Invent math (2009) 177: 463–508
DOI 10.1007/s00222-009-0180-z

Continuum limits of random matrices and the Brownian carousel

Benedek Valkó · Bálint Virág

Received: 29 February 2008 / Accepted: 26 January 2009 / Published online: 27 February 2009
© Springer-Verlag 2009

Abstract We show that at any location away from the spectral edge, the eigenvalues of the Gaussian unitary ensemble and its general β siblings converge to Sine_β, a translation invariant point process. This process has a geometric description in term of the Brownian carousel, a deterministic function of Brownian motion in the hyperbolic plane.

The Brownian carousel, a description of the a continuum limit of random matrices, provides a convenient way to analyze the limiting point processes. We show that the gap probability of Sine_β is continuous in the gap size and β, and compute its asymptotics for large gaps. Moreover, the stochastic differential equation version of the Brownian carousel exhibits a phase transition at $\beta = 2$.

The Brownian carousel and the winding angle α_λ

B. Valkó
Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53705, USA
e-mail: valko@math.wisc.edu

B. Virág (✉)
Department of Mathematics, University of Toronto, Toronto, ON M5S 2E4, Canada
e-mail: balint@math.toronto.edu

Contents

Abstract . . . 463
1 Introduction . . . 465
2 The Brownian carousel and the stochastic sine equation . . . 467
2.1 Definitions . . . 467
2.2 Properties of the Brownian carousel . . . 469
2.3 Large gap probabilities . . . 471
2.4 A phase transition at $\beta = 2$ . . . 472
3 Breakdown of the proof of Theorem 1 . . . 474
4 The hyperbolic description of the phase evolution . . . 478
4.1 The hyperbolic point of view . . . 478
4.2 Phase evolution equations . . . 479
4.3 Slowly varying phase evolution for a scaling window . . . 481
4.4 The discrete carousel . . . 484
5 The stochastic sine equation as a limit . . . 485
5.1 Single-step asymptotics . . . 485
5.2 Continuum limit of the phase evolution . . . 488
5.3 Why are the limits in different windows the same? Universality and non-universality . . . 491
6 Asymptotics in the uneventful stretch . . . 494
6.1 The uneventful middle stretch . . . 494
6.2 Bounds for oscillations in the middle stretch . . . 498
6.3 Why does the right boundary condition disappear? . . . 500
6.4 The uneventful ending . . . 501
Acknowledgements . . . 503
Tools . . . 503
A.1 Angular shift bounds . . . 503
A.2 Oscillatory sums . . . 504
A.3 A limit theorem for random difference equations . . . 506
References . . . 507

1 Introduction

The Gaussian orthogonal and unitary ensembles are the most fundamental objects of study in random matrix theory. In the past decades, their eigenvalue distribution has shown to be important in several areas of probability, combinatorics, number theory, operator algebras, even engineering (see [1] for an overview). For dimension n, the ordered eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n \in \mathbb{R}$ have joint density

$$\frac{1}{Z_{n,\beta}} e^{-\beta \sum_{k=1}^{n} \lambda_k^2/4} \prod_{j<k} |\lambda_j - \lambda_k|^\beta, \tag{1}$$

where $\beta = 1, 2$ for the Gaussian orthogonal and unitary ensembles, respectively. The above density makes sense for any $\beta \geq 0$, and the point process is often called Coulomb gas in Gaussian potential at inverse temperature β. The goal of this paper is to study its $n \to \infty$ point process limit away from the spectral edge.

The limit is described via a special case of the *hyperbolic carousel*. Let

- b be a path in the hyperbolic plane
- z be a point on the boundary of the hyperbolic plane, and
- $f : \mathbb{R}_+ \to \mathbb{R}_+$ be an integrable function.

To these three objects, the hyperbolic carousel associates a multi-set of points on the real line defined via its counting function $N(\lambda)$ taking values in $\mathbb{Z} \cup \{-\infty, \infty\}$. As time increases from 0 to ∞, the boundary point z is rotated about the center $b(t)$ at angular speed $\lambda f(t)$. $N(\lambda)$ is defined as the integer-valued total winding number of the point about the moving center of rotation.

The *Brownian carousel* is defined as the hyperbolic carousel driven by hyperbolic Brownian motion b. See Sect. 2 for more details.

In order to study the $n \to \infty$ limit of (1) we need to pick the center μ_n of the scaling window for each n. Then the scaling factor follows the Wigner semicircle law. Our main theorem gives necessary and sufficient conditions on μ_n to get a bulk-type limit.

Theorem 1 *For $\beta > 0$, let Λ_n denote the point process given by (1), and let μ_n be a sequence so that $n^{1/6}(2\sqrt{n} - |\mu_n|) \to +\infty$. Then*

$$\sqrt{4n - \mu_n^2}(\Lambda_n - \mu_n) \Rightarrow \mathrm{Sine}_\beta,$$

where Sine_β is the discrete point process given by the Brownian carousel with parameters $f(t) = (\beta/4)e^{-\beta t/4}$ and arbitrary z.

The convergence here is in law with respect to vague topology for the counting measure of the point process. The limit and convergence for the special values $\beta = 1, 2, 4$ under more restrictive scaling conditions has been well-studied, see [14] or [8]. The Brownian carousel description is novel even in these special cases. We note that the ensemble (1) may be generalized by replacing the $\sum_k \lambda_k^2$ in the exponent by a similar sum involving a fixed function V of the eigenvalues. Assuming certain growth conditions on V the corresponding problem in the $\beta = 2$ case can be treated using orthogonal polynomials and Riemann-Hilbert methods, see e.g. [1, 3].

Together with the following theorem, Theorem 1 gives a complete characterization of the possible limiting processes for the ensembles (1).

Theorem 2 (Ramírez, Rider, and Virág [15]) *For $\beta > 0$, let Λ_n denote the point process given by* (1), *and let μ_n be a sequence so that $n^{1/6}(2\sqrt{n} - \mu_n) \to a \in \mathbb{R}$. Then*

$$n^{1/6}(\Lambda_n - \mu_n) \Rightarrow \mathrm{Airy}_\beta + a.$$

Here Airy_β is defined as -1 times the point process of eigenvalues of the stochastic Airy operator, see Ramírez, Rider, and Virág [15] for more details. A straightforward diagonalization argument gives the following corollary, which is proved in Sect. 3.

Corollary 3 *As $a \to \infty$ we have $2\sqrt{a}(\mathrm{Airy}_\beta + a) \Rightarrow \mathrm{Sine}_\beta$.*

The proof of Theorem 1 is based on the tridiagonal matrix models introduced by Trotter [19] and Dumitriu and Edelman [4]. Sutton [18] and Edelman and Sutton [6] present heuristics that the operators given by the tridiagonal matrices have a limit whose eigenvalues give the Sine and Airy processes. Theorem 2 shows that this is indeed the case at the spectral edge. The bulk case, however, is fundamentally different: there seems to be no natural limiting operator with the spectrum given by the Sine point process. Rather than taking a limit of the operator itself, we consider limits of discrete variants of the phase functions in the Sturm-Liouville theory. This connection is explored further in Sect. 5.3, where we describe how the Sine point process appears as a universal limit for a large class of one-dimensional Schrödinger operators.

The eigenvalue equation of a real tridiagonal matrix gives a three-term linear recursion for the eigenvectors. This becomes a two-term recursion for the ratios of consecutive entries, which then evolves by linear fractional transformations fixing the real line. So in our case, the evolution operators perform a time-inhomogeneous random process in $\mathrm{PSL}(2, \mathbb{R})$, the group of orientation-preserving isometries of the hyperbolic plane. To get the Brownian carousel, we regularize this evolution and take limits. An important tool is Proposition 23 (based on the results of Stroock and Varadhan [17]), which yields stochastic differential equation limits of Markov processes with heavy local oscillations.

