Log correlated fields in random matrices and their extremes
Notes for Lectures

Ofer Zeitouni

Department of Mathematics
Weizmann Institute, Rehovot 76100, Israel
September 5, 2025. Version 1.04

DISCLAIMER: There is no attempt in these notes to give proper historical credit and references.

1 Lecture I: the prototype: Branching Random Walks

Branching random walks (BRWs), and their continuous time counterparts, branching Brownian motions (BBMs), form a natural model that describe the evolution of a population of particles where spatial motion is present. Groundbreaking work on this, motivated by biological applications, was done in the 1930's by Kolmogorov-Petrovsky-Piskounov and by Fisher. The model itself exhibit a rich mathematical structures; for example, rescaled limits of such processes lead to the study of superprocesses, and allowing interactions between particles creates many challenges when one wants to study scaling limits.

Our focus is slightly different: we consider only particles in $\mathbb{R}$, and are mostly interested in the atypical particles that "lead the pack". We will restrict attention to Gaussian centered increments. Some of the exercises extend this to more general situations, of relevance to random matrices.

1.1 Definitions and models

We begin by fixing notation. Let $\mathcal{T}$ be a binary tree rooted at a vertex o, with vertex set V and edge set E. We denote by $|v|$ the distance of a vertex v from the root, i.e. the length of the geodesic (=shortest path, which is unique) connecting v to o, and we write $o \leftrightarrow v$ for the collection of vertices on that geodesic (including o and v). With some abuse of notation, we also write $o \leftrightarrow v$ for the collection of *edges* on the geodesic connecting o and v. Similarly, for $v, w \in V$, we write $\rho(v, w)$ for the length of the unique geodesic connecting v and w, and define $v \leftrightarrow w$ similarly. The nth generation of the tree is the collection $D_n := \{v \in V : |v| = n\}$, while for $v \in D_m$ and $n > m$, we denote by

$$D_n^v = \{w \in D_n : \rho(w, v) = n - m\}$$

the collection of descendants of v in D_n. We call the descendants of v which are neighbors of v the *children* of v. Finally, we designate (somewhat arbitrarily) one of the children of a vertex as the "left child" and the other as a "right child".

Let $\{X_e\}_{e\in E}$ denote a family of independent (real valued) standard Gaussian random variables attached to the edges of the tree $\mathcal{T}$. For $v \in V$, set $S_v = \sum_{e\in o\leftrightarrow v} X_e$. The *Branching Random Walk* (BRW) is simply the collection of random variables $\{S_v\}_{v\in V}$.

1.2 The log-correlated structure

Because the BRW is Gaussian, the collection $\{S_v\}_{v\in D_n}$ is completely characterized by its mean (= 0), and its covariance function

$$R(v, w) = EX_vX_w = n - \rho(v, w)/2 = |a_{v,w}|,$$

where $a_{v,w}$ is the *common ancestor* of v, w, defined as the (unique) vertex of largest distance that belongs to both $o \leftrightarrow v$ and $o \leftrightarrow w$. We also write $a_{v,k}$ for the *k-th ancestor of* v (with $k < n$), ie the unique vertex on $o \leftrightarrow v$ with distance k from the root.

To understand why we think of this process as log-correlated, we embed the vertices of D_n in the interval $[0, 1]$ as follows: each vertex v determines a binary string of length n, denotes $[v]_n$, whose k-th digit is 0 or 1 according to whether the descendent of the k-th ancestor of v is the left or right child. We identify $[v]_n$ with a point in $[0, 1]$ in the natural way. Note that this identification is consistent, in the sence that if we take an infinite geodesic and consider $v \in D_n$ along that geodesic, then as $n \to \infty$ this identified point converges. In particular, for each n we obtain a Gaussian process $S_v = Y_n(v)$, $v = i2^{-n}$, $i = 0, \ldots, 2^n - 1$, with $\mathrm{Var}(S_v) = n$. Now, fix a scale ℓ. Choose a dyadic interval $I_\ell = [j2^\ell, (j + 1)2^\ell]$ at random.

Exercise 1. Pick uniformly and randomly two points x, y in I_ℓ that correspond to $v, w \in D_n$. Then, as $n \to \infty$, $R(v, w) = \ell + O(1) = \log \frac{1}{|x-y|} + O(1)$.

The (embedded) BRW is thus "on average" a log correlated field, but of course not truly: for two points $1/2 + \epsilon$ and $1/2 - \epsilon$, the covariance is 0 even though the distance is ϵ!

This model of log-correlated field can be used to construct many analogues of what we will see for random matrices. For example, consider the function, for $x \in [i2^{-n}, (i + 1)2^{-n})$ corresponding to $v \in D_n$:

$$\mathcal{M}_n(x) = e^{\gamma S_v - \gamma^2 n/2}.$$

Then $\mathcal{M}_n(x)dx$ is a positive measure (called a multiplicative cascade) which is a martingale and therefore converges a.s. It was proved (by Kahane and Peyriere, building on work of Mandelbrodt) that the limit is non-degenerate iff $\gamma < \sqrt{2\log 2} =: x^*$; that limit is called the Gaussian Multiplicative Chaos associated with the BRW; I will not discuss in these lectures the GMC, but see Lambert's course. We will see in a short while the reason for the appearance of the constant x^*.

Exercise 2. Let $\bar{\mathcal{M}}_n := \int_0^1 \mathcal{M}_n(x)dx$ denote the *total mass* of $\mathcal{M}_n(x)dx$. Prove that for $\gamma < x^*/\sqrt{2}$ one has that $\sup_n E\bar{\mathcal{M}}_n^2 < \infty$. Conclude that the martingale $\bar{\mathcal{M}}_n$ has a nontrivial limit, a.s. Hint: you need to use also Kolmogorov's 0-1 law.

Remark 1. The case $\gamma = x^*$ is special, as the martingale $\bar{\mathcal{M}}_n$ can be shown to converge to 0 a.s. However the (non-positive!) martingale measure $\widetilde{\mathcal{M}}_n(x)dx$ with $\widetilde{\mathcal{M}}_n(x) = 2^{-n}e^{\gamma S_v - \gamma^2 n/2}(\frac{\gamma^2}{2}n - \gamma S_v)$ does converge to a non-degenerate *positive* measure, whose total mass is the *derivative martingale*. This will play an important role in the study of the maximum.

1.3 The maximum

We will be interested in the *maximal displacement* of the BRW, defined as

$$M_n = \max_{v \in D_n} S_v \,.$$

Warm up: getting rid of dependence We begin with a warm-up computation. Note that M_n is the maximum over a collection of 2^n variables, that are not independent. Before tackling computations related to M_n, we first consider the same question when those 2^n variables are independent. That is, let $\{\tilde{S}_v\}_{v\in D_n}$ be a collection of i.i.d. random variables, with $\tilde{S}_v$ distributed like S_v, and let $\tilde{M}_n = \max_{v\in D_n} \tilde{S}_v$. We then have the following. (The statement extends to the Non-Gaussian, non-lattice case by using the Bahadur-Rao estimate.)

Theorem 1. *With notation as above, there exists a constant C so that*

$$P(\tilde{M}_n \leqslant \tilde{m}_n + x) \to \exp(-Ce^{-x^* x})\,, \tag{1.3.1}$$

where

$$\tilde{m}_n = nx^* - \frac{1}{2x^*}\log n. \tag{1.3.2}$$

In what follows, we write $A \sim B$ if A/B is bounded above and below by two universal positive constants (that do not depend on n).

Proof. The key is the estimate, valid for $a_n = o(\sqrt{n})$,

$$P(\tilde{S}_v > nx^* - a_n) \sim \frac{C}{\sqrt{n}} \exp(-n(x^* - a_n/n)^2/2)\,, \qquad (1.3.3)$$

which is trivial in the Gaussian case. Therefore,

$$P(\tilde{M}_n \leqslant nx^* - a_n) \sim \left(1 - \frac{C}{2^n\sqrt{n}} e^{x^* a_n + o(1)}\right)^{2^n}$$
$$\sim \exp(-Ce^{x^* a_n + o(1)}/\sqrt{n})\,.$$

Choosing now $a_n = \log n/2x^* - x$, one obtains

$$P(\tilde{M}_n \leqslant m_n + x) \sim \exp(-Ce^{-x^* x + o(1)})\,.$$

□

Remark 2. With some effort, the constant C can also be evaluated to be $1/\sqrt{2\pi}x^*$, but this will not be of interest to us. On the other hand, the constant in front of the $\log n$ term will play an important role in what follows.

Remark 3. Note the very different asymptotics of the right and left tails: the right tail decays exponentially while the left tail is doubly exponential. This is an example of extreme distribution of the Gumbel type.

BRW: the law of large numbers As a further warm up, we will attempt to obtain a law of large numbers for M_n. Recall, from the results of Section 1.3, that $\tilde{M}_n/n \to x^*$. Our goal is to show that the same result holds for M_n.

Theorem 2 (Law of Large Numbers). *We have that*

$$\frac{M_n}{n} \to_{n\to\infty} x^*\,, \quad \textit{almost surely} \qquad (1.3.4)$$

Proof. While we do not really need in what follows, we remark that the almost sure convergence can be deduced from the subadditive ergodic theorem.

The upper bound Let $\bar{Z}_n = \sum_{v\in D_n} \mathbf{1}_{S_v > (1+\epsilon)x^* n}$ count how many particles, at the nth generation, are at location greater than $(1+\epsilon)nx^*$. We apply a first moment method: we have, for any $v \in D_n$, that

$$E\bar{Z}_n = 2^n P(S_v > n(1+\epsilon)x^*) \leqslant 2^n e^{-n((1+\epsilon)x^*)^2/2}\,,$$

where we applied Chebyshev's inequality in the last inequality. Thus,

$$P(M_n > (1+\epsilon)nx^*) \leqslant E\bar{Z}_n \leqslant e^{-c(\epsilon)n}\,.$$

It follows that

$$\limsup_{n\to\infty} \frac{M_n}{n} \leqslant x^*\,, \quad \text{almost surely}\,.$$

The lower bound A natural way to proceed would have been to define

$$\underline{Z}_n = \sum_{v \in D_n} \mathbf{1}_{S_v > (1-\epsilon)x^* n}$$

and to show that with high probability, $\underline{Z}_n \geqslant 1$. Often, one handles this via the second moment method: recall that for any nonegative, integer valued random variable Z,

$$EZ = E(Z\mathbf{1}_{Z \geqslant 1}) \leqslant (EZ^2)^{1/2}(P(Z \geqslant 1))^{1/2}$$

and hence

$$P(Z \geqslant 1) \geqslant \frac{(EZ)^2}{E(Z^2)}. \tag{1.3.5}$$

In the case of independent summands, this would work.

Exercise 3. Check that the vanilla second moment method works for the LLN lower for $\tilde{M}_n$ using $\underline{Z}_n$, while it does not work for M_n.

Because of Exercise 3, we need to reduce correlations. At the level of LLN, a simple method to achieve that is to introduce the event for $v \in D_n$,

$$B_v^\epsilon = \{|S_v(t) - x^*(1-\epsilon)t| \leqslant \epsilon n, t = 1, \ldots, n\}\,.$$

We now recall a basic large deviations result.

Theorem 3 (Varadhan, Mogulskii). *With notation and assumption as above,*

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log P(B_v^\epsilon) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log P(B_v^\epsilon) = -(x^*)^2/2\,.$$

Define now

$$Z_n = \sum_{v \in D_n} \mathbf{1}_{B_v^\epsilon}\,.$$

Exercise 4. Check that the second moment method works with this definition of Z_n, ie $EZ_n \to \infty$ and $EZ_n^2/(EZ_n)^2 \to 1$.

Tightness of the centered maximum We continue to refine results for the BRW, in the spirit of Theorem 1; we will not deal yet with convergence in law, rather, we will deal with finer estimates on EM_n, as follows.

Theorem 4. *With notation and assumption as before, we have*

$$EM_n = nx^* - \frac{3}{2x^*} \log n + O(1)\,. \tag{1.3.6}$$

Further, $(M_n - EM_n)$ is a tight sequence.

Remark 4. It is instructive to compare the logarithmic correction term in (1.3.6) to the independent case, see (1.3.2): the constant 1/2 coming from the Bahadur-Rao estimate (1.3.3) is replaced by 3/2. As we will see, this change is due to extra constraints imposed by the tree structure, and ballot theorems that are close to estimates on Brownian bridges conditioned to stay positive.

Theorem 4 was first proved by Bramson [Br78] in the context of Branching Brownian Motions. The branching random walk case was discussed in [ABR09], who stressed the importance of certain ballot theorems. Roberts [Rob13] significantly simplified Bramson's original proof. The proof we present combines ideas from these sources. To reduce technicalities, we consider only the case of Gaussian increments in the proofs.

Before bringing the proof, we start with some preliminaries related to Brownian motion and random walks with Gaussian increments.

Lemma 1. *Let $\{W_t\}_t$ denote a standard Brownian motion. Then*

$$P(W_t \in dx, W_s \geqslant -1 \text{ for } s \leqslant t) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \left(1 - e^{-(x+2)/2t}\right) dx \,. \quad (1.3.7)$$

Note that the right side in (1.3.7) is of order $(x+2)/t^{3/2}$ for all $x = O(\sqrt{t})$ positive. Further, by Brownian scaling, for $y = O(\sqrt{t})$ positive,

$$P(W_t \in dx, W_s \geqslant -y \text{ for } s \leqslant t) = O\left(\frac{(x+1)(y+1)}{t^{3/2}}\right) . \quad (1.3.8)$$

Proof: This is D. André's reflection principle. Alternatively, the pdf in question is the pdf of a Brownian motion killed at hitting -1, and as such it solves the PDE $u_t = u_{xx}/2\,, u(t,-1) = 0$, with solution $p_t(0,x) - p_t(-2,x)$, where $p_t(x,y)$ is the standard heat kernel. □

Remark: An alternative approach to the proof of Lemma 1 uses the fact that a BM conditioned to remain positive is a Bessel(3) process. This is the approach taken in [Rob13].

We next bring a ballot theorem; for general random walks, this version can be found in [ABR08, Theorem 1].

Theorem 5 (Ballot theorem). *Let X_i be iid random variables of zero mean, finite variance, with $P(X_1 \in (-1/2, 1/2)) > 0$. Define $S_n = \sum_{i=1}^n X_i$. Then, for $0 \leqslant k \leqslant \sqrt{n}$,*

$$P(k \leqslant S_n \leqslant k+1, S_i > 0, 0 < i < n) = \Theta\left(\frac{k+1}{n^{3/2}}\right) , \quad (1.3.9)$$

and the upper bound in (1.3.9) *holds for any $k \geqslant 0$.*

Here, we write that $a_n = \Theta(b_n)$ if there exist constanst $c_1, c_2 > 0$ so that

$$c_1 \leqslant \liminf_{n\to\infty} \frac{a_n}{b_n} \leqslant \limsup_{n\to\infty} \frac{a_n}{b_n} \leqslant c_2 \,.$$

Exercise 5. Show that there exists a constant c_b so that

$$\lim_{x,y\to\infty}\lim_{n\to\infty}\frac{n^{3/2}}{xy}P^x(S_n\in[y,y+1],S_i>0,i=1,\ldots,n)=c_b\,. \tag{1.3.10}$$

A lower bound on the right tail of M_n Fix $y>1$ independent of n and set

$$a_n=x^*n-\frac{3}{2x^*}\log n.$$

For $v\in D_n$, define the event

$$A_v=A_v(y)=\{S_v\in[y+a_n-1,y+a_n],S_v(t)\leqslant a_nt/n+y,t=1,2,\ldots,n\},$$

and set

$$Z_n=\sum_{v\in D_n}\mathbf{1}_{A_v}\,.$$

In deriving a lower bound on EM_n, we first derive a lower bound on the right tail of the distribution of M_n, using a second moment method. For this, we need to compute $P(A_v)$. Recall that we have $(x^*)^2/2=\log 2$. Introduce the new parameter $\lambda_n^*=a_n/n$. Let μ denote the standard Gaussian law on $\mathbb{R}$.

Define a new probability measure Q on $\mathbb{R}$ by

$$\frac{d\mu}{dQ}(x)=e^{-\lambda_n^*x+(\lambda_n^*)^2/2}\,,$$

and with a slight abuse of notation continue to use Q when discussing a random walk whose iid increments are distributed according to Q. Note that in our Gaussian case, Q only modifies the mean of P, not the variance.

We can now write

$$\begin{aligned}P(A_v)&=E_Q(e^{-\lambda_n^*S_v+n(\lambda_n^*)^2/2}\mathbf{1}_{A_v})\\&\geqslant e^{-n[\lambda_n^*(a_n+y)/n-(\lambda_n^*)^2/2]}Q(A_v)\\&=e^{-n((a_n+y)/n)^2/2}Q(\tilde S_v\in[y-1,y],\tilde S_v(t)\geqslant 0,t=1,2,\ldots,n)\,.\end{aligned} \tag{1.3.11}$$

where $\tilde S_v(t)=a_nt/n-S_v(t)$ is a random walk with iid Gaussian increments of variance 1, whose mean vanishes under Q. Thus, $\{\tilde S_v(t)\}_t$ is distributed like $\{S_v(t)\}_t$.

Applying Theorem 5, we get that

$$P(A_v)\geqslant C\frac{y+1}{n^{3/2}}e^{-n((a_n+y)/n)^2/2}\,. \tag{1.3.12}$$

Since

$$((a_n+y)/n)^2=(x^*)^2-2x^*\left(\frac{3}{2x^*}\cdot\frac{\log n}{n}-\frac{y}{n}\right)+O\left(\left(\frac{\log n}{n}\right)^2\right)\,,$$

we conclude that

$$P(A_v) \geqslant C(y+1)2^{-n}e^{-x^* y}\,,$$

and therefore

$$EZ_n = 2^n P(A_v) \geqslant c_1 y e^{-x^* y}\,. \tag{1.3.13}$$

We next need to provide an upper bound on

$$EZ_n^2 = 2^n P(A_v) + \sum_{v \neq w \in D_n} P(A_v \cap A_w) = EZ_n + 2^n \sum_{s=1}^{n} 2^s P(A_v \cap A_{v_s})\,, \tag{1.3.14}$$

where $v_s \in D_n$ and $\rho(v, v_s) = 2s$.

