Rev. A, 72, (2005).

[Sze] G. Szegö. Orthogonal polynomials. Colloquium Publications XXIII. Amer. Math Soc. (1975).

[Ta1] M. Talagrand. On Russo's approximate zero-one law. Ann. Probab. 22, 1576–1587 (1994).

[Ta2] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. Publications Mathématiques de l'I.H.E.S. 81, 73–205 (1995).

[Ta3] M. Talagrand. The generic chaining. Springer (2005).

[T-W1] C. Tracy, H. Widom. Level-spacing distribution and the Airy kernel. Comm. Math. Phys. 159, 151–174 (1994).

[T-W2] C. Tracy, H. Widom. On Orthogonal and Symplectic Matrix Ensembles. Comm. Math. Phys. 177, 727–754 (1996).

[T-W3] C. Tracy, H. Widom. On the distributions of the lengths of the longest monotone subsequences in random words. Probab. Theor. Rel. Fields 119, 350–380 (2001).

[T-W4] C. Tracy, H. Widom.Distribution functions for largest eigenvalues and their applications. Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), 587–596. Higher Ed. Press, Beijing (2002).

[Vo1] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory, I. Comm. Math. Phys. 155, 71–92 (1993).

[Vo2] D. Voiculescu. Lecture on free probability theory. Ecole d'Eté de Probabilités de St-Flour. Lecture Notes in Math. 1738, 279–349 (2000). Springer.

[V-D-N] D. Voiculescu, K. Dykema, A. Nica. Free random variables. CRM Monograph Ser. 1. Amer. Math. Soc. (1992).

[Wid1] H. Widom. On the relation between orthogonal, symplectic and unitary matrix ensembles. J. Statist. Phys. 94, 347–363 (1999).

[Wid2] H. Widom. On convergence of moments for random Young tableaux and a random growth model. Int. Math. Res. Not. 9, 455–464 (2002).

[Wig] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. Ann. Math. 62, 548–564 (1955).

Institut de Mathématiques, Université Paul-Sabatier, 31062 Toulouse, France

E-mail address: `ledoux@math.ups-tlse.fr`