The Brownian carousel description gives a simple way to analyze the limiting point process. The hyperbolic angle of the rotating boundary point as measured from $b(t)$ follows the *stochastic sine equation*, a coupled one-parameter family of stochastic differential equations

$$d\alpha_\lambda = \lambda f\, dt + \mathrm{Re}((e^{-i\alpha_\lambda} - 1)dZ), \qquad \alpha_\lambda(0) = 0, \tag{2}$$

driven by a two-dimensional standard Brownian motion. For a single λ, this reduces to the one-dimensional stochastic differential equation

$$d\alpha_\lambda = \lambda f\, dt + 2\sin(\alpha_\lambda/2)dW, \qquad \alpha_\lambda(0) = 0, \tag{3}$$

which converges as $t \to \infty$ to an integer multiple $\alpha_\lambda(\infty)$ of 2π. A direct consequence of the definition of Sine_β is the following.

Proposition 4 *The number of points $N(\lambda)$ of the point process Sine_β in $[0, \lambda]$ has the same distribution as $\alpha_\lambda(\infty)/(2\pi)$.*

Convergence to the solution of the coupled SDEs is the result formally announced in the lecture by Virág [20]. In independent work, Killip and Stoiciu [13] present a related but different description of the limit processes in the setting of circular ensembles (see, e.g. [8], Chap. 2 for discussion of these models and [12] for further related results).

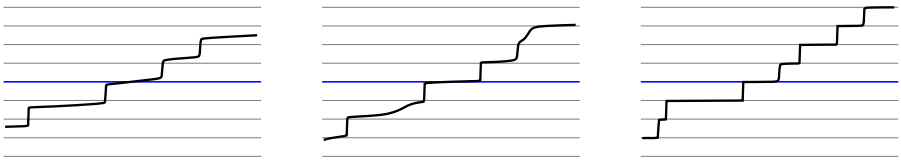

Fig. 1 The $\beta = 1$ stochastic sine equation as a function of λ at three times

Proposition 4 allows us to analyze the point process Sine_β, for example to determine the asymptotics of large gap probabilities. This has been predicted by Dyson [5] and proved for the cases $\beta = 2$ by Widom [21] and for $\beta = 1, 4$ by Jimbo et al. [10]; there, more refined asymptotics are presented; see also Deift et al. [2].

Theorem 5 *For $k \geq 0$ fixed and $\lambda \to \infty$, we have*

$$\mathrm{P}(\#\textit{ of points in } [0, \lambda] \leq k) = \exp\bigl(-\lambda^2(\beta/64 + o(1))\bigr).$$

This is shown in Sect. 2.3 for the case of a more general parameter f. Several similar asymptotic identities can be computed this way, and continuity properties can be studied. For the Sine_β processes we have

Proposition 6 *The probability distribution of $N(\lambda)$ is a continuous function of λ and β.*

In contrast, the stochastic sine equation exhibits a phase transition at $\beta = 2$.

Theorem 7 *For any $\lambda > 0$ we have a.s.*

$$\textit{for all } t \textit{ large enough} \quad \alpha_\lambda(t) \geq \alpha_\lambda(\infty) \tag{4}$$

if and only if $\beta \leq 2$. In particular, the probability of the event (4) is not analytic at $\beta = 2$ as a function of β.

Deift (personal communication, 2007) asked whether this phase transition also appears on the level of gap probabilities. This question remains open.

2 The Brownian carousel and the stochastic sine equation

2.1 Definitions

In the Poincaré disk model of the hyperbolic plane a boundary point can be described by an angle. The *Brownian carousel ODE* with parameters $f(t)$ and z_0 describes the evolution of the lifted angle $\gamma_\lambda(t)$ with $e^{i\gamma_\lambda(0)} = z_0$ as it is rotated about the center $B(t)$ at angular speed $\lambda f(t)$. Here $B(t)$ is hyperbolic Brownian motion, that is the strong solution of the SDE

$$dB = \frac{(1 - |B|^2)}{2}\, d\tilde{Z}$$

driven by complex Brownian motion $\tilde{Z}$ with standard real and imaginary parts. The speed of γ_λ, as measured in units of boundary harmonic measure from B, is $\lambda f/(2\pi)$. To change to an angle measured from 0, we need to divide by the Poisson kernel

$$\mathrm{Poi}(e^{i\gamma_\lambda}, w) = \frac{1}{2\pi}\,\mathrm{Re}\,\frac{e^{i\gamma_\lambda}+w}{e^{i\gamma}-w} = \frac{1}{2\pi}\,\frac{1-|w|^2}{|e^{i\gamma_\lambda}-w|^2},$$

which yields the ODE

$$\partial_t \gamma_\lambda = \frac{\lambda f}{2\pi\,\mathrm{Poi}(e^{i\gamma_\lambda}, B)} = \lambda f\,\frac{|e^{i\gamma_\lambda}-B|^2}{1-|B|^2}. \tag{5}$$

The most convenient way to define the winding number $N(\lambda)$ of $e^{i\gamma_\lambda}$ about the moving center of rotation $B(t)$ is to follow the corresponding angle. Let $\alpha_\lambda(t)$ denote the hyperbolic angle determined by the points z_0, $B(t)$ and $e^{i\gamma_\lambda(t)}$. As we will check at the end of this section, Itô's formula shows that α satisfies the *stochastic sine equation* (2), i.e.

$$d\alpha_\lambda = \lambda f\,dt + \mathrm{Re}((e^{-i\alpha_\lambda}-1)dZ), \qquad \alpha_\lambda(0)=0, \tag{6}$$

where dZ is simply complex white noise with standard real and imaginary parts. The name of the SDE comes from the fact that the last term equals $2\sin(\alpha_\lambda/2)\,\mathrm{Im}(e^{-i\alpha_\lambda/2}dZ)$. Since $dW = \mathrm{Im}(e^{-i\alpha_\lambda/2}dZ)$ is 1-dimensional white noise, we get the SDE (3) for the single λ marginals.

Proposition 9 of the next section shows that

$$\frac{1}{2\pi}\lim_{t\to\infty}\alpha_\lambda(t) \tag{7}$$

exists for every λ a.s. and for every $\lambda_1<\lambda_2$ a.s. $N(\lambda_1)\le N(\lambda_2)$. Thus $N(\lambda)$ can be defined as the unique random right-continuous function which agrees with (7) for every λ a.s.

To deduce (6), let $\mathcal{T}(w,z)$ denote the Möbius automorphism of the unit disk taking z_0 to 1 and taking w to 0. It is given by the formula

$$\mathcal{T}(w,z) = \frac{S(w,z)}{S(w,z_0)}, \quad S(w,z) = \frac{z-w}{1-\overline{w}z}. \tag{8}$$

Then α is defined as the continuous solution of

$$\alpha(0)=0, \qquad e^{i\alpha(t)} = \mathcal{T}(B_t, e^{i\gamma(t)}). \tag{9}$$

The stochastic sine equation (6) follows from taking logarithms and applying Itô's formula. For the driving Brownian motion we get the explicit expression

$$dZ = 2\partial_2\mathcal{T}(B,B)dB = \frac{2}{1-|B|^2}\,\frac{1-\overline{B}}{1-B}\,dB.$$

Remark 8 By Itô's formula applied to the logarithm of (9), the noise term in (6) can be interpreted as the infinitesimal movement of the angle α under the difference of transformations $d\mathcal{T} = \mathcal{T}(B+dB,\gamma)\mathcal{T}(B,\gamma)^{-1}$. This infinitesimal Möbius transformation $d\mathcal{T}$ moves 0 to $\mathcal{T}(B,B+dB) = \partial_2\mathcal{T}(B,B)dB$, a standard complex Brownian motion increment. Such a transformation $d\mathcal{T}$ changes the angle of any two points on the boundary by a Brownian increment with standard deviation proportional to their distance. This gives a more conceptual explanation of the noise term in (6).