The strategy in computing $P(A_v \cap A_{v_s})$ is to condition on the value of $S_v(n-s)$. More precisely, with a slight abuse of notation, writing $I_{j,s} = a_n(n-s)/n + [-j, -j+1] + y$, we have that

$$\begin{aligned} &P(A_v \cap A_{v_s}) \\ &\leqslant \sum_{j=1}^{\infty} P(S_v(t) \leqslant a_n t/n + y, t = 1, 2, \ldots, n-s, S_v(n-s) \in I_{j,s}) \\ &\quad \times \max_{z \in I_{j,s}} (P(S_v(s) \in [y + a_n - 1, y + a_n], \\ &\qquad\qquad S_v(t) \leqslant a_n(n-s+t)/n, t = 1, 2, \ldots, s | S_v(0) = z))^2\,. \end{aligned} \tag{1.3.15}$$

Repeating the computations leading to (1.3.12) (using time reversibility of the random walk) we conclude that

$$P(A_v \cap A_{v_s}) \leqslant \sum_{j=1}^{\infty} \frac{j^3(y+1)}{s^3(n-s)^{3/2}} e^{-(j+y)x^*} n^{3(n+s)/2n} 2^{-(n+s)}\,. \tag{1.3.16}$$

Substituting in (1.3.14) and (1.3.15), and performing the summation over j first and then over s, we conclude that $EZ_n^2 \leqslant cEZ_n$, and therefore, using again (1.3.5),

$$P(M_n \geqslant a_n - 1) \geqslant P(Z_n \geqslant 1) \geqslant cEZ_n \geqslant c_0(y+1)e^{-I'(x^*)y} = c_0(y+1)e^{-x^* y}\,. \tag{1.3.17}$$

This completes the evaluation of a lower bound on the right tail of the law of M_n.

An upper bound on the right tail of M_n A subtle point in obtaining upper bounds is that the first moment method does not work directly - in the first moment one cannot distinguish between the BRW and independent random walks, and the displacement for these has a different logarithmic corrections (the maximum of 2^n independent particles is larger).

To overcome this, note the following: a difference between the two scenarios is that at intermediate times $0 < t < n$, there are only 2^t particles in the

BRW setup while there are 2^n particles in the independent case treated in Section 1.3. Applying the first moment argument at time t shows that there cannot be any BRW particle at time t which is larger than $x^*t + C\log n$, while this constraint disappears in the independent case. One thus expect that imposing this constraint in the BRW setup (and thus, pick up an extra $1/n$ factor from the ballot theorem 5) will modify the correction term.

Carrying out this program thus involves two steps: in the first, we consider an upper bound on the number of particles that never cross a barrier reflecting the above mentioned constraint. In the second step, we show that with high probability, no particle crosses the barrier. The approach we take combines arguments from [Rob13] and [ABR09]; both papers build on Bramson's original argument.

Turning to the actual proof, fix a large constant $\kappa > 0$, fix $y > 0$, and define the function

$$h(t) = \begin{cases} \kappa \log t, & 1 \leqslant t \leqslant n/2 \\ \kappa \log(n-t+1), & n/2 < t \leqslant n\,. \end{cases} \tag{1.3.18}$$

Recall the definition $a_n = x^*n - \frac{3}{2x^*}\log n$ and let

$$\tau(v) = \min\{t > 0 : S_v(t) \geqslant a_n t/n + h(t) + y - 1\} \wedge n\,,$$

and $\tau = \min_{v \in D_n} \tau(v)$. (In words, τ is the first time in which there is a particle that goes above the line $a_n t/n + h(t) + y$.)

Introduce the events

$$B_v = \{\tau \geqslant n, S_v \in [y + a_n - 1, y + a_n]\}$$

and define $Y_n = \sum_{v \in D_n} \mathbf{1}_{B_v}$. Repeating arguments as we already saw (with a slightly modified barrier, which is dealt with by a k-dependent change of measure), one obtains the following.

Lemma 2. *There exists a constant c_2 independent of y so that*

$$P(B_v) \leqslant c_2(y+1)e^{-x^* y}2^{-n}\,. \tag{1.3.19}$$

We need to consider next the possibility that $\tau = t < n$. Assuming that κ is large enough ($\kappa > 3/2x^*$ will do), an application of the lower bound (1.3.17) to the descendants of the parent of the particle v with $\tau_v < n$ reveals that for some constant c_3 independent of y,

$$E[Y_n | \tau < n] \geqslant c_3\,.$$

(Recall that $Y_n = \sum_{v \in D_n} \mathbf{1}_{B_v}$.) We conclude that

$$P(\tau < n) \leqslant \frac{E(Y_n)P(\tau < n)}{E(Y_n \mathbf{1}_{\tau < n})} = \frac{EY_n}{E(Y_n | \tau < n)} \leqslant cEY_n\,. \tag{1.3.20}$$

One concludes from this and Lemma 2 that

$$P(M_n \geqslant a_n + y) \leqslant P(\tau < n) + EY_n \leqslant c_5(y+1)e^{-x^* y}\,. \tag{1.3.21}$$

In particular, this also implies that

$$EM_n \leqslant x^* n - \frac{3}{2x^*}\log n + O(1)\,. \tag{1.3.22}$$

Remark 5. An alternative approach to the argument in (1.3.20), which is more in line with Bramson's original proof, is as follows. Note that

$$P(\tau \leqslant n - n^{\kappa'}) \leqslant \sum_{i=1}^{n-n^{\kappa'}} 2^i P(S_n(i) \geqslant a_n i/n + h(i) + y) \leqslant Ce^{-x^* y}\,,$$

where κ' can be taken so that $\kappa' \to_{\kappa\to\infty} 0$, and in particular for κ large we can have $\kappa' < 1$. Assume now κ large enough so that $\kappa' \leqslant 1/2$. For $t \geqslant n - n^{1/2}$, one repeats the steps in Lemma 2 as follows. Let N_t be the number of vertices $w \in D_t$ (out of 2^t) whose path $S_w(s)$ crosses the barrier $(a_n s/n + h(s) + y - 1)$ at time $s = t$. We have

$$P(\tau = t) \leqslant EN_t \leqslant c(y+1)e^{-x^* yt/n}\frac{1}{(n-t)^{c_1\kappa - c_2}}$$

for appropriate constants c_1, c_2. Taking κ large enough ensures that

$$\sum_{t=n-n^{1/2}}^{n} EN_t \leqslant c(y+1)e^{-x^* y}\,.$$

Combining the last two displays leads to the same estimate as in the right side of (1.3.21), and hence to (1.3.22).

We finally prove a complementary lower bound on the expectation. Recall, see (1.3.17), that for any $y > 0$,

$$P(M_n \geqslant a_n(y)) \geqslant c(y+1)e^{-x^* y}\,,$$

where $a_n(y) = a_n + y$. In order to have a lower bound on EM_n that complements (1.3.22), we need only show that

$$\lim_{z\to-\infty}\limsup_{n\to\infty} P(M_n \leqslant a_n(z)) = 0\,. \tag{1.3.23}$$

Toward this end, fix $\ell > 0$ integer, and note that by the first moment argument used in the proof of the LLN (Theorem 2 applied to $\max_{w\in D_\ell}(-S_w)$), there exist positive constants c, c' so that

$$P(\min_{w\in D_\ell}(S_w) \leqslant -c\ell) \leqslant e^{-c'\ell}\,.$$

On the other hand, for each $v \in D_n$, let $w(v) \in D_\ell$ be the ancestor of v in generation ℓ. We then have, by independence,

$$P(M_n \leqslant -c\ell + (n-\ell)x^* - \frac{3}{2x^*}\log(n-\ell)) \leqslant (1-c_0)^{2^\ell} + e^{-c'\ell},$$

where c_0 is as in (1.3.17). This implies (1.3.23). Together with (1.3.22), this completes the proof of Theorem 4. □

Remark 6. The idea of using curved boundaries also helps in the second moment argument: the arguments we showed can help in showing that

$$P(S_n(i) \geqslant a_n i/n - \bar{h}(i) - y | \tau \geqslant n, S_n \in [y + a_n - 1, y + a_n]) = o_y(1). \quad (1.3.24)$$

That is, one can pass from an upward slopping "banana" to a downward sloping one without cost. This helps in the proof of the lower bound, as one needs less precision in the estimates.

Exercise 6. Prove (1.3.24).

Exercise 7. a) Suppose that the increments X_e are not Gaussian but satisfy that with $e = (v, v+1)$,

$$E(e^{\theta X_e}) = e^{\theta^2/2(1+O(e^{-|v|^\alpha}))}. \quad (1.3.25)$$

Check that the estimates in this section still apply.
b) Check that if $X_e = \sum_{j=2^k}^{2^{k+1}-1} \frac{W_j}{\sqrt{j \log 2}}$, where the W_j are independent, mean 0 and variance 1, and possess a uniformly bounded exponential moment, then (1.3.25) holds.

1.4 Convergence of maximum and Gumbel limit law

We begin with a lemma, whose proof we only sketch.

Lemma 3. *There exists a constant $\bar{c}$ such that*

$$\lim_{y\to\infty} \limsup_{n\to\infty} \frac{e^{x^* y}}{y} P(M_n \geqslant m_n + y) = \lim_{y\to\infty} \liminf_{n\to\infty} \frac{e^{x^* y}}{y} P(M_n \geqslant m_n + y) = \bar{c}. \quad (1.4.26)$$

Note that the lemma is consistent with the upper and lower estimates on the right tail that we already derived. The main issue here is the convergence.
Proof (sketch): The key new idea in the proof is a variance reduction step. To implement it, fix k (which will be taken function of y, going to infinity but so that $k << y$) and define, for any $v \in D_n$,

$$W_{v,k} = \max_{w \in D_k(v)} (S_w - S_v).$$

Here, $D_k(v)$ denote the vertices in D_{n+k} that are descendants of $v \in D_n$. Now,

$$P(M_{n+k} > m_{n+k} + y) = P(\max_{v \in D_n}(S_v \geqslant m_n + (x^* k - W_{v,k} + y)))\,.$$

For each $v \in D_n$, we consider the event

$$A_v(n) = \{S_v(t) \leqslant t m_n/n + y, t = 1, \ldots, n; S_v \geqslant m_n + (x^* k - W_{v,k} + y)\}\,.$$

Note that the event in $A_v(n)$ forces to have $W_{v,k} \geqslant x^* k + (m_n - S_v + y) \geqslant x^* k$, which (for k large) is an event of small probability. Now, one employs a curve $h(t)$ as described when deriving an upper bound on the right tail of M_n to show that

$$P(M_{n+k} > m_{n+k} + y) = (1 + o_y(1))P(\cup_{v \in D_n} A_v(n))\,.$$

Next, using Exercise 5, one shows that

$$\lim_{y\to\infty} \lim_{n\to\infty} \frac{e^{x^* y}}{y} P(A_v(n)) 2^{-n} = \bar{c}\,,$$

for some constant $\bar{c}$. This is very similar to computations we already did.

Finally, note that conditionally on $\mathcal{F}_n = \sigma(S_v, v \in D_j, j \leqslant n)$, the events $\{W_{v,k} \geqslant x^* k + (m_n - S_v + y)\}_{v \in D_n}$ are independent. This introduces enough decorelation so that even when $v, w \in D_n$ are neighbors on the tree, one gets that

$$P(A_v(n) \cap A_w(n)) \leqslant o_y(1) P(A_v(n))\,.$$

Because of that, defining $Z_n = \sum_{v \in D_n} 1_{A_v(n)}$, one obtains that $EZ_n^2 \leqslant (1 + o_y(1))EZ_n + CEZ_n^2$ for some constant C and therefore, using that $\limsup_{n\to\infty} EZ_n \to_{y\to\infty} 0$, one has

$$EZ_n \geqslant P(\cup_{v \in D_n} A_v(n)) \geqslant \frac{(EZ_n)^2}{EZ_n^2} \geqslant \frac{(EZ_n)^2}{EZ_n(1 + o_y(1))} \geqslant EZ_n(1 - o_y(1))\,.$$

Combining these three facts gives the lemma. □

We now finally are ready to state the following.

Theorem 6. *There exists a random variable Θ such that*

$$\lim_{n\to\infty} P(M_n \leqslant m_n + y) = E(e^{-\Theta e^{-\lambda^* y}})\,. \tag{1.4.27}$$

Thus, the law of $M_n - m_n$ converges to the law of a randomly shifted Gumbel distribution.

Remark: In fact, the proof we present will show that the random variable Θ is the limit in distribution of a sequence of random variables Θ_k. In reality, that sequence forms a martingale (the so called derivative martingale, see

Remark 1) with respect to $\mathcal{F}_k$, and the convergence is a.s.. We will neither need nor use that fact. For a proof based on the derivative martingale convergence, see Lalley and Sellke [LS87] for the BBM case and Aïdekon [Aïd13] for the BRW case.

Proof (sketch): This time, we cut the tree at a fixed distance k from the root. Use that for n large, $\log(n+k) = \log(n) + O(1/n)$. Write

$$\begin{aligned} P(M_{n+k} \leqslant m_{n+k} + y) &= E(\prod_{v \in D_k} 1_{(S_v + W_{v,k}(n)) \leqslant m_n + x^* k + y} \\ &\sim E(\prod_{v \in D_k} P(W_{v,k} \leqslant m_n + (x^* k - S_v) + y | S_v) \\ &\sim \epsilon(k) + E(\prod_{v \in D_k} \left(1 - \bar{c}(x^* k - S_v + y) e^{-\lambda^*(x^* k - S_v + y)}\right) , \end{aligned}$$

where the symbol $a \sim b$ means that $a/b \to_{n \to \infty} 1$, and we used that with high probability $(1 - \epsilon(k))$, $x^* k - S_v \geqslant x^*_k - S^*_k$ is of order $\log k$ and therefore we could apply Lemma 3 in the last equivalence. Fixing $\Theta_k = \bar{c} \sum_{v \in D_k} (x^* k - S_v) e^{-\lambda^*(x^* k - S_v)}$ and using that y is fixed while k is large, we conclude that

$$P(M_{n+k} \leqslant m_{n+k} + y) \sim \epsilon'(k) + E(e^{-\Theta_k e^{-\lambda^* y}}) .$$

Since the right side does not depend on n, the convergence of the left side follows by taking $n \to \infty$ and then taking k large. Finally, the convergence also implies that the moment generating function of Θ_k converges, which in terms implies the convergence in distribution of Θ_k. □

Extremal process We give a description of (a weak form of) a theorem due to [ABK11] and [ABBS13] in the Branching Brownian motion case and to [Ma17] in the (not necessarily Gaussian) BRW case, describing the distribution of the point process $\eta_n = \sum_{v \in D_n} \delta_{S_v - m_n}$. Our proof will follow the approach of Biskup and Louidor [BL13], and is tailored to the Gaussian setup we are considering.

We begin with a preliminary lemma. For a fixed constant R, set $\mathcal{M}_n(R) = \{v \in D_n : S_v > m_n - R\}$.

Lemma 4. *There exist functions $r(R) \to_{R \to \infty} \infty$ and $\epsilon(R) \to_{R \to \infty} 0$ so that*

$$\limsup_{n \to \infty} P(\exists u, v \in \mathcal{M}_n(R) : r(R) < \rho(u,v)/2 < n - r(D)) \leqslant \epsilon(R) . \quad (1.4.28)$$

The proof is immediate from the second moment computations we did; we omit details.

Fix now R (eventually, we will take $R \to \infty$ slowly with n) and define the *thinned* point process $\eta^s_n = \sum_{v \in D_n, S_v = \max_{w : d_T(v,w) \leqslant R} S_w} \delta_{S_v - m_n}$. In words, η^s_n is the point process obtained by only keeping points that are leaders of their respective "clan", of depth R.

Theorem 7. *(a) The process η_n^s converges, as $n \to \infty$, to a random shift of a Poisson Point Process (PPP) of intensity $Ce^{-\lambda^* x}$, denoted η^s.*

(b) The process η_n converges, as $n \to \infty$, to a decorated version of η^s, which is obtained by replacing each point in η_s by a random cluster of points, independently, shifted around z.

A description of the decoration process is also available. We however will not bother with it. Instead, we will only sketch the proof of part (a) of Theorem 7.

Before the proof, we state a general result concerning invariant point processes, due to Liggett [Li78]. The setup of Liggett's theorem (narrowed to our needs; the general version replaces $\mathbb{R}$ by a locally compact second countable topological space) is a point process η on $\mathbb{R}$ (i.e., a random, integer valued measure on $\mathbb{R}$ which is finite a.s. on each compact), with each particle evolving individually according to a Markov kernel Q. For m a locally finite positive measure on $\mathbb{R}$, let μ_m denote the PPP of intensity m. For a random measure M on $\mathbb{R}$, we set $\bar{\mu}_M = \int \mu_m P(M \in dm)$ (a more suggestive notation would be $\bar{\mu}_M = E\mu_M$ where μ_M is, conditioned on M, a PPP of intensity M). We say that the law of a point process is invariant for Q if it does not change when each particle makes independently a move according to the Markov kernel Q.

One has the following. Througout, we assume that $Q^n(x, K) \to_{n\to\infty} 0$ uniformly in x for each fixed $K \subset\subset \mathbb{R}$.

Theorem 8 (Liggett [Li78]).

(a) $\bar{\mu}_M$ is invariant for Q iff $MQ = M$ in distribution.
(b) Every invariant probability measure is of the form $\bar{\mu}_M$ for some M.
(c) The extremal invariant probability measures for the point process are of the form μ_m with m satisfying $mQ = m$ iff $MQ = M$ in distribution implies $MQ = M$ a.s.
(d) In the special case where $Q(x, dy) = g(y-x)dy$ where g is a density function with finite exponential moments, condition (c) holds and all extreme invariant m are of the form $m(dx) = Ce^{-C'x}dx$, with C' depending on C and g.