2.2 Properties of the Brownian carousel

Let L^1_* denote the set of absolutely integrable functions of $\mathbb{R}^+$ which tend to 0 at $+\infty$. Given a hyperbolic Brownian motion and a boundary point z_0, the Brownian carousel associates a random counting function $N(\lambda)$ to each $f \in L^1_*$. More generally, it is fruitful to study how $N(\lambda)$ changes when the parameter f varies but the Brownian path remains fixed. In this case λ can be absorbed in the parameter f so we will use the notation $N_f = N_f(1)$, and α_f for the case $\lambda = 1$.

Proposition 9 (Properties of the Brownian carousel) *We have*

(i) $\alpha_f - \alpha_g$ *has the same distribution as* α_{f-g},
(ii) $\alpha_f(t)$ *is increasing in* f,
(iii) $\lfloor \alpha_f(t) \rfloor_{2\pi}$ *is nondecreasing in* t *when* $f \geq 0$. *Here* $\lfloor x \rfloor_{2\pi} = \max(2\pi\mathbb{Z} \cap (-\infty, x])$,
(iv) $N_f = \frac{1}{2\pi} \lim_{t\to\infty} \alpha_f(t)$ *exists and is an integer a.s.*,
(v) $N_0 = 0$ *and* N_f *is increasing in* f,
(vi) $\mathrm{E}|N_f| \leq \frac{1}{2\pi}\|f\|_1$,
(vii) $\mathrm{E}N_f = \frac{1}{2\pi}\int_0^\infty f(x)\,dx$, *and*
(viii) N_f *has exponential tails. For integers* $a, k > 0$ *we have*

$$\mathrm{P}(|N_f| \geq ak) \leq 2\left[\frac{\|f\|_1}{2\pi a}\right]^k.$$

Proof Claim (i) holds because $\alpha_f - \alpha_g$ solves

$$d\alpha_\lambda = \lambda(f-g)\,dt + \mathrm{Re}((e^{-i\alpha_\lambda} - 1)dZ^*), \qquad \alpha_\lambda(0) = 0,$$

with $dZ^* = e^{-i\alpha_g}dZ$. The standard coupling argument shows that the solution of the stochastic sine equation is monotone in the drift term, so we get (ii).

Now assume that $f \geq g \equiv 0$. Then by the above $\alpha_f(t) \geq \alpha_g(t) = 0$. Claim (iii) follows by repeating this argument for the process after the hitting time of $2k\pi$.

Assume $f \geq 0$, and let $F(t) = \int_0^t f(s)\,ds$. Then $\alpha_f - F$ is a continuous local martingale which is uniformly bounded below by $-\|f\|_1$. Thus it a.s. converges to a random limit. So α also converges, but it can only converge to a location where the noise term vanishes; we get (iv), and (v) also follows from (ii). Now (vi) follows from

$$\begin{aligned} 2\pi\mathrm{E}N_f - \|f\|_1 = \mathrm{E}\alpha_f(\infty) - F(\infty) &\leq \lim_{t\to\infty}(\mathrm{E}\alpha_f(t) - F(t)) \\ &= \alpha_f(0) - F(0) = 0, \end{aligned}$$

where the inequality is by Fatou's Lemma. By (iii) the function $t \mapsto \lfloor \alpha_f(t) \rfloor_{2\pi}$ is nondecreasing, hence the above inequality implies that $\lfloor \alpha_f(t) \rfloor_{2\pi}$ is uniformly integrable, and so is $\alpha_f(t)$. Thus $\alpha_f - F$ is a uniformly integrable martingale and so $\mathrm{E}\alpha_f(\infty) = F(\infty)$, as required for (vii).

For general $f \in L^1_*$, monotonicity (ii) gives

$$\alpha_{-f^-} \leq \alpha_f \leq \alpha_{f^+}, \tag{10}$$

where $x^+ = \max(x, 0)$ and $x^- = (-x)^+$. Now α_{-f^-} has the same distribution as $-\alpha_{f^-}$. By the previous argument α_{f^+} and α_{f^-} are uniformly integrable. Hence $\alpha_f - F$ is a uniformly

integrable martingale, and (iv), (vii) follow. Claim (v) also follows via (ii). We take positive and negative part of (10) to get $\alpha_f^+ \le \alpha_{f^+}$, and $\alpha_f^- \le -\alpha_{-f^-}$, where the latter has the same distribution as α_{f^-}. Taking limits and expectations gives

$$\mathrm{E}|\alpha_f(\infty)| = \mathrm{E}\alpha_f(\infty)^+ + \mathrm{E}\alpha_f(\infty)^- \le \mathrm{E}\alpha_{f^+}(\infty) + \mathrm{E}\alpha_{f^-}(\infty) = \|f\|_1$$

which gives (vi).

Returning to $f \ge 0$, Markov's inequality implies that $\mathrm{P}(N_f \ge a) \le \frac{1}{2\pi}\|f\|_1/a$. Stopping the process at time τ when and if α hits $2\pi ka$ we note that

$$\mathrm{P}\left(N_f \ge (k+1)a \,\middle|\, N_f \ge ka, \mathcal{F}_\tau\right) \le \frac{1}{2\pi}\|f_\tau\|_1/a \le \frac{1}{2\pi}\|f\|_1/a,$$

where f_τ is f shifted to the left by τ. It follows that for integer k we have

$$\mathrm{P}(N_f \ge ka) \le \left[\frac{1}{2\pi}\|f\|_1/a\right]^k.$$

For general $f \in L^1_*$, we consider the positive and negative parts separately and use monotonicity (v) to get (viii). □

Remark 10 The previous lemma shows that for a fixed $f \in L^1_*$ the random function $N(\lambda)$ is a.s. finite, integer valued, monotone increasing with stationary increments. Thus $N(\lambda)$ is the counting function of a translation invariant point process. Since $f \mapsto \alpha_f$ and $f \mapsto -\alpha_{-f}$ have the same distribution, the distribution of the point process is symmetric with respect to reflections.

Corollary 11 *For any $f \in L^1_*$ the point process defined by $N(\lambda)$ is a.s. simple.*

Proof The tail estimate of the lemma implies that the probability that there are two points or more in a fixed interval of length ε is at most $c\varepsilon^2$. Breaking the interval $[0, 1]$ into pieces of length ε, and using translation invariance, we see that the chance that there is a double point in $[0, 1]$ is at most $2c\varepsilon$. Letting $\varepsilon \to 0$ shows that a.s. there are no double points in $[0, 1]$. The claim now follows from translation invariance. □

Let $\mathcal{M}$ denote the space of probability distributions on $\mathbb{Z}$ with expectation. For $\mu_1, \mu_2 \in \mathcal{M}$ let $d(\mu_1, \mu_2)$ be the first Wasserstein distance, i.e. the infimum of $\mathrm{E}|X_1 - X_2|$ over all realizations where the joint distribution of (X_1, X_2) has marginals μ_1 and μ_2. The topology induced by d is stronger than weak convergence of probability measures. Let $\mathcal{L}(N_f)$ denote the distribution of N_f. The following proposition is a stronger version of Proposition 6 in the introduction.

Proposition 12 *The map $f \mapsto \mathcal{L}(N_f)$ is Lipschitz-1 continuous in f: for $f, g \in L^1_*$ we have $d(\mathcal{L}(N_f), \mathcal{L}(N_g)) \le \|f - g\|_1$.*

Proof Proposition 9 gives that $N_g - N_f$ has the same distribution as N_{g-f} which implies

$$d(\mathcal{L}(N_f), \mathcal{L}(N_g)) \le \mathrm{E}|N_g - N_f| = \mathrm{E}|N_{g-f}| \le \|g - f\|_1. \qquad \square$$

2.3 Large gap probabilities

Theorem 13 *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ satisfy $f(t) \le c/(1+t^2)$ for all t and $\int_0^\infty |df| < \infty$. Let $k \ge 0$. As $\lambda \to \infty$, for the point process given by the Brownian carousel with parameter f we have*

$$\mathrm{P}(\#\textit{ of points in } [0,\lambda] \le k) = \exp\bigl(-\lambda^2(\|f\|_2^2/8 + o(1))\bigr). \tag{11}$$

Lemma 14 *Let Y be an adapted stochastic process with $|Y_t| < m$, and let X satisfy the SDE $dX = Y\,dB$ where B_t is a Brownian motion. Then for each $a, t > 0$ we have*

$$\mathrm{P}(X(t) - X(0) \ge a) \le \exp\left(-a^2/(2tm^2)\right).$$

Proof We may assume $X(0) = 0$. Then $X_t = B_\tau$ where τ is the random time change $\tau = \int_0^t Y^2(s)ds$. Since $\tau < m^2 t$ the inequality now follows from

$$\mathrm{P}(B_r > a) \le \exp\bigl(-a^2/(2r)\bigr). \qquad \square$$

Proof of Theorem 13 The event in (11) is given in terms of the stochastic sine equation as $\lim_{t\to\infty} \alpha_\lambda(t) \le 2k\pi$. We will give upper and lower bounds on its probability.

Upper bound. By Proposition 9(iii) it is enough to give an upper bound on the probability that α stays less than $x = 2(k+1)\pi$. For $0 < s < t$ we have

$$\mathrm{P}(\alpha(t) < x \mid \mathcal{F}_s) = \mathrm{P}\left(-\int_s^t 2\sin(\alpha/2)dB > \lambda \int_s^t f\,dt - x + \alpha(s) \,\middle|\, \mathcal{F}_s\right).$$

We may drop the $\alpha(s)$ from the right hand side and use Lemma 14 with $Y = -2\sin(\alpha/2)$, $m = 2$, $a = \lambda(\int_s^t f\,dt - x/\lambda)$ to get the upper bound

$$\mathrm{P}(\alpha(t) < x \mid \mathcal{F}_s) \le \exp(-\lambda^2 r(s,t)), \quad r(s,t) = \frac{(\int_s^t f\,dt - x/\lambda)^2}{8(t-s)}.$$

Then, by just requiring $\alpha(t) < x$ for times $\varepsilon, 2\varepsilon, \ldots \in [0, K]$ the probability that α stays less than $x = 2(k+1)\pi$ is bounded above by

$$\mathrm{E}\prod_{k=0}^{K/\varepsilon} \mathrm{P}(\alpha((k+1)\varepsilon) < x \mid \mathcal{F}_{k\varepsilon}) \le \exp\left\{-\lambda^2 \sum_{k=0}^{K/\varepsilon} r(\varepsilon k, \varepsilon k + \varepsilon)\right\}.$$

A choice of ε so that $x/\lambda = o(\varepsilon)$ as $\lambda \to \infty$ yields the asymptotic Riemann sum

$$\sum_{k=0}^{K/\varepsilon} r(\varepsilon k, \varepsilon k + \varepsilon) = \frac{1}{8}\int_0^K f^2(t)dt + o(1).$$

Letting $K \to \infty$ provides the desired upper bound.

Lower bound. Consider the solution $\tilde{\alpha}(t)$ of (3) with the same driving Brownian motion, but with initial condition $\tilde{\alpha}(0) = \pi$. Then $\tilde{\alpha} \ge \alpha$. For $\varepsilon < \pi/4$, let A_s be the event that $\tilde{\alpha}(t) \in (0, \pi + \varepsilon]$ for $t \in [0, s]$. Then

$$\mathrm{P}(\alpha(\infty) < 2\pi) \ge \mathrm{P}(A_s) \sup_{y \in (0,\pi+\varepsilon)} \mathrm{P}(\tilde{\alpha}(\infty) < 2\pi \mid \tilde{\alpha}(s) = y).$$

The sup is bounded below via Markov's inequality by

$$1 - \frac{\pi + \varepsilon + \lambda \int_s^\infty f(t)dt}{2\pi} \geq 1/4,$$

where the last inequality holds if s is set to be a large constant multiple of λ. The event A_s is equivalent to $R = \log\tan(\tilde{\alpha}/4)$ staying in the interval $I = (-\infty, \log\tan((\pi+\varepsilon)/4)]$ where the evolution of R is given by Itô's formula as

$$dR = \frac{\lambda}{2} f \cosh R\, dt + \frac{1}{2}\tanh R\, dt + dB, \qquad R(0) = 0.$$

Let $I^* = [-\varepsilon, \log\tan((\pi+\varepsilon)/4)]$, and consider a process R^* so that (i) the noise terms of R and R^* are the same and (ii) the drift term of R^* at every time is greater than the spatial maximum over I^* of the drift term of R. Let A_s^* denote the event that for $t \in [0, s]$ we have $R_t^* \in I^*$. On this event $R^* \geq R$, and therefore A_s also holds. With an appropriate choice of c we may set

$$q(t) = (1/2 + c\varepsilon)\lambda f(t) + c\varepsilon, \qquad dR^* = q(t)dt + dB.$$

Let A_s^* also denote the corresponding set of paths. Girsanov's theorem gives

$$\mathrm{P}(R^* \in A_s^*) = \mathrm{E}\left[\mathbf{1}(B \in A_s^*)\exp\left(-\frac{1}{2}\int_0^s q(t)^2 dt + \int_0^s q(t)dB(t)\right)\right]. \tag{12}$$

Integration by parts transforms the second integral:

$$q(s)B(s) - \int_0^s B(t)dq(t) \geq -c\varepsilon\lambda\left(f(s) + \int_0^\infty |df|\right) - c\varepsilon \geq -c'\varepsilon(1+\lambda)$$

on the event $B \in A_s^*$. Here we also used that f is bounded. The probability of this event, i.e. that Brownian motion stays in an interval of width $c\varepsilon$, is at least $\exp(-c's/\varepsilon^2)$. In summary, (12) is bounded below by

$$\exp\big(-c\varepsilon(1+\lambda) - cs/\varepsilon^2 - (1/8 + c\varepsilon)\lambda^2\|f\|_2^2\big).$$

The choice $s = c\lambda$, $\varepsilon = \lambda^{-1/3}$ gives the desired lower bound. □

2.4 A phase transition at $\beta = 2$

The goal of this section is to prove Theorem 7 that at $\beta = 2$ there is a phase transition in the behavior of the stochastic sine equation.

As α converges to an integer multiple of 2π and it can never go below an integer multiple of 2π that it has passed (Proposition 9(iii)), eventually it must converge either from above or from below. Theorem 7 says that α converges from above with probability 1 if and only if $\beta \leq 2$. Otherwise, it converges from below with positive probability.