(Part (d) of the theorem is an application of the Choquet-Deny theorem that characterizes the exponential distribution).

Proof of Theorem 7(a) (sketch): We write $\eta_n = \eta_n(\{S_v\})$ to emphasize that η_n depends on the Gaussian field $\{S_v\}_{v\in D_n}$. Note that due to the Gaussian structure,

$$\eta_n \stackrel{d}{=} \eta_n(\{\sqrt{1-1/n}S_v\} + \{\sqrt{1/n}S'_v\}), \qquad (1.4.29)$$

where $\{S'_v\}$ is an independent copy of $\{S_v\}$ and the equality is in distribution. Now, $\{\sqrt{1/n}S'_v\}$ is a Gaussian field with variance *of order* 1, while $\sqrt{1-1/n}S_v = S_v - \frac{1}{2n}S_v + o(1)$.

Note that for any fixed $v \in D_n$, we have that $\max_{w:d_T(v,w)\leqslant R}(S'_w - S'_v)/\sqrt{n} \leqslant \delta(R)$ with probability going to 1 as $n \to \infty$, for an appropriate function $\delta(R) \to_{R\to\infty} 0$. By a diagonalization argument, one can then choose $R = R(n)$ so that

$$\max_{v,w\in D_n:S_v>m_n-R,d_T(v,w)\leqslant R}(S'_w - S'_v)/\sqrt{n} \leqslant \delta(R)$$

with probability going to 1 as $n \to \infty$.

Consider the right side of (1.4.29) as a (random) transformation on η_n; when restricting attention to the interval $(-R(n), \infty)$, which a.s. contains only finitely many points (first moment!), the transformation, with probability approaching 1 does the following:

– Replaces each point S_v by $S_v - x^*$.
– Adds to each *clan* an *independent* centered Gaussian random variable of variance 1.
– Adds a further small error of order $\delta(R_n)$

When thinning, one notes that the same transformation applies to the thinned process. Thus, any weak limit of the thinned process is invariant under the transformation that adds to each point an independent normal of mean $-x^*$. By the last point of Liggett's theorem, we conclude that any limit point of η_n^s is a random mixture of PPP with exponential intensity. The convergence of the maximum then determines both the exponent in the exponential intensity (must be λ^*) as well as the mixture (determined by the maximum). This completes the proof. □

2 Lecture II: The CβE

The goal of this section is to provide us with a road-map for reading [PZ25]. Due to the length and technical complexity of the latter, I chose to emphasize the main ideas and the parallels (and differences!) with the treatment of the BRW. Most proofs, therefore, will only be sketched.

The Circular-β ensemble (CβE) is a distribution on n points $(e^{i\omega_1}, e^{i\omega_2}, \ldots, e^{i\omega_n})$ on the unit circle with a joint density given by

$$\frac{1}{Z_{n,\beta}} \prod_{1\leqslant j<k\leqslant n} |e^{i\omega_j} - e^{i\omega_k}|^\beta d\omega_1 \cdots d\omega_n. \tag{2.1.1}$$

In the special case of $\beta = 2$ this is the joint distribution of eigenvalues of a Haar-distributed unitary random matrix. The characteristic polynomial $X_n(z) := \prod_{j=1}^n (1 - e^{i\omega_j} z)$ of the CβE has attracted a considerable interest, for its connections to the theories of logarithmically–correlated fields and (when $\beta = 2$) analytic number theory.

A particular quantity of interest is $M_n := \max_{|z|=1} \log|X_n(z)|$. Let

$$m_n = \log n - \tfrac{3}{4} \log\log n. \tag{2.1.2}$$

The random matrix part of the Fyodorov–Hiary–Keating conjecture [FHK12b] states that in the special case that $\beta = 2$, $M_n - m_n$ converges in distribution towards a limiting random variables R_2, with

$$P(R_2 \in dx) = 4e^{2x} K_0(2e^x) dx. \tag{2.1.3}$$

It was later observed in [SZ15] that the probability density in (2.1.3) is the law of the sum of two independent Gumbel random variables.

For general $\beta > 0$, an important step forward was obtained by [CMN18], who proved that $M_n - \sqrt{2/\beta} m_n$ is tight. The goal of this lecture is to sketch the proof of the following.

Theorem 9 ([PZ25]). *The sequence of random variables $M_N - \sqrt{2/\beta} m_N$ converges in distribution to a random variables R_β. Further,*

$$R_\beta = C_\beta + G_\beta + \frac{1}{\sqrt{2\beta}} \log(\mathscr{B}_\infty(\beta)), \tag{2.1.4}$$

where C_β is an (implicit) constant, G_β is Gumbel distributed with parameter $1/\sqrt{2\beta}$, and $\mathscr{B}_\infty(\beta)$ is a random variable that is independent of G_β.

Remark 7. [PZ25] give a description of $\mathscr{B}_\infty(\beta)$ as the total mass of a certain *derivative martingale*. For a specific log-correlated field on the circle, [Rem20] computes the law of the total mass of the associated GMC and confirms the Fyodorov-Bouchaud prediction [FB08] for it. It is possible (and even anticipated, especially in light of [LN24], see Lambert's talks) but not yet proved, that the distribution of $\mathscr{B}_\infty$ is also Gumbel. If true (even if only for $\beta = 2$), Theorem 9 would then yield a proof of the random matrix side of the Fyodorov-Hiary-Keating conjecture [FHK12b].

Theorem 9 is a consequence of a more general result, which gives the convergence of the distance between certain marked point processes built from a sequence of orthogonal polynomials, and a sequence of (n-independent) decorated Poisson point process. This general result also applies to the imaginary part of $\log X_n(z)$ (and thus, allows for control on maximal fluctuation of eigenvalue count on intervals).

2.2 OPUC preliminaries and formulation of main results

A major advance in the study of M_n was achieved in [CMN18], who used the Orthogonal Polynomials on the Unit Circle (OPUC) representation of the CβE measure due to [KN04]; we refer to Lambert's course and to [Sim04] for an encyclopedic account of the OPUC theory. Let $\{\gamma_k\}$ be independent, complex, rotationally invariant random variables for which

$|\gamma_k|^2 = \text{Beta}(1, \beta(k+1)/2)$, that is with density on $[0,1]$ proportional to $(1-x)^{\beta(k+1)/2-1}$. The *Szegő* recurrence is, for all $k \geqslant 0$,

$$\begin{pmatrix} \Phi_{k+1}(z) \\ \Phi^*_{k+1}(z) \end{pmatrix} := \begin{pmatrix} z & -\overline{\gamma_k} \\ -\gamma_k z & 1 \end{pmatrix} \begin{pmatrix} \Phi_k(z) \\ \Phi^*_k(z) \end{pmatrix}, \quad \left\{ \begin{matrix} \Phi_0(z) \equiv 1, \\ \Phi^*_k(z) = z^k \overline{\Phi_k(1/\overline{z})}. \end{matrix} \right\} \tag{2.2.5}$$

where Φ^*_k and Φ_k are polynomials of degree at most k. Define in terms of these coefficients the *Prüfer phases*

$$\Psi_{k+1}(\theta) = \Psi_k(\theta) + \theta - 2\Im\left(\log(1 - \gamma_k e^{i\Psi_k(\theta)})\right), \quad \Psi_0(\theta) = \theta, \tag{2.2.6}$$

where here and below we take the principal branch of the logarithm with discontinuity along the negative real line. Then, $\Psi_k(\cdot)$ may be identified as a continuous version of the logarithm of $\theta \mapsto \frac{1}{i}\log(e^{i\theta}\frac{\Phi_k(e^{i\theta})}{\Phi^*_k(e^{i\theta})})$. Let α be a uniformly distributed element of the unit circle, independent of $\{\gamma_k : k \geqslant 0\}$, and set for any $\theta \in \mathbb{R}$,

$$X_n(e^{i\theta}) := \Phi^*_{n-1}(e^{i\theta}) - \alpha e^{i\theta}\Phi_{n-1}(e^{i\theta}) = \Phi^*_{n-1}(e^{i\theta})\big(1 - \alpha e^{i\Psi_{n-1}(\theta)}\big). \tag{2.2.7}$$

Then X_n has the law of the characteristic polynomial (as process in θ). Note that this means that to understand $|X_n(e^{i\theta})|$, we need to understand both the real and imaginary part of $\log \Phi^*_{n-1}$.

The polynomials $\{\Phi^*_k\}$ satisfy the recurrence

$$\log \Phi^*_{k+1}(e^{i\theta}) = \log \Phi^*_k(e^{i\theta}) + \log(1 - \gamma_k e^{i\Psi_k(\theta)}), \quad \Phi^*_0(e^{i\theta}) = 1. \tag{2.2.8}$$

We also recall the *relative Prüfer phase* [CMN18, Lemma 2.4] given by the recurrence

$$\psi_{k+1}(\theta) = \psi_k(\theta) + \theta - 2\Im\left(\log(1 - \gamma_k e^{i\psi_k(\theta)}) - \log(1-\gamma_k)\right), \quad \psi_0(\theta) = \theta. \tag{2.2.9}$$

In law $\{\psi_k(\theta) : k \in \mathbb{N}, \theta \in [0, 2\pi]\}$ is equal to $\{\Psi_k(\theta) - \Psi_k(0) : k \in \mathbb{N}, \theta \in [0, 2\pi]\}$.

We will be interested in the extreme values of the fluctuations of real and imaginary parts of $\log \Phi^*$, for which reason we will formulate our results in terms of the recurrence

$$\varphi_{k+1}(\theta) = \varphi_k(\theta) + 2\Re\big\{\sigma\left(\log(1 - \gamma_k e^{i\Psi_k(\theta)})\right)\big\}, \quad \varphi_0(\theta) = 0, \tag{2.2.10}$$

where σ is one of $\{1, \pm i\}$. Then, for $\sigma = 1$, $\varphi_k(\cdot) = 2\Re \log \Phi^*_k(\cdot)$ while for $\sigma = i$, $\varphi_k(\cdot) = -2\Im \log \Phi^*_k(\cdot)$. To alleviate notation, we will in these notes mostly consider the real part, i.e. $\sigma = 1$, and explain at the end how results are transfered to the characteristic polynomial.

The following result, which complements Theorem 9, yields the convergence in law of the centered maxima of φ_n.

Theorem 10. *For any* $\sigma \in \{1, \pm i\}$ *, the centered maximum* $\max_{\theta\in[0,2\pi]} \varphi_n(\theta) - \sqrt{8/\beta}\, m_n$ *converges in law to a randomly shifted Gumbel of parameter* $\sqrt{2/\beta}$*. In the notation of Theorem 9, the limit is* $C^\sigma_\beta + 2G^\sigma_\beta + \sqrt{2/\beta}\log \mathscr{B}_\infty$*, where* G^σ_β *has the same law as* G_β *and* C^σ_β *is an (implicit) constant.*

The following result appears in [CMN18], with m_n as in (2.1.2). We use N because, to link with Lecture 1, we will later use $n = \lfloor \log_2 N \rfloor$.

Theorem 11. *For any* $\sigma \in \{1, \pm i\}$ *, with* $M_N := \max_{\theta\in[0,2\pi]} \varphi_N(\theta)$*, the centered maximum* $M_N - \sqrt{8/\beta}\, m_N$ *is tight. The same holds for the real and imaginary parts of the logarithm of the characteristic polynomial.*

Before discussing the proofs of Theorems 9 and 10 on convergence, we wish to explain the proof of Theorem 11 and relate it to what we saw in Lecture I.

2.3 OPUC as log-correlated fields

In a first step, observe that instead of discussing $\sup_\theta \varphi_N(\theta)$, one can discuss $\sup_{\theta\in U_N} \varphi_N(\theta)$, where U_N is an equispaced net in $[0, 2\pi]$ of cardinality mN, with m fixed and large. This follows from an interpolation lemma, which slightly expands and quantifies a result in [CMN18], and the fact that φ_N is a trigonometric polynomial of degree N.

Lemma 5. *For any polynomial* Q *of degree* $k \geqslant 1$*, and any natural number* $m \geqslant 2$*,*

$$\max_{|z|=1} |Q(z)|^2 \leqslant \frac{m}{m-1} \cdot \max_{\omega:\omega^{2mk}=1} |Q(\omega)|^2.$$

Furthermore, if for any $b > 0$ *we partition the* $(2mk)$*-th roots of unity into* $\mathcal{N}$ *and* $\mathcal{F}$ *so that* $\mathcal{N}$ *are all those roots of unity* ω *so that* $|\omega - 1| \leqslant \frac{2b}{k}$*, then there is an absolute constant* $C > 0$ *so that*

$$\max_{\substack{|z-1|\leqslant \frac{b}{k},\\ |z|=1}} |Q(z)|^2 \leqslant \frac{m}{m-1} \cdot \max_{\omega\in\mathcal{N}} |Q(\omega)|^2 + \frac{C}{b(m-1)} \cdot \max_{\omega\in\mathcal{F}} |Q(\omega)|^2 \quad \text{and}$$

$$\min_{\substack{|z-1|\leqslant \frac{b}{k},\\ |z|=1}} |Q(z)|^2 \geqslant \frac{m}{m-1} \cdot \min_{\omega\in\mathcal{N}} |Q(\omega)|^2 - \left(1 + \frac{C}{b}\right) \frac{1}{(m-1)} \cdot \max_{\omega:\omega^{2mk}=1} |Q(\omega)|^2.$$

I do not discuss the (classical) proof and refer instead to [PZ25, Theorem 8.2].

We now begin to explain the hierarchical structure of the processes $\varphi_k(\theta)$. Consider first (2.2.10) for a *fixed* θ. Note that the increment $\varphi_{k+1}(\theta) - \varphi_k(\theta)$ is independent of $\varphi_j(\theta))_{j\leqslant k, \theta\in[0,2\pi]}$, has mean 0 (if $\sigma = 1$) and mean θ (if $\sigma = i$), and variance c_β/k. Further, one has the following.

Exercise 8. Check that $\varphi_{2^{k+1}}(\theta) - \varphi_{2^k}(\theta)$, up to the centering $i2^k\theta$ and scaling by $\sqrt{c_\beta}$, satisfies the assumptions of Exercise 7.

From Exercise 8, we see that $\varphi_{2^k}(\theta)$ behaves essentially like a Gaussian random walk (in k). We would like to proceed as in the case of BRW, and introduce barriers. Note that in the introduction of barriers for BRW, we used the fact that the cardinality of vertices at depth k was 2^k; Here, this is not the case: if we are interested in M_N with $N = 2^n$, there are still (essentially) 2^n points at level $k < n$, and not 2^k. This destroys the option of using a crude union bound in order to introduce the barrier. Instead, we have the following.

Lemma 6 (MetaTheorem). *The processs $\varphi_{2^k}(\theta)$ is continuous at scale 2^{-k}.*

We cannot provide in these notes a proof of Lemma 6, but we can explain why one can expect it to be correct: indeed, we see that the means of $E[\varphi_{2^k}(\theta) - \varphi_{2^k}(\theta')] = \theta 2^k$ if $\sigma = i$, and it is not hard to evaluate the variance of this difference and show it is of order 1. This of course is not enough, as we need results at the level of large deviations, and a big technical difficulty is to introduce good enough approximations that are strong enough at the tail.

Equipped with Lemma 6, Exercise 8, and Lemma 5, one can obtain the analogue of the upper tail (1.3.21), and prove the upper tail upper bound part of Theorem 11.

To obtain a right tail lower bound, we again need to apply a second moment method with barrier. For that, a first step is the claim, stated here for $\sigma = 1$:

$$E(\phi_{2^k}(\theta)\phi_{2^k}(\theta')) = -\log(|\theta - \theta'| \vee 2^{-k}) + O(1). \tag{2.3.11}$$

We have already explained why $\varphi_{2^k}(\theta)$ and $\varphi_{2^k}(\theta')$ are close together if $|\theta - \theta'| \sim 2^{-k}$. Suppose now that $|\theta - \theta'| = 2^{-\ell}$ with $\ell < k$. Then, we need to understand why $(\varphi_{2^k} - \varphi_{2^\ell})(\theta)$ and $(\varphi_{2^k} - \varphi_{2^\ell})(\theta')$ are essentially independent. For that, note that the *Prüfer phases* satisfy

$$\Psi_j(\theta) \sim \Psi_j(\theta') + (\theta - \theta')j + O(\sqrt{\log j}). \tag{2.3.12}$$

Now these *Prüfer phases* enter into the evolution of φ_k (see (2.2.10)) as

$$\varphi_{k+1}(\theta) - \varphi_k(\theta) \sim |\gamma_k| \cos(\Psi_k(\theta) + \alpha_k), \tag{2.3.13}$$

with α_k uniform on $[0, 1]$ and independent of φ_k.

Exercise 9. Assume that in (2.3.13), one replaces $\sim$ by an equality and removes the error term in (2.3.12). Show that this implies (2.3.11).

The actual proof, of course, has to justify these approximation steps, at the level of large deviations. Once this is done, the same second moment method that worked for BRW can be applied.

2.4 Back to convergence: the landscape

There are several places where the description above fails to be precise enough to yield convergence:

1. Locally, ie at microscopic scales (meaning that $|\theta - \theta'| \sim O(1/N) = O(2^{-n})$), the polynomial $\varphi_N(\theta)$ is very different from the piecewise constant field obtained by embedding a BRW on the interval $[0,1]$.
2. Similarly, the macroscopic description of the beginning of the recursions is very different from a BRW. In particular, there is no exact independence of increments as in the BRW case.
3. The log-determinant is not exactly φ_N, and in fact one needs to use an extra randomization and the "shape" of φ_N in a neighborhood of near maxima.

Points 1,2 already appeared in the study of the maximum of the two dimensional GFF [BDZ16], and its generalization to log-correlated Gaussian fields [DRZ17]. In a nutshell, the approach of [BDZ16] used a spatial Markov property that gave independence of increments (which is not present here), and the general [DRZ17] used in a strong way the Gaussianity of the field, which also fails here. We will see however that the ideas from these references, together with some new elements, are behind the proof of Theorem 9.