Proof Case $\beta \leq 2$. It suffices to prove that if $\alpha_\lambda(t_0) \in (2\pi k - y, 2\pi k)$ with $0 < y < \pi$ then $\alpha_\lambda(t)$ leaves this interval a.s. in finite time. As $\alpha_\lambda(t + t_0)$ also evolves according to the stochastic sine equation with $\lambda' = \lambda e^{-\beta t_0/4}$, we may set $t_0 = 0$ and we are also free to set $k = 1$. Let $\mathcal{B}$ denote the event that the process α_λ started at $\alpha_\lambda(0) = x$ in $(2\pi - y, 2\pi)$ will stay in this interval forever. It suffices to show that $\mathcal{B}$ has zero probability.

Consider $R = \log\tan(\alpha_\lambda/4)$ and set y so that $\log\tan((2\pi - y)/4) = 1$. While $\alpha_\lambda \in (0, 2\pi)$, Itô's formula gives the evolution of the process R:

$$R(0) = r_0 > 1, \qquad dR = q(R,t)dt + dB,\ q(r,t) = \frac{\lambda}{2}\cosh r\, e^{-\beta t/4} + \frac{1}{2}\tanh r. \tag{13}$$

Then $\mathcal{B}$ is the event that $R(t) \in (1, \infty)$ for all t. On $\mathcal{B}$ we have

$$q(R,t) > \frac{1}{2}\tanh R \geq \frac{1}{2} - e^{-2R} \geq 1/4,$$

which gives $R(t) \geq t/4 - B(t) + r_0$ from (13). Set

$$Q(t) = \int_0^t \frac{\tanh(R) - 1}{2} ds, \tag{14}$$

by the previous inequality, on the event $\mathcal{B}$ we have

$$Q(t) \geq -\int_0^t e^{-t/2+2B(s)-2r_0} ds > -\int_0^\infty e^{-t/2+2B(s)-2r_0} ds = -M,$$

where the integral M is a.s. finite. Let

$$L(t) = R(t) - t/2 - B(t) - Q(t). \tag{15}$$

Then on $\mathcal{B}$ we have

$$L'(t) = \frac{\lambda}{2}\cosh(L(t) + t/2 + B(t) + Q(t))e^{-\beta t/4} \tag{16}$$

$$\geq \frac{\lambda}{4}\exp\big[L(t) + B(t) + t(1/2 - \beta/4) - M\big]. \tag{17}$$

The equation follows from Itô's formula and the inequality uses $\cosh r \geq e^r/2$. Multiplying (17) by e^{-L} and integrating we get that on the event $\mathcal{B}$

$$e^{-L(0)} - e^{-L(t)} \geq C\int_0^t \exp\big[B(s) + s(1/2 - \beta/4)\big] ds$$

with a random $0 < C < \infty$. As the exponent is a Brownian motion with nonnegative drift, the limit of the integral on the right is a.s. infinite, thus the probability of $\mathcal{B}$ is 0.

Case $\beta < 2$. It suffices to prove that for a large t_0 if $\alpha_\lambda(t_0) \in (2\pi - \varepsilon, 2\pi)$ then $\alpha_\lambda(t)$ stays in the slightly larger interval $(2\pi - \delta, 2\pi)$ with positive probability. Choosing the values of ε and δ appropriately it suffices to show that if $R(0) > 2$ and λ is small enough then the event $\mathcal{B}$ that $R \in (1, \infty)$ for $t \geq 1$ has positive probability.

Recall the definition of Q and L from (14) and (15). On the event $\mathcal{B}$ we have

$$1/4 \leq \tanh(R) \leq 1/2, \quad \text{and} \quad -t/4 \leq Q(t) \leq 0.$$

Using this with (16) and the fact that for r nonnegative $\cosh r \leq e^r$ we get

$$L'(t) \leq \frac{\lambda}{2}\exp\big[L(t) + B(t) + t(1/2 - \beta/4)\big]. \tag{18}$$

From (18) we get

$$e^{-L(0)} - e^{-L(t)} \le \frac{\lambda}{2} \int_0^t \exp\big[B(s) + s(1/2 - \beta/4)\big]\, ds.$$

Let M^* denote the above integral for $t = \infty$. Then M^* is almost surely finite. Moreover, $L(t)$ and thus $R(t)$ remain finite if

$$M^* < 2e^{-L(0)}/\lambda. \tag{19}$$

From (16) we get $L'(t) > 0$ and $L(t) > L(0) = 2$ which gives

$$R(t) > 2 + B(t) + t/2 + Q(t) \ge 2 + B(t) + t/4.$$

So $R(t)$ stays above 1 if

$$B(t) \ge -t/4 - 1 \quad \text{for all } t. \tag{20}$$

This has positive probability, so the conditional distribution of and M^* given (20) is supported on finite numbers. This means that the intersection of the events (20) and (19) holds with positive probability for a sufficiently small choice of λ, and it implies $\mathcal{B}$. □

3 Breakdown of the proof of Theorem 1

The goal of this section is to divide the proof of the main theorem into independent pieces, which in turn will be proved in the later sections. The proof presented here also serves as an outline of the later sections.

Proof of Theorem 1 Fix $\beta > 0$, and consider the $n \times n$ random tridiagonal matrix

$$M(n) = \frac{1}{\sqrt{\beta}} \begin{pmatrix} \mathcal{N}_0 & \chi_{(n-1)\beta} & & \\ \chi_{(n-1)\beta} & \mathcal{N}_1 & \chi_{(n-2)\beta} & \\ & \chi_{(n-2)\beta} & \mathcal{N}_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}, \tag{21}$$

where the $\chi_{j\beta}$ and $\mathcal{N}_j$ entries are independent, $\mathcal{N}_j$ has normal distribution with mean 0 and variance 2, and $\chi_{j\beta}$ has chi distribution with $j\beta$ degrees of freedom. (For integer values of its parameter, χ_d is the length of a d-dimensional vector with independent standard normal entries.) We let Λ_n be the multi-set of eigenvalues of this matrix, which by Dumitriu and Edelman [4] has the desired distribution (1).

First, we may assume that $\mu_n \ge 0$; indeed, for a tridiagonal matrix, changing the sign of all diagonal elements changes the spectrum to its negative. In our case the diagonal elements have symmetric distributions, and by Remark 10 the limiting Sine_β process is also symmetric.

We set

$$n_0 = n_0(n) = n - \mu_n^2/4 - \frac{1}{2}.$$

The assumption $n^{1/6}(2\sqrt{n} - |\mu_n|) \to \infty$ implies

$$n_0^{-1}\mu_n^{2/3} \to 0, \qquad \frac{4n - \mu_n^2}{4n_0} \to 1.$$

So it suffices to show that

$$\text{if} \quad n_0^{-1}\mu_n^{2/3} \to 0 \quad \text{as } n \to \infty, \quad \text{then} \quad 2n_0^{1/2}(\Lambda_n - \mu_n) \Rightarrow \mathrm{Sine}_\beta, \tag{22}$$

an equivalent version of the claim which makes computations nicer. Recall that the counting function $N(\lambda)$ of a set of points in $\mathbb{R}$ is the number of points in $(0, \lambda]$ for $\lambda \geq 0$ or negative the number of points in $(\lambda, 0]$ for $\lambda < 0$.