We begin with the description of the random shift. Define the random measure and its total mass

$$\mathscr{D}_k(\theta)d\theta := \tfrac{1}{2\pi} e^{\sqrt{\frac{\beta}{2}}\varphi_k(\theta) - \log k}\big(\sqrt{2}\log k - \sqrt{\tfrac{\beta}{4}}\varphi_k(\theta)\big)_+ d\theta, \mathscr{B}_k := \int_0^{2\pi} \mathscr{D}_k(\theta)d\theta. \tag{2.4.14}$$

This is not exactly a martingale, due to the truncation, but the truncation becomes meaningless for large k. We equip the space of finite measures with the weak-* topology, and then we show that this measure converges almost surely.

Theorem 12. *For any $\sigma \in \{1, \pm i\}$ and any $\beta > 0$, there is an almost surely finite random variable $\mathscr{B}_\infty$ and an almost surely finite, nonatomic random measure $\mathscr{D}_\infty$ so that*

$$\mathscr{D}_{2^j} d\theta \xrightarrow[j\to\infty]{\text{a.s.}} \mathscr{D}_\infty \quad \text{and} \quad \mathscr{B}_{2^j} \xrightarrow[j\to\infty]{\text{a.s.}} \mathscr{B}_\infty.$$

Furthermore for any $\epsilon > 0$ there is a compact $K \subset (0,\infty)$ so that with

$$\chi(\theta) = \mathbf{1}\left\{(\sqrt{2}\log k - \sqrt{\tfrac{\beta}{4}}\varphi_k(\theta))/\sqrt{\log k} \notin K,\right\}$$

it holds that for any $k \in \mathbb{N}$,

$$\mathbb{P}\left(\int_0^{2\pi} e^{\sqrt{\frac{\beta}{2}}\varphi_k(\theta) - \log k}\big|\sqrt{2}\log k - \sqrt{\tfrac{\beta}{4}}\varphi_k(\theta)\big|\chi(\theta)d\theta > \epsilon\right) < \epsilon. \tag{2.4.15}$$

The meaning of (2.4.15) is that only angles θ with $\varphi_k(\theta)$ near the maximal value contribute to $\mathscr{D}_\infty$.

Theorem 12 is very useful in carrying out the analogue of the computation in the proof of Theorem 6, thus addressing effectively point 2 above. Dealing with point 1 is however more delicate. We turn to describing this.

We introduce parameters $\{k_p : p \in \mathbb{N}\}$ which will be chosen large but independent of n. These parameters will be taken large after n is sent to infinity. Moreover, they will be ordered in a decreasing fashion, so that $k_j \gg k_{j+1}$.

We formulate a sequential extremal process, indexed by k_1, of approximation for the process of near maxima, which we refer to as the *landscape*. Divide the unit circle into consecutive arcs $\{\widehat{I_{j,N}}\}$ by the formula that for any $j, N \in \mathbb{N}$,

$$\hat{I}_j = \widehat{I_{j,N}} := 2\pi\big[\tfrac{(j-1)k_1}{N}, \tfrac{jk_1}{N}\big). \tag{2.4.16}$$

Let $\mathcal{D}_{N/k_1}$ denote the collection of indices $j = 1, 2, \ldots, \lceil \frac{N}{k_1} \rceil$. We let $\theta_j = \theta_{j,N}$ be the supremum of $\widehat{I_{j,N}}$. Over each of these intervals, we define the process

$$D_j = D_{j,N} : [-2\pi k_1, 0] \to \mathbb{C},$$

$$D_j(\theta) := \begin{cases} \big(\Phi_N^*\big)^2(\exp(i(\theta_j + \frac{\theta}{N}))) \cdot \exp(-i(N+1)\theta_j - \sqrt{\frac{8}{\beta}} m_N), & \text{if } \sigma = 1, \\ \exp\big(\varphi_N(\theta_j + \frac{\theta}{N}) - \sqrt{\frac{8}{\beta}} m_N\big), & \text{if } \sigma = i. \end{cases} \tag{2.4.17}$$

This will serve as the *decoration process*, although we will not prove their convergence as $k_1 \to \infty$. We next define for all $j \in \mathcal{D}_{n/k_1}$ random variables

$$\widehat{W}_j = \widehat{W}_{j,N} := \max_{\theta \in \widehat{I}_j} \{\varphi_N(\theta)\} - \sqrt{\frac{8}{\beta}} m_N \tag{2.4.18}$$

which is a local maximum, appropriately centered, and set

$$\mathrm{Ex}_n = \mathrm{Ex}_n^{k_1} := \sum_{j \in \mathcal{D}_{N/k_1}} \delta_{(\theta_j, \widehat{W}_j, D_j)}. \tag{2.4.19}$$

A central technical challenge will be to show that φ_k and Ψ_k are essentially constant on the interval $\hat{I}_j$ for $k \approx N/k_1$, and that hence it suffices to track both φ_k and Ψ_k only at the point $\theta_j \in \hat{I}_j$. We gloss over this detail in these notes. Our goal will be to approximate the process Ex_N by a Poisson processes with random intensity, which we describe next.

Toward this end, recall that an important strategy used throughout the analysis of extrema of branching processes is effectively conditioning on the initial portion of the process, wherein the extrema gain a nontrivial correlation. We will do the same and condition on the first Verblunsky coefficients. We use the parameter k_2, which we assume is a power of 2 (to apply Theorem 12), to refer to how many Verblunsky coefficients on which we condition. We also use $(\mathscr{F}_k : k \in \mathbb{N}_0)$ to refer to the natural σ-algebra generated by the sequence of Verblunsky coefficients $(\gamma_k : k \in \mathbb{N}_0)$.

Introduce the law $\mathfrak{p}_{k_1}(v,\cdot)$, which is a law of a random function on $\theta \in (-2\pi k_1, 0)$ which is related to the exponential of the solution of a family of coupled diffusions $\mathfrak{U}^o_s(\theta)$ in an auxiliary time parameter (We discuss below in Section 2.6 these equations).

The measure Ex_N will be approximated by a Poisson random measure with a random intensity on the same space. This intensity on $\Gamma := [0, 2\pi] \times \mathbb{R} \times \mathcal{C}([-2\pi k_1, 0], \mathbb{C})$ will take the form of a product measure $\mathscr{D}_\infty \times \widehat{\mathfrak{p}_{k_1}}$, where $(\widehat{\mathfrak{p}_{k_1}} : k_1 \in \mathbb{N})$ is a deterministic Radon measure on $\mathbb{R} \times \mathcal{C}([-2\pi k_1, 0], \mathbb{C})$, which is constructed as follows. Let

$$\iota(v, f) := \Big(\max_{x \in [-2\pi k_1, 0]} (-\sqrt{4/\beta}v + \log|f(x)|), f e^{-\sqrt{4/\beta}v}\Big) \tag{2.4.20}$$

be a map of $\mathbb{R} \times \mathcal{C}([-2\pi k_1, 0], \mathbb{C})$ to itself, let

$$I(v) = \sqrt{\tfrac{2}{\pi}} v e^{\sqrt{2}v} \mathbf{1}\left\{(\log k_1)^{1/10} \leqslant v \leqslant (\log k_1)^{9/10}\right\},$$

and let

$$\widehat{\mathfrak{p}_{k_1}}(dv, df) \text{ denote the push-forward of } I(v)dv \times \mathfrak{p}_{k_1}(v, df) \text{ by } \iota. \tag{2.4.21}$$

We let Π^{k_1} be a Poisson random measure on Γ with intensity $\mathscr{D}_\infty \times \widehat{\mathfrak{p}_{k_1}}$. It is worthwhile explaining what exactly is this process: the first coordinate marks the "height" of the recursion at angle θ, ie generates "high points" (corresponding to some very large level k_2) according to the intensity $\mathscr{D}_\infty$. At these high values, one generates pieces of trajectories (the *decoration*) around height v, with intensity $\widehat{\mathfrak{p}_{k_1}}(dv, df)$.

To compare point processes on Γ, we endow the latter with the distance

$$\partial_0((\theta_1, z_1, f_1), (\theta_2, z_2, f_2)) := \Big(d_{\mathbb{T}}(\theta_1, \theta_2) + |z_1 - z_2| + \sup_{t \in [-2\pi k_1, 0]} |f_1(t) - f_2(t)|\Big) \wedge 1.$$

In terms of this we define a Wasserstein distance on point configurations $\xi_1 = \sum_{i=1}^m \delta_{y_i}$ and $\xi_2 = \sum_{i=1}^n \delta_{z_i}$

$$\partial_1(\xi_1, \xi_2) := \begin{cases} 0, & \text{if } m = n = 0, \\ \min_\pi \max_{i=1,\dots,n} \partial_0(y_i, z_{\pi(i)}), & \text{if } m = n > 0, \\ 1, & \text{if } m \neq n. \end{cases}$$

with the minimum being the distance over all permutations π of $\{1, 2, \dots, n\}$. Finally, for two point processes Q_1 and Q_2 we define the pseudometric

$$\partial_2(Q_1, Q_2) := \inf_{(\xi_1, \xi_2)} \mathbb{E}(\partial_1(\xi_1, \xi_2)),$$

with the infimum over couplings (ξ_1, ξ_2) in which $\xi_1 \sim Q_1$ and $\xi_2 \sim Q_2$. To make a comparison between Π^{k_1} and $\mathrm{Ex}_N^{k_1}$, we will only make a comparison in which their second coordinate is in a compact set. Hence we shall further restrict the space Γ to

$$\widehat{\Gamma}_{k_7} := [0, 2\pi] \times [-k_7, k_7] \times \mathcal{C}([-2\pi k_1, 0], \mathbb{C}). \tag{2.4.22}$$

The main approximation result is the following.

Theorem 13. *For any $k_7 > 0$, we have*

$$\limsup_{k_1, N \to \infty} \partial_2 \big(\Pi^{k_1} \cap \widehat{\Gamma}_{k_7}, \mathrm{Ex}_N^{k_1} \cap \widehat{\Gamma}_{k_7} \big) = 0. \tag{2.4.23}$$

Theorem 13 implies Theorems 9 and 10.

2.5 A high level description of the proof of Theorem 10.

We now provide a high level description of the proof, that glosses over many important details.

We write

$$\begin{aligned} \varphi_n(\theta) &= \varphi_{k_2}(\theta) + (\varphi_{N/k_1}(\theta) - \varphi_{k_2}(\theta)) + (\varphi_N(\theta) - \varphi_{N/k_1}(\theta)) \\ &=: \varphi_{k_2}(\theta) + \Delta_{k_2, N/k_1}(\theta) + \Delta_{N/k_1, n}(\theta), \end{aligned}$$

and

$$\max_{\theta \in [0, 2\pi]} \varphi_N(\theta) = \max_j \max_{\theta \in \widehat{I_{j,N}}} \Big(\varphi_{k_2}(\theta) + \Delta_{k_2, N/k_1}(\theta) + \Delta_{N/k_1, N}(\theta) \Big).$$

We claim that the last expression can be approximated as

$$\max_j \Big(\varphi_{k_2}(\theta_j) + \Delta_{k_2, N/k_1}(\theta_j) + \max_{\theta \in \widehat{I_{j,N}}} \Delta_{N/k_1, N}(\theta) \Big). \tag{2.5.24}$$

To analyze the maximum in (2.5.24), we introduce the field $f_{n,j}(\eta) := \Delta_{N/k_1, N}(\theta_j + \eta/N)$, with $\eta \in [-2\pi k_1, 0]$ and Δ defined above (2.5.24), and write (2.5.24) as

$$\begin{aligned} &\max_j \Big(\varphi_{k_2}(\theta_j) + \Delta_{k_2, N/k_1}(\theta_j) + \max_{\eta \in [-2\pi k_1, 0]} f_{n,j}(\eta) \Big) \qquad (2.5.25) \\ &=: \max_j \Big(\varphi_{k_2}(\theta_j) + \Delta_{k_2, N/k_1}(\theta_j) + \Delta'_{N/k_1, N}(j) \Big). \end{aligned}$$

The main contribution to the maximum comes from js with $\Delta_{k_2, N/k_1}(\theta_j)$ large, of the order of $\sqrt{8/\beta}(m_N - \log(k_1 k_2))$. However, the $\Delta_{k_2, N/k_1}(j)$ are far from independent for different j. In order to begin controlling this, we introduce two "good events": a global good event $\mathscr{G}_n$, which allows us to replace the recursion by one driven by Gaussian variables (called $\mathfrak{z}_t(\theta)$, and taken for convenience in continuous time) and also impose an *a priori* upper limit on the recursion, and a barrier event $\widehat{\mathscr{R}}$, which ensures that the Gaussian-driven recursion $\mathfrak{z}_t(\theta)$ stays within a certain entropic envelope. We will also insist that $\mathfrak{z}_{N/k_1}(\theta_j)$ stays within an appropriate window. These steps are similar to what is done in [CMN18] and prepare the ground for the application of the second moment method.

We next claim that the fields $f_{n,j}(\eta)$ converge in distribution to the solution of a system of coupled stochastic differential equations as in (2.6.29) (this is not literally the case, and requires some pre-processing in the form of restriction to appropriate events and using a change of k_1, that we gloss over here). In particular, the law of those fields are determined by the Markov kernel $\mathfrak{p}_{k_1}$. Further and crucially, the fields $f_{n,j}$ can be constructed so that for well separated js, they are independent. This latter independence is what makes the proof work. Of course, for adjoining arcs it is actually hopeless to make them truly independent, but it turns out we need the independence only for far away intervals, and this can be achieved due to the fast rotating Prüfer phases.

As in many applications of the second moment method, to allow for some decoupling it is necessary to condition on $\mathscr{F}_{k_2}$. We need to find high points of the right side of (2.5.25). The basic estimate, for a given j, is that with $w_j = \sqrt{8/\beta}\log k_2 - \varphi_{k_2}(\theta_j)$,

$$\mathbb{P}\Big(\varDelta_{k_2,N/k_1}(\theta_j) \sim \sqrt{8/\beta}(m_N - \log(k_1k_2) - v), \text{Barrier} \mid \mathscr{F}_{k_2}\Big) \sim C\frac{ve^{2v}w_je^{-2w_j}}{N}. \tag{2.5.26}$$

This estimate, already appearing in [CMN18], is nothing but a barrier estimate.

If the variables $\{\varDelta_{k_2,n/k_1}(\theta_j) + \varDelta'_{n/k_1,n}(j) : j\}$ were an independent family, we would be at this point done, for then we would have that

$$\begin{aligned}&\mathbb{P}\Big(\varphi_{k_2}(\theta_j) + \varDelta_{k_2,N/k_1}(\theta_j) + \varDelta'_{N/k_1,N}(j) > \sqrt{8/\beta}(m_N + x) \;\mid\; \mathscr{F}_{k_2}\Big)\\ &\sim C\frac{w_je^{-2w_j}}{N/k_1}\mathbb{E}_{\mathfrak{p}_{k_1}}\Big((V_j - x)e^{2(V_j - x)}\Big) \sim C\frac{\mathscr{D}_{k_2}(\theta_j)}{N/k_1}e^{-2x}.\end{aligned} \tag{2.5.27}$$

Hence, we have using independence over different j that

$$\begin{aligned}\mathbb{P}(\max_{\theta\in[0,2\pi]}\varphi_N(\theta) \leqslant \sqrt{8/\beta}(m_N + x) \;\mid\; \mathscr{F}_{k_2}) &\sim \prod_j\Big(1 - C\frac{\mathscr{D}_{k_2}(\theta_j)}{N/k_1}e^{-2x}\Big)\\ &\sim \exp\big(-C\mathscr{B}_{k_2}e^{-2x}\big),\end{aligned}$$

which would then yield Theorem 10.

Unfortunately, different js are not independent. We handle that through several Poisson approximations. First, we condition on $\mathscr{F}_{n/k_1}$ and use the "two moments suffice" method of [AGG89] to show that the process of near maxima (together with the shape $(\varphi_N(\theta) - \varphi_{N/k_1}(\theta), \theta \in \widehat{I_{j,n}})$) can be well approximated, as $k_1 \to \infty$, by a Poisson point process of intensity $\mathfrak{m}$ which depends still on k_1. In the proof, the independence for well separated js and the second moment computations play a crucial role.

Using these steps we obtain a Poisson process with random intensity, measurable on $\mathscr{F}_{N/k_1^+}$. Our final step is another standard use of the second moment method to show that this random intensity concentrates, yielding Theorem 13.

2.6 Decoupling and a diffusion approximation

We now describe the SDE's alluded to before. Recall the recursions (2.2.8) and (2.2.9). Writing t instead of k, this is similar to the SDE

$$dX_t = i\theta dt + e^{c_\beta \Im X_t} dW_t,$$

where W_t is a complex Brownian motion. After a time change $t \mapsto e^t$ with now $t = \log N - \log k_1 + s$ and a change of scaling for the angles to correspond to small arcs, ie $\theta \mapsto \theta/N$ and we relativise near 0, we obtain the SDE

$$dY_s = \frac{i\theta e^s}{k_1} ds + c_\beta e^{i\Im Y_s} dW_s.$$

Note that for other intervals, W_s gets multiplied by an extra term $e^{i\Psi_t(\theta_j)}$, which oscillates rapidly, and therefore creates independence between far away k_1-intervals. This leads to the following. Set $T_+ = \log k_1$. Consider the family of standard complex Brownian motions

$$\{\mathfrak{W}_t^j : T_- \leqslant t \leqslant T_+, j \in \mathcal{D}_{N/k_1}\}, \tag{2.6.28}$$

which will have the property that they are independent from one another when the relevant arcs are separated by a small power of N. Now with respect to these Brownian motions we define the complex diffusions $(\mathfrak{L}_t^j : t \in [0, T_+], \theta \in \mathbb{R}, j \in \mathcal{D}_{n/k_1})$ as the (strong) solution of the stochastic differential equation

$$\begin{aligned} d\mathfrak{L}_t^j(\theta) = d\mathfrak{L}_t(\theta) &= i\theta e^t k_1^{-1} dt + \sqrt{\tfrac{4}{\beta}} e^{i\Im \mathfrak{L}_t(\theta)} d\mathfrak{W}_t^j, \text{ and} \\ \mathfrak{U}_t^j(\theta) &= -\Re\big(\sigma\big(\mathfrak{L}_t^j(\theta) - i\theta k_1^{-1} e^t\big)\big) - \sqrt{\tfrac{8}{\beta}} m_n. \\ \mathfrak{L}_{T_-}^j(\theta) &= -2\log \Phi^*_{n_1^+}(\exp(i(\theta_j + \tfrac{\theta}{n}))) + \tfrac{i\theta}{k_1} e^{T_-}, \quad \text{for} \quad \theta \in \mathbb{R}. \end{aligned} \tag{2.6.29}$$

As explained above, the diffusion $\mathfrak{L}^j(\theta)$ will serve as a proxy for the evolution of $-2\log \Phi^*_{k(t)}(\theta_j + \theta/N) + i\theta e^t/k_1$ where $k(t) \approx n_1 e^t$ up to rounding errors, and in particular its imaginary part will mimick the evolution of the Prüfer phases. When $\sigma = 1$, the diffusion $\mathfrak{U}_t^j(\theta)$ is designed to be a proxy for $2\log|\Phi^*_{k(t)}(\theta_j + \theta/N)|$.

In reality, the initial condition for the diffusion is not constant. However, one can move back to negative time (by a power smaller than 1 of $\log k_1$) and initialize the SDE from flat initial conditions, so as to decouple from the past. A technical part of the proof shows that no harm is done by this initialization.

Exercise 10. To get a feeling for the Poisson approximations in this section, consider a BRW of depth n, so that each vertex v is associated with a point $x_v = j2^{-n}$, and to each $x \in [j2^{-n}, (j+1)2^{-n}]$ we define $S(x) = S_v + \Re\mathfrak{L}_1^j((x-x_v)2^n)$, and $\mathfrak{L}_t^j$ is as in (2.6.29) with $k_1 = 1$ and $\mathfrak{L}_0^j = 0$. Prove that Theorem 6 continues to hold, with Θ replaced by $\Theta + C$ with some deterministic constant C. State a Poisson convergence for the local landscape.

3 Lecture III: Jacobi matrices and GβE

We consider in this section a class of random matrices that are tridiagonal; these include matrices whose eigenvalue distruibutions mimick the GβE ensembles. Our emphasis is on methods that work for all β, and therefore we do not discuss results specific to $\beta = 2$.

3.1 Tridiagonal representation of β ensembles

The following material will be skipped and serves only as motivation. It is taken from [AGZ10]. We begin by recalling the definition of χ random variables (with t degrees of freedom).

Definition 3.1.1 *The density on* $\mathbb{R}_+$

$$f_t(x) = \frac{2^{1-t/2} x^{t-1} e^{-x^2/2}}{\Gamma(t/2)}$$

is called the χ distribution *with* t *degrees of freedom, and is denoted* χ_t.

If t is integer and X is distributed according to χ_t, then X has the same law as $\sqrt{\sum_{i=1}^t \xi_i^2}$ where ξ_i are standard Gaussian random variables.

Let ξ_i be independent i.i.d. standard Gaussian random variables of zero mean and variance 1, and let $Y_i \sim \chi_{i\beta}$ be independent and independent of the variables $\{\xi_i\}$. Define the tridiagonal symmetric matrix $H_N \in \mathrm{Mat}_N(\mathbb{R})$ with entries $H_N(i,j) = 0$ if $|i-j| > 1$, $H_N(i,i) = \sqrt{2/\beta}\xi_i$ and $H_N(i,i+1) = Y_{N-i}/\sqrt{\beta}$, $i = 1, \ldots, N$. The main result of this section is the following.

Theorem 3.1.2 (Edelman–Dumitriu) *The joint distribution of the eigenvalues of* H_N *is given by*

$$C_N(\beta)\Delta(\lambda)^\beta e^{-\frac{\beta}{4}\sum_{i=1}^N \lambda_i^2}\,. \tag{3.1.3}$$

We begin by performing a preliminary computation that proves Theorem 3.1.2 in the case $\beta = 1$ and also turns out to be useful in the proof of the theorem in the general case.

Proof of Theorem 3.1.2 ($\beta = 1$**)** Let X_N be a matrix distributed according to the GOE law (and in particular, its joint distribution of eigenvalues has the density coinciding with (3.1.3)). Set $\xi_N = X_N(1,1)/\sqrt{2}$, which is a standard Gaussian variable. Let $X_N^{(1,1)}$ denote the matrix obtained from X_N by striking the first column and row, and let $Z_{N-1}^{\mathrm{T}} = (X_N(1,2), \ldots, X_N(1,N))$. Then Z_{N-1} is independent of $X_N^{(1,1)}$ and ξ_N. Let $\tilde{H}_N$ be an orthogonal $N-1$-by-$N-1$ matrix, measurable on $\sigma(Z_{N-1})$, such that $\tilde{H}_N Z_{N-1} = (\|Z_{N-1}\|_2, 0, \ldots, 0)$, and set $Y_{N-1} = \|Z_{N-1}\|_2$, noting that Y_{N-1} is independent of ξ_N and is distributed according to χ_{N-1}. (A particular choice of $\tilde{H}_N$ is the *Householder reflector* $\tilde{H}_N = I - 2uu^{\mathrm{T}}/\|u\|_2^2$, where $u = Z_{N-1} - \|Z_{N-1}\|_2(1, \ldots, 0)$.) Let

$$H_N = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{H}_N \end{pmatrix} .$$

Then the law of eigenvalues of $H_N X_N H_N^{\mathrm{T}}$ is still (3.1.3), while

$$H_N X_N H_N^{\mathrm{T}} = \begin{pmatrix} \sqrt{2}\xi_N & Y_{N-1} & \mathbf{0}_{N-2} \\ Y_{N-1} & & \\ & X_{N-1} & \\ \mathbf{0}_{N-2} & & \end{pmatrix},$$

where X_{N-1} is again distributed according to the GOE and is independent of ξ_N and Y_{N-1}. Iterating this construction $N-1$ times (in the next step, with the Householder matrix corresponding to X_{N-1}), one concludes the proof (with $\beta = 1$). □

We next prove some properties of the eigenvalues and eigenvectors of tridiagonal matrices. Let $\mathcal{D}_N$ denote the collection of diagonal N-by-N matrices with real entries, $\mathcal{D}_N^{\mathrm{d}}$ denote the subset of $\mathcal{D}_N$ consisting of matrices with distinct entries, and $\mathcal{D}_N^{\mathrm{do}}$ denote the subset of matrices with decreasing entries. Let $\mathcal{U}_N^{(1)}$ denote the collection of N-by-N orthogonal matrices, and let $\mathcal{U}_N^{(1),+}$ denote the subset of $\mathcal{U}_N^{(1)}$ consisting of matrices whose first row has all elements strictly positive.

We parametrize tridiagonal matrices by two vectors of length N and $N-1$, $\mathbf{a} = (a_1, \ldots, a_N)$ and $\mathbf{b} = (b_1, \ldots, b_{N-1})$, so that if $H \in \mathcal{H}_N^{(1)}$ is tridiagonal then $H(i,i) = a_{N-i+1}$ and $H(i, i+1) = b_{N-i}$. Let $\mathcal{T}_N \subset \mathcal{H}_N^{(1)}$ denote the collection of tridiagonal matrices with all entries of $\mathbf{b}$ strictly positive.

Lemma 3.1.4 *The eigenvalues of any $H \in \mathcal{T}_N$ are distinct, and all eigenvectors $v = (v_1, \ldots, v_N)$ of H satisfy $v_1 \neq 0$.*

Proof. The null space of any matrix $H \in \mathcal{T}_N$ is at most one dimensional. Indeed, suppose $Hv = 0$ for some nonzero vector $v = (v_1, \ldots, v_N)$. Because all entries of $\mathbf{b}$ are nonzero, it is impossible that $v_1 = 0$ (for then, necessarily all $v_i = 0$). So suppose $v_1 \neq 0$, and then $v_2 = -a_N/b_{N-1}$. By solving recursively the equation

$$b_{N-i}v_{i-1} + a_{N-i}v_i = -b_{N-i-1}v_{i+1},\ i = 2, \ldots, N-1, \tag{3.1.5}$$

which is possible because all entries of $\mathbf{b}$ are nonzero, all entries of v are determined. Thus, the null space of any $H \in \mathcal{T}_N$ is one dimensional at most. Since $H - \lambda I \in \mathcal{T}_N$ for any λ, the first part of the lemma follows. The second part follows because we showed that if $v \neq 0$ is in the null space of $H - \lambda I$, it is impossible to have $v_1 = 0$. □

Let $H \in \mathcal{T}_N$, with diagonals $\mathbf{a}$ and $\mathbf{b}$ as above, and write $H = UDU^{\mathrm{T}}$ with $D \in \mathcal{D}_N^{\mathrm{do}}$ and $U = [v^1, \ldots, v^N]$ orthogonal, such that the first row of U, denoted $\mathbf{v} = (v_1^1, \ldots, v_1^N)$, has nonnegative entries. (Note that $\|\mathbf{v}\|_2 = 1$.) Write $\mathbf{d} = (D_{1,1}, \ldots, D_{N,N})$. Let $\mathrm{d}_N^{\mathrm{c}} = \{(x_1, \ldots, x_N) : x_1 > x_2 \cdots > x_N\}$ and let

$$S_+^{N-1} = \{\mathbf{v} = (v_1, \ldots, v_N) \in \mathbb{R}^N : \|\mathbf{v}\|_2 = 1,\ v_i > 0\}\,.$$

(Note that $\mathrm{d}_N^{\mathrm{c}}$ is similar to d_N, except that the ordering of coordinates is reversed.)

Lemma 3.1.6 *The map*

$$(\mathbf{a}, \mathbf{b}) \mapsto (\mathbf{d}, \mathbf{v}) : \mathbb{R}^N \times \mathbb{R}_+^{(N-1)} \to \Delta_N^{\mathrm{c}} \times S_+^{N-1} \tag{3.1.7}$$

is a bijection, whose Jacobian J is proportional to

$$\frac{\Delta(\mathbf{d})}{\prod_{i=1}^{N-1} b_i^{i-1}}\,. \tag{3.1.8}$$

Proof. That the map in (3.1.7) is a bijection follows from the proof of Lemma 3.1.4, and in particular from (3.1.5) (the map $(\mathbf{d}, \mathbf{v}) \mapsto (\mathbf{a}, \mathbf{b})$ is determined by the relation $H = UDU^{\mathrm{T}}$).

To evaluate the Jacobian, we recall the proof of the $\beta = 1$ case of Theorem 3.1.2. Let X be a matrix distributed according to the GOE, consider the tridiagonal matrix with diagonals $\mathbf{a}, \mathbf{b}$ obtained from X by the successive Householder transformations employed in that proof. Write $X = UDU^*$ where U is orthogonal, D is diagonal (with elements $\mathbf{d}$), and the first row $\mathbf{u}$ of U consists of nonnegative entries (and strictly positive except on a set of measure 0). Check that $\mathbf{u}$ is independent of D and that, by a simple Jacobian computation, the density of the distribution of the vector $(\mathbf{d}, \mathbf{u})$ with respect to the product of the Lebesgue measure on $\mathrm{d}_N^{\mathrm{c}}$ and the the uniform measure on S_+^{N-1} is proportional to $\Delta(\mathbf{d})e^{-\sum_{i=1}^N d_i^2/4}$. Using Theorem 3.1.2 and the first part of the lemma, we conclude that the latter (when evaluated in the variables $\mathbf{a}, \mathbf{b}$) is proportional to

$$Je^{-\sum_{i=1}^N \frac{a_i^2}{4} - \sum_{i=1}^{N-1} \frac{b_i^2}{2}} \prod_{i=1}^{N-1} b_i^{i-1} = Je^{-\sum_{i=1}^N d_i^2/4} \prod_{i=1}^{N-1} b_i^{i-1}\,.$$

The conclusion follows. □

We will also need the following useful identity.

Lemma 3.1.9 *With notation as above, we have the identity*

$$\Delta(\mathbf{d}) = \frac{\prod_{i=1}^{N-1} b_i^i}{\prod_{i=1}^{N} v_1^i} . \tag{3.1.10}$$

Proof. Write $H = UDU^{\mathrm{T}}$. Let $e_1 = (1, 0, \dots, 0)^{\mathrm{T}}$. Let w^1 be the first column of U^{T}, which is the vector made out of the first entries of $v^1, \dots, v^n$. One then has

$$\prod_{i=1}^{N-1} b_i^i = \det[e_1, He_1, \dots, H^{N-1}e_1] = \det[e_1, UDU^{\mathrm{T}}e_1, \dots, UD^{N-1}U^{\mathrm{T}}e_1]$$
$$= \pm \det[w^1, Dw^1, \dots, D^{N-1}w^1] = \pm\Delta(\mathbf{d}) \prod_{i=1}^{N} v_1^i .$$

Because all terms involved are positive by construction, the $\pm$ is actually a $+$, and the lemma follows. □We can now conclude.

Proof of Theorem 3.1.2 (general $\beta > 0$) The density of the independent vectors $\mathbf{a}$ and $\mathbf{b}$, together with Lemma 3.1.6, imply that the joint density of $\mathbf{d}$ and $\mathbf{v}$ with respect to the product of the Lebesgue measure on $\mathrm{d}_N^{\mathrm{c}}$ and the uniform measure on S_+^{N-1} is proportional to

$$J \prod_{i=1}^{N-1} b_i^{i\beta-1} e^{-\frac{\beta}{4}\sum_{i=1}^{N} d_i^2} . \tag{3.1.11}$$

Using the expression (3.1.8) for the Jacobian, one has

$$J \prod_{i=1}^{N-1} b_i^{i\beta-1} = \mathrm{d}(\mathbf{d}) \left(\prod_{i=1}^{N-1} b_i^i\right)^{\beta-1} = \mathrm{d}(\mathbf{d})^{\beta} \left(\prod_{i=1}^{N} v_1^i\right)^{\beta-1} ,$$

where (3.1.10) was used in the second equality. Substituting in (3.1.11) and integrating over the variables $\mathbf{v}$ completes the proof. □

3.2 Characteristic polynomials for GβE and Jacobi matrices

We will be interested in the characteristic polynomial of Jacobi matrices similar to H_N. By Section 3.1, the law of the characteristic polynomial for H_N is the same as that for the GβE ensembles. There are several methods for handling that. In particular, [BMP22], using loop equations and z complex, showed that $\log|\det(zI - H_N)|$ is a logarithmically correlated field (more on that in Lecture IV). We will take here a different route, closer to what was done in the case of the CβE; in doing so, we follow [ABZ23] and [AZ25]; see [LP23, LP20] for other related results.

We consider matrices of the form

$$J_n = \begin{pmatrix} b_n & a_{n-1} & 0 & \cdots & 0 \\ a_{n-1} & b_{n-1} & a_{n-2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a_2 & b_2 & a_1 \\ 0 & \cdots & \cdots & a_1 & b_1 \end{pmatrix}. \tag{3.2.12}$$

where the coefficients of J_n satisfy the following assumptions.

Assumption 3.2.13 *$(a_k)_{k\geqslant 1}$ and $(b_k)_{k\geqslant 1}$ are two independent sequences of independent random variables whose law is absolutely continuous with respect to the Lebesgue measure and such that*

$$\mathbb{E}(a_k^2) = k + O(1),\ \mathrm{Var}(a_k^2) = kv + O(1),\quad \mathbb{E}(b_k) = 0,\ \mathrm{Var}(b_k) = v + O\Big(\frac{1}{k}\Big), \tag{3.2.14}$$

where v is some positive constant. Further, there exists $\mathfrak{h}_0 > 0$ such that

$$\sup_{k\geqslant 1} \mathbb{E}\big(e^{\mathfrak{h}_0 |b_k|}\big) < +\infty,\ \sup_{k\geqslant 1} \mathbb{E}\big(e^{\frac{\mathfrak{h}_0}{\sqrt{k}}|a_k^2 - \mathbb{E}(a_k^2)|}\big) < +\infty. \tag{3.2.15}$$

We denote by p_n the characteristic polynomial of the scaled Jacobi matrix $J_n/\sqrt{n}$ defined by $p_n(z) = \det(zI_n - J_n/\sqrt{n})$ for any $z \in \mathbb{R}$. Our main result reads as follows.

Theorem 14. *Let $\eta > 0$ and denote by $I_\eta := \{z \in \mathbb{R} : \eta \leqslant |z| \leqslant 2 - \eta\}$. In probability,*

$$\frac{\max_{z\in I_\eta} \big(\log|p_n(z)| - n\big(\frac{z^2}{4} - \frac{1}{2}\big)\big) - \sqrt{v}\log n}{\log\log n} \underset{n\to+\infty}{\longrightarrow} -\frac{3\sqrt{v}}{4}.$$

The GβE can be checked to be a particular case of Theorem 14. This partially confirms the conjecture Fyodorov and Simm stated for the GUE [FS16]. We note that the linear in n term (i.e., $n(z^2/4 - 1/2)$) corresponds to the logarithmic potential of the semicircle law.