Denote the counting function of the random multiset $2n_0^{1/2}(\Lambda_n - \mu_n)$ by $N_n(\lambda)$, and that of Sine_β by $N(\lambda)$. Claim (22) follows if for every $d \geq 1$ and $(\lambda_1, \lambda_2, \ldots, \lambda_d) \in \mathbb{R}^d$ we have

$$\big(N_n(\lambda_1), N_n(\lambda_2), \ldots, N_n(\lambda_d)\big) \overset{d}{\Longrightarrow} \big(N(\lambda_1), N(\lambda_2), \ldots, N(\lambda_d)\big).$$

The proof of this consists of several steps, these are verified in detail in the subsequent sections with the help of the Appendix.

Consider the one-parameter family of SDEs defining the Sine_β process:

$$d\tilde{\alpha}_\lambda = \lambda \frac{\beta}{4} e^{-\beta t/4} dt + \mathrm{Re}((e^{-i\tilde{\alpha}_\lambda} - 1) dZ), \tag{23}$$

where Z is complex Brownian motion on $[0, \infty)$ with standard real and imaginary parts. The time-change $t \to -\frac{2}{\beta}\log(1-t)$ transforms (23) to

$$2\sqrt{\beta(1-t)}\, d\alpha_\lambda = \lambda \beta^{1/2} dt + 2\sqrt{2}\,\mathrm{Re}((e^{-i\alpha_\lambda} - 1) dW), \tag{24}$$

where W_t is complex Brownian motion for $t \in [0, 1)$ with standard real and imaginary parts. Proposition 9 of Sect. 2 shows that the counting function $N(\lambda)$ of the process Sine_β can be represented as the right-continuous version of $(2\pi)^{-1} \lim_{t\to\infty} \tilde{\alpha}_\lambda(t)$, a limit which exists for every $\lambda \in \mathbb{R}$ a.s. This gives

Step 1 *For every $\lambda \in \mathbb{R}$, a.s. we have $2\pi N(\lambda) = \lim_{\varepsilon \to 0^+} \alpha_\lambda(1-\varepsilon)$.*

The eigenvalue equation for a tridiagonal matrix gives a three-term recursion for the eigenvector entries. This can be solved for any value of λ, but the boundary condition given by the last equation is only satisfied for eigenvalues.

The ratios of consecutive eigenvector entries $r_{\ell,\lambda}$ evolve via transformations of the form $r \mapsto a - b/r$, $b > 0$. These transformations are isometries of the Poincaré half plane model of the hyperbolic plane. The hyperbolic framework is introduced in Sect. 4.1 for the study of these recursions. In particular, $r_{\ell,\lambda}$ moves on the boundary of the hyperbolic plane which can be represented as a circle, eigenvalues can be counted by tracking the winding number of $r_{\ell,\lambda}$ as a function of λ. The rough phase function $\hat{\varphi}_{\ell,\lambda}$ (introduced in Sect. 4.2) transforms $r_{\ell,\lambda}$ to an angle through $2\arctan(r_{\ell,\lambda})$. Taking always the appropriate inverse of tan we get a continuous function of λ taking values in $\mathbb{R}$, the universal cover of the circle.

Our goal is to take limits of the evolution of $\hat{\varphi}_{\ell,\lambda}$. Since it has fast oscillations first it needs to be regularized. In order to remove the oscillations, we follow a shifted version of the hyperbolic angle of $r_{\ell,\lambda}$ around the fixed point of a simplified version of the transformation $r \mapsto a - b/r$. As we will see later, the important part of the evolution takes place in the interval $0 \leq \ell \leq \lfloor n_0 \rfloor$, which is exactly when this transformation is a hyperbolic rotation.

The precise regularization is done in Sect. 4.3; there we introduce the (regularized) phase function $\varphi_{\ell,\lambda}$ and target phase function $\varphi^{\odot}_{\ell,\lambda}$ with parameters $0 \leq \ell \leq \lfloor n_0 \rfloor$ and $\lambda \in \mathbb{R}$. These correspond to solving the eigenvalue equations starting from the two ends, 1 and n. As

$n \to \infty$, these two parts will require completely different treatment, so it is natural to break the evolution into two parts this way. Proposition 18 shows how we can count the eigenvalues using the zeroes of these phase functions mod 2π; this is a discrete analogue of the Sturm-Liouville oscillation theory.

Let $\#A$ denote the number of elements of A.

Step 2 *For $\ell = 1, 2, \ldots, \lfloor n_0 \rfloor$, the function $\varphi_{\ell,\lambda}$ is monotone increasing, and is independent of $\varphi^{\odot}_{\ell,\lambda}$. For any $\lambda < \lambda'$ and $1 \le \ell \le \lfloor n_0 \rfloor$ almost surely we have*

$$N_n(\lambda') - N_n(\lambda) = \#\big((\varphi_{\ell,\lambda} - \varphi^{\odot}_{\ell,\lambda}, \varphi_{\ell,\lambda'} - \varphi^{\odot}_{\ell,\lambda'}] \cap 2\pi\mathbb{Z}\big). \tag{25}$$

Since $\hat{\varphi}$ counts all eigenvalues below a certain level, its regularized version φ will encode all the fluctuations in the number of such eigenvalues. So the continuum limit of φ is expected to have large oscillations as its time-parameter converges to ∞. In order to deal with this problem, we introduce the *relative phase function* $\alpha_{\ell,\lambda} = \varphi_{\ell,\lambda} - \varphi_{\ell,0}$. This is related to the number of eigenvalues in an interval, so it is expected that its scaling limit will have nice behavior at $+\infty$. Note that $\alpha_{\ell,\lambda}$ has the same sign as λ by Step 2. Let

$$m_1 = \lfloor n_0(1-\varepsilon) \rfloor, \qquad m_2 = \lfloor n - \mu_n^2/4 - \kappa\,(\mu_n^{2/3} \vee 1) \rfloor,$$

where the constants $\varepsilon, \kappa > 0$ will be specified later in a way that the chain of inequalities $0 \le m_1 \le m_2$ holds.

Next we will describe the limiting behavior of $\varphi_{\ell,\lambda}$ and $\alpha_{\ell,\lambda}$ when ℓ is in the intervals $[0, m_1]$ and $[m_1, m_2]$, respectively. This is the content of the next three steps. Section 5.1 studies the behavior of the relative phase function on $[0, m_1]$. In Corollary 27 we will prove that $\alpha_{\ell,\lambda}$ converges to the SDE (24) in this stretch.

Step 3 *For every $0 < \varepsilon \le 1$*

$$\alpha_{m_1,\lambda} \overset{d}{\Longrightarrow} \alpha_\lambda(1-\varepsilon), \quad \text{as } n \to \infty \tag{26}$$

in the sense of finite dimensional distributions for λ.

Proposition 28 of Sect. 6.1 shows that $\alpha_{\ell,\lambda}$ does not change much in the second stretch, if it is already close to 0 mod 2π at the beginning of the stretch.

Step 4 *There exists constants c_0, c_1 depending only on $\bar{\lambda}$ and β such that if $\kappa = \kappa_n > c_0$, $\lambda \le |\bar{\lambda}|$ then*

$$\mathrm{E}\big[\big|(\alpha_{m_1,\lambda} - \alpha_{m_2,\lambda})\big| \wedge 1\big] \le c_1(\mathrm{E}\,\mathrm{dist}(\alpha_{m_1,\lambda}, 2\pi\mathbb{Z}) + \varepsilon^{1/2} + n_0^{-1/2}(\mu_n^{1/3} \vee 1) + \kappa^{-1}). \tag{27}$$

In Proposition 33 of Sect. 6.3 we show that $\varphi_{m_2,0}$ becomes uniform mod 2π.