3.3 The three term recursion

Owing to the tridiagonal structure of J_n, its characteristic polynomial is naturally linked to a certain three term recursion. More precisely, for any $z \in (-2, 2)$ and $n \in \mathbb{N}$, $n \geqslant 1$, let $(q_k(z))_{k\in\{-1,\dots,n\}}$ be defined by the recursion:

$$q_{-1}(z) = 0,\ q_0(z) = 1,\ q_k(z) = \big(z\sqrt{n} - b_k\big)q_{k-1} - a_{k-1}^2 q_{k-2}(z),\ k \geqslant 1, \tag{3.3.16}$$

where $a_0 = 0$ by convention. Then $q_k(z) = \det(z\sqrt{n}I_k - J_k)$, and in particular $q_n(z) = n^{n/2}p_n(z)$. Let $(\phi_k)_{k\in\{0,\dots,n\}}$ denote the scaled variables defined by

$$\phi_k(z) = \frac{q_k(z)}{\sqrt{k!}},\ k \in \{0,\dots,n\}, z \in (-2,2). \tag{3.3.17}$$

The choice of this scaling is motivated by the fact that these new polynomials now satisfy the recursion

$$\phi_k(z) = \Big(z_k - \frac{b_k}{\sqrt{k}}\Big)\phi_{k-1}(z) - \frac{a_{k-1}^2}{\sqrt{k(k-1)}}\phi_{k-2}(z), \quad k \geqslant 2, z \in (-2,2), \tag{3.3.18}$$

where $z_k = z\sqrt{n/k}$ and where the second order coefficients are of order 1, given that by assumption $a_{k-1}^2 \asymp k$ typically. Setting

$$X_k^z = \begin{pmatrix} \phi_k(z) \\ \phi_{k-1}(z) \end{pmatrix} \quad T_k^z = \begin{pmatrix} z_k - \frac{b_k}{\sqrt{k}} & -\frac{a_{k-1}^2}{\sqrt{k(k-1)}} \\ 1 & 0 \end{pmatrix}, \quad k \geqslant 2, \tag{3.3.19}$$

the recursion (3.3.18) is equivalent to

$$X_k^z = T_k^z X_{k-1}^z, \quad k \geqslant 2, z \in (-2,2), \tag{3.3.20}$$

where $X_1^z = (z\sqrt{n} - b_1, 1)^{\mathsf{T}}$. To get a feel for this recursion, note that under our Assumptions 3.2.13, T_k^z is a small random perturbation of the matrix A_k^z defined by

$$A_k^z := \begin{pmatrix} z_k & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.3.21}$$

As A_k^z belongs to $\mathbb{SL}_2(\mathbb{R})$, the dynamics of the system will highly depend in which of the three classes of $\mathbb{SL}_2(\mathbb{R})$, hyperbolic, parabolic or elliptic, it belongs to. Since this classification is determined by respectively the value of $|\mathrm{Tr}(A_k^z)|$ being strictly greater than 2, equal to 2 or strictly smaller than 2, this leads us to define the critical time $k_{0,z}$ and - although less obvious for now - the critical window ℓ_0 as

$$k_{0,z} := \lfloor \frac{z^2 n}{4} \rfloor, \quad \ell_0 := \lfloor \kappa n^{1/3} \rfloor, \quad \kappa \geqslant 1, \tag{3.3.22}$$

and to decompose the recursion into three regimes: a *hyperbolic regime* (until time $k_{0,z} - \ell_0$) where the eigenvalues of A_k^z are real, a *parabolic regime* (between time $k_{0,z} - \ell_0$ and $k_{0,z} + \ell_0$), and an *elliptic regime* (after time $k_{0,z} + \ell_0$) where the eigenvalues of A_k^z are complex conjugated and of modulus 1.

Set $\alpha_{k,z}$ when $k < k_{0,z}$, the spectral radius of A_k^z, given by

$$\alpha_{k,z} := \frac{|z_k| + \sqrt{z_k^2 - 4}}{2}, \quad 1 \leqslant k \leqslant k_{0,z}. \tag{3.3.23}$$

With this notation, Theorem 14 is actually equivalent, modulu some anti-concentration results that we do not detail, to the following:

Theorem 15. *Let $\eta > 0$ and denote by $I_\eta := \{z \in \mathbb{R} : \eta \leqslant |z| \leqslant 2-\eta\}$. Under Assumptions 3.2.13,*

$$\frac{\max_{z\in I_\eta}\left(\log\|X_n^z\| - \sum_{k=1}^{k_{0,z}}\log\alpha_{k,z}\right) - \sqrt{v}\log n}{\log\log n} \underset{n\to+\infty}{\longrightarrow} -\frac{3\sqrt{v}}{4}, \qquad (3.3.24)$$

in probability.

In these notes, we neglect completely the parabolic regime, and focus instead on the hyperbolic and elliptic regimes. The parabolic regime contributes a variance of order 1 and does not affect the analysis except if one wants to obtain convergence results, which we do not aim for.

3.4 The hyperbolic regime

Conceptually, the easiest part to handle is the hyperbolic phase: indeed, in that region, we expect that $\log|\phi_k| \sim \sum_{i=1}^k \log\alpha_i$, and therefore it is natural to define $\psi_k = \phi_k/\prod_{i=1}^k \alpha_i$. Writing now $\nu_k = \psi_k/\psi_{k-1}$, the recursion (3.3.18) reads

$$\nu_k(z) = \frac{(z_k - b_k/\sqrt{k})}{\alpha_k} - \frac{a_{k-1}^2}{\sqrt{k(k-1)}\alpha_k\alpha_{k-1}}\frac{1}{\nu_{k-1}}(z). \qquad (3.4.25)$$

We expect to have $\nu_k \sim 1$, and therefore, writing $\nu_k = 1+\delta_k$, we obtain the recursion

$$\delta_k = u_k + v_k\frac{\delta_{k-1}}{1+\delta_{k-1}} \sim u_k + v_k\delta_{k-1} - v_k\delta_{k-1}^2, \qquad (3.4.26)$$

where, up to negligible terms,

$$u_k = \frac{z_k}{\alpha_k} - 1 - \frac{1+g_k/\sqrt{k}}{\alpha_k\alpha_{k-1}} - \frac{b_k}{\alpha_k\sqrt{k}}, \quad v_k = \frac{1+g_k/\sqrt{k}}{\alpha_k\alpha_{k-1}}. \qquad (3.4.27)$$

It is not hard to check, using (3.3.23), that for $k \in [\epsilon n, k_0]$, we have up to negligible terms that

$$u_k = \frac{1}{2\sqrt{k_0(k_0-k)}} + \tilde{g}_k\sqrt{\frac{2v}{k}}, \quad v_k = \frac{(1+g_k/\sqrt{k})}{1+\sqrt{(k_0-k)/k_0}},$$

where the independent variables $\tilde{g}_k$ have mean 0 and variance 1. One can now solve (3.4.26) in two steps: first, disregard the term δ_{k-1}^2 and solve the resulting linear equation, and then computing the perturbation due to the quadratic term (replacing δ_{k-1}^2 by the solution to the linearized equation). It is not hard to verify that the approximate solution obtained by this procedure is within $O(1)$, with high probability, from the exact solution.

Eventually, we need to compute

$$\sum_{j=\epsilon n}^{k_0-\ell_0} \log(1+\delta_j) \sim \sum_{j=\epsilon n}^{k_0-\ell_0} \delta_j - \frac{1}{2}\sum_{j=\epsilon n}^{k_0-\ell_0} \delta_j^2.$$

The second term will turn out to have fluctuations of order 1, and its fluctuations are therefore negligible (it does contribute to the mean!). However, the first term is very far from being a random walk, due to the high correlation between δ_j and δ_{j-1}, see (3.4.26). However, the correlation length is of the order of $\sqrt{k_0/(k_0-k)}$. Choosing sequences L_j so that $L_j = k_0^{1/3} j^{2/3}$, $j = 1, \ldots, \epsilon n$, one finds that the variables

$$\Delta_j = \sum_{i=k_0-L_{j+1}}^{k_0-L_j} \delta_j$$

are essentially uncorrelated, and have variance $1/j$. Now we are back to the random walk setup, and the techniques we discussed in the context of the CβE can be applied!

Remark 8. The point of view described above is the one taken in [ABZ23]. There is a slightly more geometric point of view, that unifies the treatment of the hyperbolic and elliptic regimes, and that is developed in [AZ25]. Due to time constraints, we do not describe it here.

Exercise 11. Consider the linearized recursion in the hyperbolic regime, i.e. (3.4.26) without the quadratic term. Show that $\sum_{i=t}^{k_0-\ell_0} \log(1+\delta_i)$, after an exponential time change, can be coupled to a Brownian motion with drift.

3.5 Change of basis and description of the new recursions

To analyze the recursion (3.3.20), we perform a certain change of basis to leverage the geometric properties of the expected transition matrix, which is roughly A_k^z. To this end, define the following time-dependent change of basis $(P_k^z)_{1\leqslant k\leqslant n}$ as

$$P_k^z := \begin{pmatrix} 1 & \alpha_{k,z}^{-1} \\ \alpha_{k,z}^{-1} & 1 \end{pmatrix},\ 1 \leqslant k \leqslant k_{0,z}-\ell_0,\ P_k^z := \begin{pmatrix} \frac{\sqrt{4-z_k^2}}{2} & \frac{z_k}{2} \\ 0 & 1 \end{pmatrix},\ k_{0,z}+\ell_0 \leqslant k \leqslant n, \tag{3.5.28}$$

and set $P_k^z := P_{k_{0,z}-\ell_0}^z$ for $|k-k_{0,z}| < \ell_0$, where $\alpha_{k,z}$ is defined in (3.3.23) and $z_k := z\sqrt{n/k}$.

The choice of this change of basis is motivated by the fact that $\|P_k^z\| \lesssim 1$ and that $D_k^z := (P_k^z)^{-1} A_k^z P_k^z$, $1 \leqslant k \leqslant k_{0,z} - \ell_0$ is a diagonal matrix whereas $R_k^z := (P_k^z)^{-1} A_k^z P_k^z$, $k \geqslant k_{0,z} + \ell_0$ is a rotation. More precisely,

$$D_k^z := \begin{pmatrix} \alpha_{k,z} & 0 \\ 0 & \alpha_{k,z} \end{pmatrix},\ k \leqslant k_{0,z}-\ell_0,\ R_k^z := \begin{pmatrix} \cos(\theta_k^z) & -\sin(\theta_k^z) \\ \sin(\theta_k^z) & \cos(\theta_k^z) \end{pmatrix},\ k \geqslant k_{0,z}+\ell_0, \tag{3.5.29}$$

where $\theta_k^z \in [0, 2\pi)$ is such that $e^{i\theta_k^z} = (z_k + \sqrt{4 - z_k^2})/2$ for $k \geqslant k_{0,z} + \ell_0$.

In the sequel, we denote by Ξ_k^z the new transition matrix and by Y_k^z the coordinate vector of X_k^z in the basis P_k^z, that is

$$\Xi_{k+1}^z := (P_{k+1}^z)^{-1} T_{k+1}^z P_k^z, \; Y_k^z = P_k^z X_k^z, \; 1 \leqslant k \leqslant n. \tag{3.5.30}$$

With this notation, the sequence Y^z satisfies the recursion $Y_{k+1}^z = \Xi_{k+1}^z Y_k^z$ for any $1 \leqslant k \leqslant n$. Our central observable will be the field $\psi_k(z)$ defined by $\psi_0(z) = 0$ and

$$\psi_k(z) := \log \|Y_k^z\| - M_k(z), \; 1 \leqslant k \leqslant n, \; z \in I_\eta, \tag{3.5.31}$$

where $M_k(z)$ denotes the accumulated mean defined as the sum of the "instantaneous" mean $\mu_k(z)$ by

$$\mu_k(z) = \frac{v-1}{4(k - k_{0,z})}, \quad M_k(z) := \sum_{\ell = k_0 + \ell_0}^{k} \mu_\ell(z). \tag{3.5.32}$$

3.6 Recursion in the elliptic regime

In this regime where $k_{0,z} + \ell_0 \leqslant k \leqslant n$, we show that with high probability the vector Y_k^z rotates with essentially the same angle as R_k^z over short enough blocks and that the increments of the process $\psi(z)$ over these blocks are well-approximated by a sum of independent random variables. More precisely, define the blocks $(k_{i,z})_{j_o \leqslant i \leqslant j_1}$ by $k_{j_o,z} := k_{0,z} + \ell_0$, $k_{j_1,z} := n$, and

$$k_{i,z} := k_{0,z} + \lfloor i^4 k_{0,z}^{1/3} \rfloor, \; j_o < i < j_1, \tag{3.6.33}$$

and where $j_o := \max\{i : \lfloor i^4 k_{0,z}^{1/3} \rfloor \leqslant \ell_0\}$ and $j_1 := \min\{i : k_{0,z} + \lfloor i^4 k_{0,z}^{1/3} \rfloor \geqslant n\}$. For any $k \geqslant k_{0,z} + \ell_0$, denote by ζ_k^z a measure of the argument of Y_k^z in $[0, 2\pi)$. Next, say that the i^{the} block is *good* if for all $k_{i,z} \leqslant k < k_{i+1,z}$,

$$\zeta_k^z - \zeta_{k_{i,z}}^z + \sum_{\ell = k_{i,z}+1}^{k} \theta_\ell^z \in [-\delta_i, \delta_i] + 2\pi\mathbb{Z}, \tag{3.6.34}$$

where $\delta_i := i^{-1/4}$. Denote also by $\mathscr{G}_{i,z}$ the event that the i^{th} block is good. With this notation, we have the following proposition.

Proposition 3.6.35 (Probability of a "bad" block) *For any $i \geqslant \kappa^{1/4}$,*

$$\mathbb{P}\big(\mathscr{G}_{i,z}^{\complement}\big) \leqslant e^{-\mathfrak{c} i^{1/2}},$$

where $\mathfrak{c}$ is a positive constant depending on the model parameters.

As an immediate consequence, we can conclude by using a union bound that all the blocks starting from the $(\log n)^{\mathfrak{q}}$-th block on are good with overwhelming probability provided $\mathfrak{q}$ is large enough.

Corollary 3.6.36 *Let $\mathfrak{q} > 4$. Denote by $\mathscr{G}_z := \bigcap_{i \geqslant (\log n)^{\mathfrak{q}}} \mathscr{G}_{i,z}$. Then,*

$$\mathbb{P}(\mathscr{G}_z^{\complement}) \leqslant e^{-(\log n)^2},$$

where $\mathfrak{c} > 0$ depends on the model parameters.

Next, we show a representation of the increments $\Delta\psi_{k_{i+1,z}}(z) := \psi_{k_{i+1,z}}(z) - \psi_{k_{i,z}}(z)$ as a sum of independent random variables up to some small error on the event where the i^{th} block is good. This representation will be at the base of all our subsequent results in the elliptic regime. To describe the noise appearing in the new recursion in the elliptic regime we define the variables $c_{k,z}$ and d_k as

$$c_{k,z} := -\frac{2}{\sqrt{4-z_k^2}}\Big(\frac{z_k(b_k - \mathbb{E}(b_k))}{2\sqrt{k}} + \frac{a_{k-1}^2 - \mathbb{E}a_{k-1}^2}{\sqrt{k(k-1)}}\Big), \quad d_k := -\frac{b_k - \mathbb{E}(b_k)}{\sqrt{k}}. \tag{3.6.37}$$

With this notation, the "instantaneous" noise at time k is given by the function w_k^z defined as

$$w_k^z(\zeta) := c_k^z \sin(\theta_k^z + \zeta)\cos(\zeta) + d_k \sin(\theta_k^z + \zeta)\sin(\zeta), \quad \zeta \in \mathbb{R}. \tag{3.6.38}$$

We are now ready to state our result in the elliptic regime.

Proposition 3.6.39 (Representation of the increments on "good" blocks) *Let $i \geqslant \kappa^{1/4}$. On the event $\mathscr{G}_{i,z}$,*

$$\Delta\psi_{k_{i+1,z}}(z) = \sum_{k=k_{i,z}+1}^{k_{i+1,z}} \big(w_k^z(\zeta_{i,k-1}^z) + \mathcal{P}_k^z\big) + O(i^{-5/4}),$$

where $\zeta_{i,\ell}^z := \zeta_{k_{i,z}}^z + \sum_{k=k_{i,z}+1}^{\ell} \theta_k^z$ for any $\ell \in [k_{i,z}, k_{i+1,z}]$, w_k^z is defined in (3.6.38) and $\mathcal{P}_k^z$ is a $\mathcal{F}_k$-measurable variable satisfying that

$$s_{k-1}(\mathcal{P}_k^z) \lesssim \frac{1}{i^{1/4}(k-k_0)} + \frac{(\log n)^{\mathfrak{C}}}{(k-k_{0,z})^2}, \tag{3.6.40}$$

where $\mathfrak{C} > 0$ depends on the model parameters. There exists a deterministic sequence $\sigma_{k,z}^2 > 0$, $k \geqslant k_{0,z} + \ell_0$ such that

$$\operatorname{Var}_{k_{i,z}}\Big(\sum_{k=k_{i,z}+1}^{k_{i+1,z}} w_k(\zeta_{i,k-1}^z)\Big) = \sum_{k=k_{i,z}+1}^{k_{i+1,z}} \sigma_{k,z}^2 + O(i^{-7}), \quad a.s. \tag{3.6.41}$$

Moreover,

$$\sigma_{k,z}^2 = \frac{v}{k-k_{0,z}} + O\Big(\frac{1}{\sqrt{k_0(k-k_0)}}\Big). \tag{3.6.42}$$

3.7 A priori exponential moment estimate

Finally, we state the a priori exponential estimate for increments of the process $\psi(z)$.

Proposition 3.7.43 *There exist* $\mathfrak{h} > 0$ *depending on the model parameters* $\mathfrak{C}_\delta > 0$ *depending on* $\delta > 0$ *and* $\mathfrak{C}_\kappa$ *depending on* κ *such that for any* $z \in I_\eta$, $0 \leqslant k \leqslant k' \leqslant k^-_{\delta,z}$ *and* $1 \leqslant \lambda \leqslant (\log n)^{-\mathfrak{h}} n^{1/6}$,

$$\log \mathbb{E}_k\big[e^{\lambda(\psi_{k'}(z)-\psi_k(z))}\big] \leqslant \mathfrak{C}_\delta \lambda^2,$$

and for any $k^-_{\delta,z} \leqslant k \leqslant k' \leqslant n$,

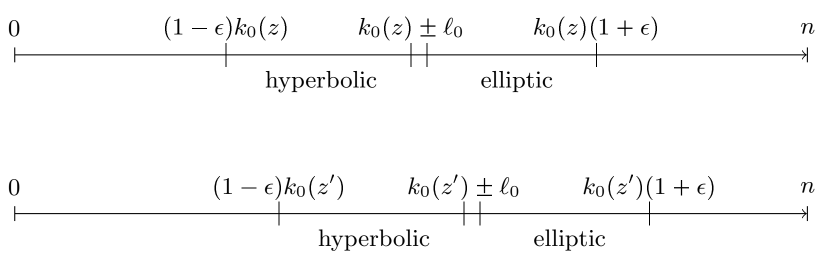
$$\log \mathbb{E}_k\big[e^{\lambda(\psi_{k'}(z)-\psi_k(z))}\big] \leqslant \mathfrak{C}_\kappa \lambda^2 \sum_{\ell=k+1}^{k'} \frac{1}{|k_{0,z}-k| \vee n^{1/3}}. \tag{3.7.44}$$

Moreover, if $k' \leqslant k_{0,z} - \ell_0$ *or* $k \geqslant k_{0,z} + \ell_0$, $\mathfrak{C}_\kappa$ *can be taken independent of* κ. *Further if* $k \geqslant k_{0,z} - \ell_0$, *then* (3.7.44) *holds for* $1 \leqslant -\lambda \leqslant (\log n)^{-\mathfrak{h}} n^{-1/6}$.

In particular, this result justifies the claim that the parabolic regime has only a contribution to the field $\psi(z)$ of order 1 depending on κ.

3.8 The log-correlated structure

We now discuss the correlation structure of the field. It is worthwhile to keep in mind the following diagram.



Depending on the distance between z, z', there are different overlaps in regimes:

– **[small distance]** If $|z - z'| \ll n^{-2/3}$, the overlap between the hyperbolic regimes for z and z' is complete, and the increments are fully correlated (this corresponds to "late branching" in the BRW tree picture). The elliptic regime overlaps, but the rotation frequency for z and z' is different. One has correlation only for such ℓ that $|\theta^z_\ell - \theta^{z'}_\ell|^{-1} < \theta^\ell_z$ (this is similar to the oscilatory phase in the CβE). This gives covariance proportional to $-\log|z - z'|$.

– **[moderate distance]** If $|z-z'| \gg n^{-2/3}$ but $|z-z'| \ll 1$ then the elliptic regime achieves full decorelation, and one only has correlation from the hyperbolic reguime. Due to the logarithmic time change (see Exercise 11), the covariance is again proportional to $-\log|z-z'|$.
– **[large distance]** If $|z-z'|$ is of order 1, there is esssentially no correlation (covariance of order 1 for variables of variance $\log n$) (this corresponds to "early branching" in the BRW tree picture).

(In the above, I talk about correlation, but in fact this holds also at the level of exponential moments.)

These correlation/decorrelation estimates are important at two places: First, the correlation over short distances allows us to introduce barriers, by reducing the exponential complexity of the field at given scales. This gives the upper bound. Second, this also gives the tree-like decorrelation .

Remark 9. We have avoided a neighborhood of $z = 0$ in these notes. The reason is that $z = 0$ implies that there is no hyperbolic regime, and in fact $\theta_\ell^z \sim \pi$ and thus is of order 1 for all ℓ. That is, the recursion, viewed at even times, essentially linearizes, even if we are at the elliptic regime!

4 Lecture IV: GβE and Wigner matrices, and their characteristic polynomial

For the (logarithm of) the determinant of random matrices, Tao-Vu proved a CLT even for Wigner matrices. In doing so, they first proved it for the GOE/GUE (using the tri-diagonal representation, see Remark 9. This begs the question, whether one can transfer some of the GβE results to the Wigner setup. Our methods in the last lecture are not really appropriate for that, since the three diagonal representation of Wigner matrices creates dependencies that are hard to control. Instead, we follow [BLZ25], which in turn is building on [BMP22]. Necessarily, here we will be mostly descriptive.

4.1 GβE and loop equations

We describe here the main input from [BMP22], which is slightly improved in [BLZ25]: a joint exponential moment for the log determinant at various points, when z is in the complex domain but only slightly above the real line. We focus here on the GβE, although the results apply to more general potentials, that is to measures

$$\frac{\mathrm{d}\mu_N(\lambda_1,\dots,\lambda_N)}{\mathrm{d}\lambda_1\dots\mathrm{d}\lambda_N} = \frac{1}{Z_N}\prod_{1\leqslant k<l\leqslant N}|\lambda_k-\lambda_l|^\beta \exp\left(-\frac{\beta N}{2}\sum_{k=1}^N V(\lambda_k)\right), \tag{4.1.1}$$

where $Z_N = Z_N(\beta)$ is a normalizing factor. We only use $V(x) = x^2$, and write μ forr the semicircle law of density $\rho(x) = c\sqrt{4-x^2}\mathbf{1}_{|x|\leqslant 2}$. Let γ_k be the kth (out of N) quantiles of the semi-circle law, i.e. $\int_0^{\gamma_k} \rho(x)dx = k/N$. Set

$$L_N(E) = \sum_{j=1}^{N} \log\left(E-\lambda_j\right) - N\int_{\mathbb{R}} \log\left(E-x\right) \mathrm{d}\mu(x), \tag{4.1.2}$$

which is the logarithm of the characteristic polynomial up to a centering shift.

Theorem 16. *[BLZ25] For any $\varepsilon, \kappa > 0$ we have*

$$\mathbb{P}\left(\sup_{2+\kappa<E<2-\kappa} \sqrt{\frac{\beta}{2}} \frac{\operatorname{Re} L_N(E)}{\log N} \in [1-\varepsilon, 1+\varepsilon]\right) = 1 - O(1),$$
$$\mathbb{P}\left(\sup_{2+\kappa<E<2-\kappa} \sqrt{\frac{\beta}{2}} \frac{\operatorname{Im} L_N(E)}{\log N} \in [1-\varepsilon, 1+\varepsilon]\right) = 1 - O(1).$$

Further,

$$\mathbb{P}\left(\max_{\kappa N \leqslant k \leqslant (1-\kappa)N} \pi\sqrt{\frac{\beta}{2}} \cdot \frac{\rho(\gamma_k)\, N\big(\lambda_k - \gamma_k\big)}{\log N} \in [1-\varepsilon, 1+\varepsilon]\right) = 1 - O(1). \tag{4.1.3}$$

Compare with Lecture III!

The main input needed for the proof of Theorem 16 is the following estimate from [BMP22, Remark 2.4]. For $z \in \mathbb{H}$, let

$$s(z) = s_N(z) = \frac{1}{N}\sum_{k=1}^{N} \frac{1}{\lambda_k - z}, \quad m(z) = \int_{\mathbb{R}} \frac{\mathrm{d}\mu(x)}{x-z}. \tag{4.1.4}$$

Theorem 17. *There exist constants $C, c, \tilde{\eta} > 0$ such that for any $q \geqslant 1$, $N \geqslant 1$ and $z = E + \mathrm{i}\eta$ with $0 < \eta \leqslant \tilde{\eta}$ and $-2-\eta \leqslant E \leqslant 2+\eta$, we have*

$$\mathbb{E}\big[\, |s(z) - m(z)|^q \,\big] \leqslant \frac{(Cq)^{q/2}}{(N\eta)^q}. \tag{4.1.5}$$

The proof of Theorem 17 follows a well trodden route, going back to Johansson [J98], with sharpening based on recursing the estimate while decreasing the imaginary part of z. As we will see, the fact that in the right hand side of (4.1.5) one has the term $q^{q/2}$ (and not q^q) is crucial, since we need to apply it with q of the order of $\log N$ and $\eta \sim \log N/N$.

We now describe the proof of Theorem 16, starting with the upper bound. By monotonicity of $\eta \mapsto \log|E + \mathrm{i}\eta - \lambda|$, $\eta > 0$, and the estimate $\int \log|E - \lambda|\, \mathrm{d}\rho(\lambda) = \int \log|E + \mathrm{i}\varepsilon - \lambda|\, \mathrm{d}\rho(\lambda) + O(\varepsilon)$ uniformly in E, there exists a fixed $C > 0$ such that

$$\sup_{E\in[-2+\kappa,2-\kappa]} \Re L_N(E) \leqslant \sup_{E\in[-2+\kappa,2-\kappa]} \Re L_N\left(E+\frac{\mathrm{i}}{N}\right)+C. \tag{4.1.6}$$

Let $J=[-2+\kappa,2-\kappa]\cap N^{-1-c}\mathbb{Z}$, where $c>0$ is an arbitrary small constant. For any $E\in[-2+\kappa,2-\kappa]$, let E' be the closest point in J, $z=E+\frac{\mathrm{i}}{N}$ and $z'=E'+\frac{\mathrm{i}}{N}$. Then from $\log(1+\varepsilon)=\varepsilon+O(\varepsilon^2)$ and recalling the definition of $s(z)$ from (4.1.4), this implies

$$\begin{aligned}\Re L_N(z)-\Re L_N(z') &= O((z-z')N(s(z')-m(z')))+O\left((z-z')^2\sum\frac{1}{|z'-\lambda_i|^2}\right)+O(1)\\ &= N^{-c}O(|s(z')-m(z')|)+N^{-2c}O(\Im s(z'))+O(1).\end{aligned} \tag{4.1.7}$$

Next, Theorem 17 (with $q=\log N$) together with Markov's inequality gives

$$\max_{E\in[-2+\kappa,2-\kappa]} \mathbb{P}\left(|s(z)-m(z)|>(\log N)^{7/10}\right)\leqslant N^{-200}. \tag{4.1.8}$$

for large enough N. Together with the boundedness of m on compact sets of $\mathbb{C}$, this gives

$$\mathbb{P}\left(\exists E'\in J: |s(z')|\geqslant(\log N)^{7/10}\right)\leqslant N^{-100}. \tag{4.1.9}$$

We conclude that

$$\mathbb{P}\left(\sup_{E\in[-2+\kappa,2-\kappa]}\Re L_N(E)\leqslant\sup_{E\in J}\Re L_N(E+\mathrm{i}N^{-1})+(\log N)^{9/10}\right)\geqslant 1-O(N^{-100}). \tag{4.1.10}$$

We now control the increments of L_N along the line segment $\{\Re z=E, N^{-1}<\Im z<\eta_0\}$ using Markov's inequality, where we set

$$\eta_0=\frac{(\log N)^{1000}}{N} \tag{4.1.11}$$

throughout this section. For $E\in J$, we denote $z=z(E)=E+\mathrm{i}/N$ and $\tilde{z}=E+\mathrm{i}\eta_0$. Then for any fixed $\varepsilon>0$ and $p\in\mathbb{N}$, we have by a union bound that

$$\begin{aligned}&\mathbb{P}\left(\exists E\in J:\Re L_N(z)>\Re L_N(\tilde z)+\varepsilon\log N\right)\\ &\qquad\leqslant CN^{1+c}(\varepsilon\log N)^{-2p}\\ &\times\max_{E\in J}\left(\int_{[N^{-1},\eta_0]^{2p}}\mathbb{E}\prod_{i=1}^{p}(N(s-m)(E+\mathrm{i}\eta_i))\prod_{i=p+1}^{2p}(N\overline{(s-m)}(E+\mathrm{i}\eta_i))\right)\mathrm{d}\eta_1\dots\mathrm{d}\eta_{2p}.\end{aligned} \tag{4.1.12}$$

We now suppose that $p=O(N(\log\log N)^{-1})$. Theorem 17 gives, for $E\in[-2+\kappa,2-\kappa]$,

$$\mathbb{E}[|(s-m)(E+\mathrm{i}\eta)|^p] \leqslant \frac{(Cp)^{p/2}}{(N\eta)^p} + C^p e^{-\tilde{c}N} \leqslant \frac{(Cp)^{p/2}}{(N\eta)^p}, \tag{4.1.13}$$

where the latter inequality holds because we assume $\eta < \eta_0$. Equation (4.1.13) and Hölder's inequality give

$$\mathbb{E}\left[\prod_{i=1}^{p} \big|N(s-m)(E+\mathrm{i}\eta_i)\big| \prod_{i=p+1}^{2p} |N\overline{(s-m)}(E+\mathrm{i}\eta_i)|\right] \leqslant (Cp)^p \prod_{i=1}^{2p} \frac{1}{\eta_i}. \tag{4.1.14}$$

Inserting the previous display in (4.1.12), we obtain

$$\begin{aligned}\mathbb{P}\left(\exists E \in J : \Re L_N(z) > \Re L_N(\tilde{z}) + \varepsilon \log N\right) &\leqslant \frac{N^{1+c}(Cp)^p(\log\log N)^{2p}}{(\varepsilon \log N)^{2p}} \\ &\leqslant N^{1+c}\frac{(Ap)^p(\log\log N)^{2p}}{(\log N)^{2p}} \leqslant N^{-100}, \end{aligned} \tag{4.1.15}$$

where A is a new constant depending on C and ε, and the latter inequality is obtained by setting $p = B\frac{\log N}{\log\log N}$ for sufficiently large B. We note that for the above reasoning, the Gaussian-like moment growth $(Cq)^{q/2}$ in Theorem 17 is crucial (as opposed to an exponential-like growth of $(Cq)^q$) .

Moreover, from Markov's inequality and an exponential moment computation *at a high enough imaginary part* η_0, for any fixed $\lambda > 0$ we have

$$\begin{aligned}&\mathbb{P}\left(\exists E \in J : \Re L_N(\tilde{z}) > (1+\varepsilon)\sqrt{\frac{2}{\beta}}\log N\right) \\ &\qquad \leqslant N^{1+c}\, e^{-\lambda(1+\varepsilon)\sqrt{\frac{2}{\beta}}\log N} \max_{E\in J} \mathbb{E}\big[e^{\lambda \Re L_N(\tilde{z})}\big] \\ &\qquad\qquad \leqslant C\, N^{1+c} \max_{E\in J} e^{\frac{\sigma(\lambda,0,\tilde{z})}{2} + \mu(\lambda,0,\tilde{z}) - \lambda(1+\varepsilon)\sqrt{\frac{2}{\beta}}\log N},\end{aligned}$$

for appropriate σ and μ satisfying $\mu(\lambda,0,\tilde{z}) = O(1)$ and $\sigma(\lambda,0,\tilde{z}) = (1+o(1))\lambda^2\frac{\log N}{\beta}$ uniformly in N, $E \in J$, and λ in any compact subset of $\mathbb{R}_+$. Choosing $\lambda = \sqrt{2\beta}$ this implies that

$$\mathbb{P}\left(\exists E \in J : \Re L_N(\tilde{z}) > (1+\varepsilon)\sqrt{\frac{2}{\beta}}\log N\right) \leqslant e^{-(2\varepsilon - c - \bar{(}1))\log N} \to 0. \tag{4.1.16}$$

The lower bound uses smoothing of the log determinant by moving into the complex plane, the evaluation of exponential moments, and a GMC method as in Lambert's lectures. We do not provide further details, except to note that this method cannot give the $\log\log N$ correction as in Lecture III!

Finally, the proof of (4.1.3) is just a restatement of the control of the imaginary part of the log-determinant, which is proved in a way similar to the real part.

4.2 Wigner matrices

We are now ready for the universality results.

Theorem 18. *[BLZ25] Let H be a symmetric Wigner matrix and set $d\nu(x) = \rho_{\rm sc}(x)\,\mathrm{d}x$ in (4.1.2). Then for any $\varepsilon, \kappa > 0$ we have*

$$\mathbb{P}\left(\sup_{|E|<2-\kappa} \frac{\operatorname{Re} L_N(E)}{\sqrt{2}\log N} \in [1-\varepsilon, 1+\varepsilon]\right) = 1 - \bar{(}1),$$

$$\mathbb{P}\left(\sup_{|E|<2-\kappa} \frac{\operatorname{Im} L_N(E)}{\sqrt{2}\log N} \in [1-\varepsilon, 1+\varepsilon]\right) = 1 - \bar{(}1).$$

The same result holds for Hermitian Wigner matrices after replacing the $\sqrt{2}$ factors with 1.

Remark 10. For the imaginary part of the logarithm, a similar estimate on the minimum holds, by considering the sup for the Wigner matrix $-H$:

$$\mathbb{P}\left(\inf_{|E|<2-\kappa} \frac{\operatorname{Im} L_N(E)}{\sqrt{2}\log N} \in [-1-\varepsilon, -1+\varepsilon]\right) = 1 - \bar{(}1).$$

No such statement holds for the real part, as $\inf_{|E|<2-\varepsilon} \operatorname{Re} L_N(E) = -\infty$.

For Gaussian-divisible Wigner matrices, universality actually holds up to tightness.

Theorem 19. *Let H be a Gaussian-divisible symmetric Wigner matrix, that is $H = \sqrt{1-\varepsilon^2}H' + \varepsilon G$, where H' is a Wigner matrix independent of a GOE matrix G. Then for any $\kappa > 0$, there exists a coupling between H and a GOE such that the following sequence of random variables is tight:*

$$\Big(\sup_{|E|<2-\kappa} \Re L_N^H(E) - \sup_{|E|<2-\kappa} \Re L_N^{\rm GOE}(E)\Big)_{N\geqslant 1}.$$

In particular, in view of the results in Lecture 3, the loglog correction term is as in the GOE case!

The idea behind the proof of Theorems 18 and 19, like that of many universality results, is to first provide the proof for a modified (Gaussian divisible with ε goind to 0 with N) model, and then use moment matching to get rid of the regularization. We only discuss briefly the first part, which builds strongly on [B22], which in turns builds on the dynamical approach of Erdős, Schlein and Yau, and on [EYY12].

Before discussing the proof, we need one more a-priori preliminary rigidity result.

Theorem 20. *[EYY12] Let H be a Wigner matrix. Then there exists $C_0 > 0$ such that the following claims hold.*

1. There exists $c>0$ such that

$$\mathbb{P}\left(\bigcup_{z\in\mathbb{H}}\left\{|m_N(z)-m(z)|\geqslant\frac{\varphi}{N\eta}\right\}\right)\leqslant c^{-1}\exp\left(-\varphi^c\right) \tag{4.2.17}$$

and

$$\mathbb{P}\left(\bigcup_{z\in\mathbb{H}}\left\{\max_{i,j\in[\![1,N]\!]}|G_{ij}(z)-\delta_{ij}(z)|\geqslant\varphi\sqrt{\frac{\Im m(z)}{N\eta}}+\frac{\varphi}{N\eta}\right\}\right)\leqslant c^{-1}\exp\left(-\varphi^c\right). \tag{4.2.18}$$

2. There exists $c>0$ such that, defining $\hat{k}=\min(k,N+1-k)$,

$$\mathbb{P}\left(\exists k\in[\![1,N]\!]:|\lambda_k-\gamma_k|\geqslant\varphi\,\hat{k}^{-\frac{1}{3}}N^{-\frac{2}{3}}\right)\leqslant c^{-1}\exp\left(-\varphi^c\right). \tag{4.2.19}$$

We next provide a quantitative relaxation of the eigenvalues (Proposition 1), which is a variant of [B22, Theorem 3.1] and relies on this work. Let H be a Wigner matrix. We first recall the definition of Dyson Brownian motion with initial data $H_0=H$.