Step 5 *If $\kappa \to \infty$ and $n_0^{-1}\kappa(\mu_n^{2/3} \vee 1) \to 0$ then*

$$\{\varphi_{m_2,0}\}_{2\pi} \overset{d}{\Longrightarrow} \mathrm{Uniform}[0, 2\pi],$$

where $\{x\}_{2\pi} = \min_{k\in\mathbb{Z}, k\le x}(x - 2\pi k)$.

Finally, in Lemma 34 of Sect. 6.4 we show that nothing interesting happens after m_2.

Step 6 *For every fixed $\kappa > 0$ and $\lambda \in \mathbb{R}$*

$$\left|\varphi^{\odot}_{m_2,\lambda} - \varphi^{\odot}_{m_2,0}\right| \xrightarrow{P} 0, \quad \text{as } n \to \infty.$$

In a metric space, if $\lim_{n\to\infty} x_{n,k} = x_k$ for every k and also $\lim_{k\to\infty} x_k = x$ then we can find a subsequence $n(k) \to \infty$ for which $\lim_{n\to\infty} x_{n,n(k)} = x$. This simple fact, together with the previous steps, allows us to choose sequences $\varepsilon = \varepsilon_n \to 0$, $\kappa = \kappa_n \to \infty$ in a way that the following limits hold simultaneously:

$$(\alpha_{m_1,\lambda_i}, i = 1, \dots, d) \xRightarrow{d} 2\pi(N(\lambda_i),\ i = 1, \dots, d), \tag{28}$$

$$\{\varphi_{m_2,0}\}_{2\pi} \xRightarrow{d} \text{Uniform}[0, 2\pi], \tag{29}$$

$$\left|\varphi^{\odot}_{m_2,\lambda_i} - \varphi^{\odot}_{m_2,0}\right| \xrightarrow{P} 0, \quad i = 1, \dots, d. \tag{30}$$

Since $\text{dist}(\cdot, 2\pi\mathbb{Z})$ is a bounded continuous function, (28) implies that the right hand side of (27) vanishes in the limit and so

$$\alpha_{m_1,\lambda_i} - \alpha_{m_2,\lambda_i} \xrightarrow{P} 0, \quad i = 1, \dots, d. \tag{31}$$

By (28) and (31) the completion of the proof only requires the following last step. Let $\langle x\rangle_{2\pi}$ denote the element of $2\pi\mathbb{Z}$ in $[x - \pi, x + \pi)$.

Step 7 *For $i = 1, \dots, d$ and $\lambda = \lambda_i$ we have* $\lim_{n\to\infty} \mathrm{P}(2\pi N_n(\lambda) = \langle\alpha_{m_2,\lambda}\rangle_{2\pi}) = 1$.

We conclude by the proof of Step 7. We suggest skipping it at the first reading, as it is the most technical part of this outline. We include it here because it uses too much of the notation and assumptions of the preceding discussion.

We will assume $\lambda > 0$, the other case follows similarly. Then $0 \le \langle\alpha_{m_2,\lambda}\rangle_{2\pi} \in 2\pi\mathbb{Z}$, for any $x \in \mathbb{R}$ we have

$$\langle\alpha_{m_2,\lambda}\rangle_{2\pi} = 2\pi\,\#\big(\big(x, x + \langle\alpha_{m_2,\lambda}\rangle_{2\pi}\big] \cap 2\pi\mathbb{Z}\big).$$

Using this with $x = \varphi_{m_2,\lambda} - \varphi^{\odot}_{m_2,\lambda} - \langle\alpha_{m_2,\lambda}\rangle_{2\pi}$ we get

$$\langle\alpha_{m_2,\lambda}\rangle_{2\pi} = 2\pi\,\#\big(\big(\varphi_{m_2,0} - \varphi^{\odot}_{m_2,\lambda} + \alpha_{m_2,\lambda} - \langle\alpha_{m_2,\lambda}\rangle_{2\pi}, \varphi_{m_2,\lambda} - \varphi^{\odot}_{m_2,\lambda}\big] \cap 2\pi\mathbb{Z}\big). \tag{32}$$

The symmetric difference between the intervals in (32) and (25) is an interval J with endpoints $\varphi_{m_2,0} - \varphi^{\odot}_{m_2,0}$ and $\varphi_{m_2,0} - \varphi^{\odot}_{m_2,\lambda} + \alpha_{m_2,\lambda} - \langle\alpha_{m_2,\lambda}\rangle_{2\pi}$. So it suffices to show

$$\lim_{n\to\infty} \mathrm{P}\left((J \cap 2\pi\mathbb{Z}) = \emptyset\right) = 1. \tag{33}$$

We will show that the length of J converges to 0 while one of its endpoints becomes uniformly distributed mod 2π. First,

$$|J| \le \left|\alpha_{m_2,\lambda} - \langle\alpha_{m_2,\lambda}\rangle_{2\pi}\right| + \left|\varphi^{\odot}_{m_2,\lambda} - \varphi^{\odot}_{m_2,0}\right| \xrightarrow{P} 0,$$

where the convergence of the first term follows from (28), (31) as $\alpha_{m_2,\lambda}$ converges to an element of $2\pi\mathbb{Z}$; the convergence of the second term is (30). Also, since $\varphi_{m_2,0}$ and $\varphi^{\odot}_{m_2,0}$

are independent, from (29) we have $\{\varphi_{m_2,0} - \varphi^{\odot}_{m_2,0}\}_{2\pi} \xRightarrow{d} \text{Uniform}[0, 2\pi]$. Equation (33), Step 7 and the theorem follows. □

Proof of Corollary 3 Note that weak convergence of point processes is metrizable. Let $a_i \to \infty$. For every i, we can find $n_i > i$ so that the point process

$$\Lambda_i^* = 2\sqrt{a_i}\, n_i^{1/6}(\Lambda_{n_i} - 2\sqrt{n_i} + a_i n_i^{-1/6})$$

is $1/i$-close to $2\sqrt{a_i}(\text{Airy}_\beta + a_i)$ by Theorem 2. By Theorem 1, Λ_i^* converges to Sine_β. □

4 The hyperbolic description of the phase evolution

4.1 The hyperbolic point of view

The eigenvector equation for a tridiagonal matrix gives a three-term recursion in which each step is of the form $u_{\ell+1} = bu_\ell - au_{\ell-1}$, in our case with $a > 0$. Let $\text{PSL}(2, \mathbb{R})$ denote the group of linear fractional transformations preserving the upper half plane $\mathbb{H}$ and its orientation. Then $r_\ell = u_{\ell+1}/u_\ell$ evolves by elements of $\text{PSL}(2, \mathbb{R})$ of the form $r \mapsto b - a/r$.

We will think of $\mathbb{H}$ as the Poincaré half-plane model for the hyperbolic plane; it is equivalent to the Poincaré disk model $\mathbb{U}$ via the bijection

$$\mathbf{U} : \bar{\mathbb{H}} \to \bar{\mathbb{U}}, \qquad z \mapsto \frac{i - z}{i + z},$$

which is also a bijection of the boundaries. Thus $\text{PSL}(2, \mathbb{R})$ acts naturally on $\bar{\mathbb{U}}$, the closed unit disk. As r moves on the boundary $\partial\mathbb{H} \equiv \mathbb{R} \cup \{\infty\}$, its image under $\mathbb{U}$ will move along $\partial\mathbb{U}$.