Let B be a symmetric matrix such that the entries $\{B_{ij}\}_{i<j}$ and $B_{ii}/\sqrt{2}$ are independent standard Brownian motions, and $B_{ij}=B_{ji}$. Consider the matrix Ornstein–Uhlenbeck process

$$\mathrm{d}H_t=\frac{1}{\sqrt{N}}\mathrm{d}B_t-\frac{1}{2}H_t\,\mathrm{d}t. \tag{4.2.20}$$

If the eigenvalues of H_0 are distinct, it is well known that the eigenvalues $(\lambda_1(t),\lambda_2(t),\ldots,\lambda_N(t))$ of H_t are given by the strong solution of the system of stochastic differential equations

$$\mathrm{d}\lambda_k=\frac{\mathrm{d}\beta_k}{\sqrt{N}}+\left(\frac{1}{N}\sum_{\ell\neq k}\frac{1}{\lambda_k-\lambda_\ell}-\frac{1}{2}\lambda_k\right)\mathrm{d}t, \tag{4.2.21}$$

where the $\{\beta_k\}_{k=1}^N$ are independent, standard Brownian motions. (See, for example, [AGZ10, Lemma 4.3.3].)

We now let $(\mu_1(t),\mu_2(t),\ldots,\mu_N(t))$ be a strong solution of the same SDE (4.2.21) with initial condition $(\mu_1,\mu_2,\ldots,\mu_N)$, where $\{\mu_k\}_{k=1}^N$ are the eigenvalues of a GOE:

$$\mathrm{d}\mu_k=\frac{\mathrm{d}\beta_k}{\sqrt{N}}+\left(\frac{1}{N}\sum_{\ell\neq k}\frac{1}{\mu_k-\mu_\ell}-\frac{1}{2}\mu_k\right)\mathrm{d}t.$$

For any $z\in\mathbb{H}$, we define

$$z_t=\frac{e^{t/2}(z+\sqrt{z^2-4})+e^{-t/2}(z-\sqrt{z^2-4})}{2}, \tag{4.2.22}$$

where $\sqrt{z^2-4}$ is defined using a branch cut in the segment $[-2,2]$. For $z \in \mathbb{R}$, we define $z_t = \lim_{\eta\to 0^+}(z+\mathrm{i}\eta)_t$. Set

$$\varphi = \exp\left(C_0(\log\log N)^2\right), \tag{4.2.23}$$

The following key estimate on the difference between $\lambda(t)$ and $\mu(t)$ follows from the main result in [B22]. Let L^H and L^{GOE} denote the observable (4.1.2) defined using the eigenvalues of H and GOE, respectively.

Proposition 1. *Fix $\kappa, \varepsilon > 0$. Then for any $D > 0$ there exist $C(\varepsilon,\kappa,D) > 0$ such that for all $t \in (\varphi^C/N, 1)$, $E \in [-2+\kappa, 2-\kappa]$, and $k \in [\![1,N]\!]$ such that $\gamma_k \in [-2+\kappa, 2-\kappa]$, we have*

$$\mathbb{P}\left(\left|\lambda_k(t)-\mu_k(t)-\frac{\Im L_N^H(E_t)-\Im L_N^{\mathrm{GOE}}(E_t)}{N\operatorname{Im} m(E_t)}\right| > \frac{N^{1+\varepsilon}\max(|E-\gamma_k|,N^{-1})}{N^2 t}\right) \leqslant CN^{-D}. \tag{4.2.24}$$

Proof. The key to the proof is [B22, Theorem 3.1], which states that there exists $C(D) > 0$ such that

$$\mathbb{P}\left(\left|\left(\lambda_k(t)-\mu_k(t)\right)-\bar{\mathrm{u}}_k(t)\right| > \frac{N^{\varepsilon}}{N^2 t}\right) \leqslant CN^{-D} \tag{4.2.25}$$

for $t \in (\varphi^C/N, 1)$, where we define

$$\bar{\mathrm{u}}_k(t) = \frac{1}{N\operatorname{Im} m(\gamma_k^t)}\sum_{j=1}^{N}\operatorname{Im}\left(\frac{1}{\gamma_j-\gamma_k^t}\right)\left(\lambda_j(0)-\mu_j(0)\right), \qquad \gamma_k^t = (\gamma_k)_t. \tag{4.2.26}$$

Moreover, from [B22, Lemma 3.4], for all $\gamma_k, \gamma_\ell \in [-2+\kappa, 2-\kappa]$ we have

$$\mathbb{P}\left(|\bar{\mathrm{u}}_k(t)-\bar{\mathrm{u}}_\ell(t)| \geqslant C\varphi\frac{|k-\ell|}{N^2 t}\right) \leqslant CN^{-D}. \tag{4.2.27}$$

Let $E \in [-2+\kappa, 2-\kappa]$ be given, and fix some $\ell = \ell(E,N)$ such that $|E-\gamma_\ell| = \min_{j\in[\![1,N]\!]}|E-\gamma_j|$. The definition of γ_k gives

$$|k-\ell| < CN|\gamma_k-\gamma_\ell| \leqslant CN\left(|\gamma_k-E|+|E-\gamma_\ell|\right) \leqslant 2CN|\gamma_k-E|, \tag{4.2.28}$$

for some constant $C > 0$. Then equations (4.2.25) and (4.2.27) together with the previous line imply that

$$\mathbb{P}\left(\left|\left(\lambda_k(t)-\mu_k(t)\right)-\bar{\mathrm{u}}_\ell(t)\right| > \frac{CN^{1+\varepsilon}\max(|E-\gamma_k|,N^{-1})}{N^2 t}\right) \leqslant CN^{-D}, \tag{4.2.29}$$

where we increased C if necessary and used $N^{\varepsilon} \geqslant \phi$ for sufficiently large N (depending on ε). We therefore just need to bound $\left|\frac{\Im L_N^H(E_t)-\Im L_N^{\mathrm{GOE}}(E_t)}{N\operatorname{Im} m(E_t)} - \bar{\mathrm{u}}_\ell(t)\right|$. We write

$$\frac{\Im L_N^H(E_t) - \Im L_N^{\mathrm{GOE}}(E_t)}{N \operatorname{Im} m(E_t)} = \frac{1}{N \operatorname{Im} m(E_t)} \Im \sum_{j=1}^{N} \log\left(1 + \frac{\lambda_j(0) - \mu_j(0)}{\mu_j(0) - E_t}\right). \tag{4.2.30}$$

On the rigidity event from (4.2.19), a Taylor expansion of the logarithm gives, with overwhelming probability,

$$\Im \sum_{j=1}^{N} \log\left(1 + \frac{\lambda_j(0) - \mu_j(0)}{\mu_j(0) - E_t}\right) = \Im \sum_{j=1}^{N} \frac{\lambda_j(0) - \mu_j(0)}{\mu_j(0) - E_t} + O\left(\sum_{j=1}^{N} \left|\frac{\lambda_j(0) - \mu_j(0)}{\mu_j(0) - E_t}\right|^2\right). \tag{4.2.31}$$

For the error term, on the rigidity event from (4.2.19) we can write

$$\sum_{j=1}^{N} \left|\frac{\lambda_j(0) - \mu_j(0)}{\mu_j(0) - E_t}\right|^2 \leqslant \frac{C\phi^2}{Nt} \cdot \frac{1}{N} \sum_{j=1}^{N} \frac{\Im E_t}{|\mu_j(0) - E_t|^2} = \frac{C\phi^2}{Nt} \operatorname{Im} m_N(E_t) \leqslant \frac{C\phi^2}{Nt}, \tag{4.2.32}$$

where we used $ct \leqslant \Im E_t \leqslant Ct$ to bound $\operatorname{Im} m_N(E_t) \leqslant C$ using (4.2.17). This estimate on $\operatorname{Im} m_N(E_t)$ also shows that the second term in (4.2.31) is negligible when inserted in (4.2.30).

Finally, we need to bound

$$\begin{aligned}
&\frac{1}{N \Im m(E_t)} \Im \sum_{j=1}^{N} \frac{\lambda_j(0) - \mu_j(0)}{\mu_j(0) - E_t} - \bar{\mathrm{u}}_\ell(t) \\
&= \frac{1}{N} \sum_{j=1}^{N} (\lambda_j(0) - \mu_j(0)) \left(\frac{1}{\Im m(E_t)} - \frac{1}{\Im m(\gamma_\ell^t)}\right) \operatorname{Im} \frac{1}{\mu_j(0) - E_t} \\
&+ \frac{1}{N} \sum_{j=1}^{N} \frac{\lambda_j(0) - \mu_j(0)}{\Im m(\gamma_\ell^t)} \operatorname{Im}\left(\frac{1}{\mu_j(0) - E_t} - \frac{1}{\gamma_j - \gamma_\ell^t}\right).
\end{aligned}$$

For the first sum, from $|\operatorname{Im} m(E_t) - \operatorname{Im} m(\gamma_\ell^t)| \leqslant C|E_t - \gamma_\ell^t| \leqslant CN^{-1}$, $\operatorname{Im} m(E_t) \geqslant c$, and $\operatorname{Im} m(\gamma_\ell^t) \geqslant c$, on the rigidity event from (4.2.19) we obtain

$$\frac{1}{N} \sum_{j=1}^{N} |\lambda_j(0) - \mu_j(0)| \left(\frac{1}{\Im m(E_t)} - \frac{1}{\Im m(\gamma_\ell^t)}\right) \operatorname{Im} \frac{1}{\mu_j(0) - E_t} \leqslant \frac{C\varphi}{N^3} \cdot \sum \frac{t}{|\gamma_j - E_t|^2} \leqslant \frac{C\varphi}{N^2}.$$

On the same rigidity event, the second sum is bounded by

$$\frac{1}{N} \sum_{j=1}^{N} |\mu_j(0) - \lambda_j(0)| \left|\operatorname{Im} \frac{E_t - \gamma_\ell^t}{(\mu_j(0) - E_t)(\gamma_j - \gamma_\ell^t)}\right| \leqslant \frac{C\varphi}{N^3} \sum_j \left(\frac{1}{|\mu_j(0) - E_t|^2} + \frac{1}{|\gamma_j - \gamma_\ell^t|^2}\right) \leqslant \frac{C\varphi}{N^2 t}.$$

We have thus obtained

$$\left|\frac{\Im L_N^H(E_t) - \Im L_N^{\mathrm{GOE}}(E_t)}{N \operatorname{Im} m(E_t)} - \bar{\mathrm{u}}_\ell(t)\right| \leqslant C \frac{\varphi^2}{N^2 t}, \tag{4.2.33}$$

which concludes the proof.

The relaxation of L_N is an a priori more intricate problem as $\Re L_N$ depends on the full spectrum, however the analysis folows the same lines, with appropriate discretizations and union bounds. The details are in [BLZ25].

References

[ABR08] L. Addario-Berry and B. Reed, Ballot theorems for random walks with finite variance. arXiv:0802.2491v2. (Updated version: http://www.math.mcgill.ca/louigi/papers/ballot.pdf; published version: Bolyai Mathematical Studies, 17:9–35, 2008, Springer).

[ABR09] L. Addario-Berry and B. Reed, Minima in branching random walks, *Annals Probab.* 37:1044–1079, 2009.

[Aïd13] Elie Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.*, 41(3A):1362–1426, 2013.

[AGZ10] G. W. Anderson, A. Guionnet and O. Zeitouni, An introduction to random matrices, Cambridge University Press, 2011.

[ABBS13] E. Aïdékon, M. Brunnet, J. Berestycki and Z. Shi, Branching Brownian motion seen from its tip, PTRF 157:405–451 (2013).

[ABK11] L.-P. Arguin, A. Bovier and N. Kistler, The extremal process of branching Brownian motion, PTRF 157:535–574 (2013).

[AGG89] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[ABZ23] F. Augeri, R. Butez and O. Zeitouni, A CLT for the characteristic polynomial of random Jacobi matrices, and the GβE, *PTRF* 186:1–89 (2023).

[AZ25] F. Augeri and O. Zeitouni, Maximum of the characteristic polynomial of random Jacobi matrices, in preparation, 2025.

[Bis20] Marek Biskup. Extrema of the two-dimensional discrete Gaussian free field. In *Random graphs, phase transitions, and the Gaussian free field*, volume 304 of *Springer Proc. Math. Stat.*, pages 163–407. Springer, Cham, 2020.

[BL13] M. Biskup and O. Louidor, Extreme local extrema of two-dimensional discrete Gaussian free field, *Comm. Math. Phys.* 345:271–304, 2016.

[BDG01] E. Bolthausen, J.-D. Deuschel and G. Giacomin, Entropic repulsion and the maximum of the two-dimensional harmonic crystal, *Annals Probab.* 29:1670–1692, 2001.

[B22] P. Bourgade, Extreme gaps between eigenvalues of Wigner matrices, *J. Eur. Math. Soc.* 8:2823–2873, 2022.

[BMP22] P. Bourgade, K. Mody and M. Pain. Optimal local law and central limit theorem for β-ensembles. *Comm. Math. Phys.* 390:1017–1079, 2022.

[BLZ25] P. Bourgade, P. Lopatto and O. Zeitouni, Optimal Rigidity and Maximum of the Characteristic Polynomial of Wigner Matrices, *Geometric and Functional Analysis* 35(1):161-253, 2025.

[Br78] M. Bramson, Maximal displacement of branching Brownian motion, *Comm. Pure Appl. Math.* 31:531–581, 1978.

[BDZ16] M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 69(1):62–123, 2016.

[BZ11] M. Bramson and O. Zeitouni, Tightness of the recentered maximum of the two–dimensional discrete Gaussian Free Field, *Comm. Pure App. Math* 65:1–20, 2011.

[CMN18] Reda Chhaibi, Thomas Madaule, and Joseph Najnudel. On the maximum of the CβE field. *Duke Mathematical Journal*, 167(12):2243–2345, 2018.

[DZ98] A. Dembo and O. Zeitouni, *Large deviation techniques and applications*, 2nd edition, Springer, New-York, 1998.

[DRZ17] J. Ding, R. Roy, and O. Zeitouni. Convergence of the centered maximum of log-correlated Gaussian fields. *Ann. Probab.*, 45(6A):3886–3928, 2017.

[EYY12] L. Erdős, L., H.-T. Yau and J. Yin, Rigidity of eigenvalues of generalized Wigner matrices, *Adv. Math.*, 229:1435–1515, 2012.

[FB08] Yan V. Fyodorov and Jean-Philippe Bouchaud. Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A*, 41(37):372001, 12, 2008.

[FHK12a] Yan V. Fyodorov, Gaith A. Hiary, and Jonathan P. Keating. Freezing transition, characteristic polynomials of random matrices, and the Riemann Zeta function. *Phys. Rev. Lett.*, 108:170601, 2012.

[FHK12b] Yan V. Fyodorov, Ghaith A. Hiary, and Jonathan P. Keating. Freezing Transition, Characteristic Polynomials of Random Matrices, and the Riemann Zeta-Function. *Phys. Rev. Lett.*, 108:170601, 2012.

[FS16] Y. V. Fyodorov and N. Simms, One the distribution of the maximum value of the characteristic polynomial of GUS random matrices, *Nonlinearity* 29:2837-2855, 2016.

[J98] K. Johansson, On fluctuations of eigenvalues of random Hermitian matrices, *Duke Math. J.* 91:151–204, 1998.

[KN04] Rowan Killip and Irina Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.*, 50:2665–2701, 2004.

[LS87] S. P. Lalley and T. Sellke, A conditional limit theorem for the frontier of a branching Brownian motion, *Annals Probab.* 15:1052–1061, 1987.

[LN24] G. Lambert and J. Najnudel, Subcritical multiplicative chaos and the characteristic polynomial of the CβE, arXiv:2407.19817, 2024.

[LP23] G. Lambert and E. Paquette. Strong approximation of Gaussian beta-ensemble characteristic polynomials: the hyperbolic regime, *Annals Appl. Prob.* 33:549–612, 2023.

[LP20] G. Lambert and E. Paquette. Strong approximation of Gaussian β-ensemble characteristic polynomials: the edge regime and the stochastic Airy function. arXiv:2009.05003, 2020.

[Li78] T. Liggett, Random invariant measures for Markov chains and independent particle systems, *Z. Wahr. verw. geb.* 45:297–313, 1978.

[Ma17] T. Madaule, Convergence in law for the branching random walk seen from its tip, *J. Theor. Prob.* 30:27–63, 2017.

[PZ25] E. Paquette and O. Zeitouni, The extremal landscape for the CβE ensemble, *Forum of Math. - Sigma* (2025).

[Rem20] Guillaume Remy. The Fyodorov-Bouchaud formula and Liouville conformal field theory. *Duke Math. J.*, 169(1):177–211, 2020.

[Rob13] Matthew I. Roberts. A simple path to asymptotics for the frontier of a Branching Brownian Motion. *Ann. Probab.*, 41(5):3518–3541, 2013.

[Sim04] Barry Simon. *Orthogonal Polynomials on the Unit Circle, Parts I, II.* American Math. Society, Providence, RI, 2004.

[SZ15] Eliran Subag and Ofer Zeitouni. Freezing and decorated Poisson point processes. *Comm. Math. Phys.*, 337(1):55–92, 2015.

[TV12] T. Tao and V. Vu, A central limit theorem for the determinant of a Wigner matrix, *Adv. Math.* 231:74–101, 2012.