In order to follow the number of times this image circles $\mathbb{U}$, we would like to extend the action of $\text{PSL}(2, \mathbb{R})$ from $\partial\mathbb{U}$ to its universal cover, $\mathbb{R}' \equiv \mathbb{R}$, where we use prime to distinguish this from $\partial\mathbb{H}$. This action is uniquely determined up to shifts by 2π, but here we have a choice. For each choice, we get an element of a larger group $\text{UPSL}(2, \mathbb{R})$ defined via its action on $\mathbb{R}'$. $\text{UPSL}(2, \mathbb{R})$ still acts on $\bar{\mathbb{H}}$ and $\bar{\mathbb{U}}$ just like $\text{PSL}(2, \mathbb{R})$, and for $\mathbf{T} \in \text{UPSL}(2, \mathbb{R})$ the three actions are denoted by

$$\bar{\mathbb{H}} \to \bar{\mathbb{H}} : z \mapsto z.\mathbf{T}, \qquad \bar{\mathbb{U}} \to \bar{\mathbb{U}} : z \mapsto z_\circ\mathbf{T}, \qquad \mathbb{R}' \to \mathbb{R}' : z \mapsto z_*\mathbf{T}.$$

We note in passing that the topological group $\text{UPSL}(2, \mathbb{R})$ is the universal cover of the hyperbolic motion group $\text{PSL}(2, \mathbb{R})$, and $\text{PSL}(2, \mathbb{R})$ is a quotient of $\text{UPSL}(2, \mathbb{R})$ by the infinite cyclic normal subgroup generated by the 2π-shift on $\mathbb{R}'$. For every $\mathbf{T} \in \text{UPSL}(2, \mathbb{R})$ the function $x \mapsto x_*\mathbf{T}$ is strictly increasing, analytic and quasiperiodic, i.e. $(x + 2\pi)_*\mathbf{T} = x_*\mathbf{T} + 2\pi$.

Given an element $\mathbf{T} \in \text{UPSL}(2, \mathbb{R})$, $x, y \in \mathbb{R}'$, we define the angular shift

$$\text{ash}_{\mathbb{R}'}(\mathbf{T}, x, y) = (y_*\mathbf{T} - x_*\mathbf{T}) - (y - x)$$

i.e. the amount the signed distance of x, y changed over the transformation $\mathbf{T}$. This only depends on the image of $\mathbf{T}$ in $\text{PSL}(2, \mathbb{R})$ and the images $v = e^{ix}$, $w = e^{iy} \in \partial\mathbb{U}$ of x, y under the covering map. This allows us to define $\text{ash}(\mathbf{T}, v, w)$; more concretely,

$$\text{ash}(\mathbf{T}, v, w) = \text{ash}_{\mathbb{R}'}(\mathbf{T}, x, y) = \text{Arg}_{[0,2\pi)}(w_\circ\mathbf{T}/v_\circ\mathbf{T}) - \text{Arg}_{[0,2\pi)}(w/v),$$

where the last equality has self-evident notation and is straightforward to check. Note also that the above formula defines $\mathrm{ash}(\mathbf{T}, v, w)$ for $\mathbf{T} \in \mathrm{PSL}(2, \mathbb{R})$, $v, w \in \partial\mathbb{U}$ as well. For explicit computations, we will rely on the following fact, whose proof is given in Appendix A.1.

Fact 15 (Angular Shift Identity) *Let* $\mathbf{T} \in \mathrm{PSL}(2, \mathbb{R})$ *be a Möbius transformation and* $v, w \in \partial\mathbb{U}$; *let* $\sigma = 0_\circ \mathbf{T}^{-1}$. *Then*

$$\mathrm{ash}(\mathbf{T}, v, w) = 2\mathrm{Arg}\left(\frac{(w-\sigma)v}{w(v-\sigma)}\right) = 2\mathrm{Arg}\left(\frac{1-\sigma\bar{w}}{1-\sigma\bar{v}}\right). \tag{34}$$

Next, we specify generators for $\mathrm{UPSL}(2, \mathbb{R})$. Let $\mathbf{Q}(\alpha)$ denote the rotation by α in $\mathbb{U}$ about 0, more precisely, the shift by α on $\mathbb{R}'$:

$$\varphi_* \mathbf{Q}(\alpha) = \varphi + \alpha. \tag{35}$$

For $a, b \in \mathbb{R}$ let $\mathbf{A}(a, b)$ be the affine map $z \mapsto a(z+b)$ in $\mathbb{H}$. If $a > 0$ then this is in $\mathrm{PSL}(2, \mathbb{R})$, it fixes the ∞ in $\partial\mathbb{H}$ and -1 in $\partial\mathbb{U}$. We specify the action of $\mathbf{A}$ on $\mathbb{R}'$ by making it fix $\pi \in \mathbb{R}'$. Then we have

$$\varphi_* \mathbf{A}(a, b) = \varphi + \mathrm{ash}(\mathbf{A}(a, b), -1, e^{i\varphi}). \tag{36}$$

The following lemma estimates the angular shift. The proof is given in Appendix A.1.

Lemma 16 *Suppose that for a* $\mathbf{T} \in \mathrm{UPSL}(2, \mathbb{R})$ *we have* $(i+z).\mathbf{T} = i$ *with* $|z| \le 1/3$. *Then*

$$\begin{aligned}\mathrm{ash}(\mathbf{T}, v, w) &= \mathrm{Re}\left[(\bar{w}-\bar{v})\left(-z - \frac{i(2+\bar{v}+\bar{w})}{4} z^2\right)\right] + \varepsilon_3 \\ &= -\mathrm{Re}\,[(\bar{w}-\bar{v})z] + \varepsilon_2 \\ &= \varepsilon_1,\end{aligned} \tag{37}$$

where for $d = 1, 2, 3$ *and an absolute constant* c *we have*

$$|\varepsilon_d| \le c|w-v||z|^d \le 2c|z|^d. \tag{38}$$

If $v = -1$ *then the previous bounds hold even in the case* $|z| > 1/3$.

4.2 Phase evolution equations

The eigenvalue equation of a tridiagonal matrix can be solved recursively. The goal of this section is to analyze this recursion in terms of phase functions.

We conjugate the matrix $M = M(n)$ in (21) by a diagonal matrix D with

$$D_{ii} = D(n)_{ii} = \prod_{\ell=1}^{i} \frac{\chi_{(n-\ell)\beta}}{\sqrt{\beta}\, s_\ell}, \quad \text{where } s_j = \sqrt{n-j-1/2}.$$

We get the tridiagonal matrix $M^D = D^{-1}MD$ given by

$$\frac{1}{\sqrt{\beta}}\begin{pmatrix} \mathcal{N}_0 & \frac{\chi^2_{(n-1)\beta}}{s_1\sqrt{\beta}} & & \\ s_1\sqrt{\beta} & \mathcal{N}_1 & \frac{\chi^2_{(n-2)\beta}}{s_2\sqrt{\beta}} & \\ & s_2\sqrt{\beta} & \mathcal{N}_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix} = \begin{pmatrix} X_0 & s_0+Y_0 & & \\ s_1 & X_1 & s_1+Y_1 & \\ & s_2 & X_2 & \ddots \\ & & \ddots & \ddots \end{pmatrix}. \tag{39}$$

Then M^D and M have the same eigenvalues, but M^D has the property that the eigenvalue equations are independent. (A similar conjugation appears in [6].) The moments of the independent random variables

$$X_j = \frac{\mathcal{N}_j}{\sqrt{\beta}}, \qquad Y_j = \frac{\chi^2_{(n-j-1)\beta}}{\beta s_{j+1}} - s_j, \quad 0 \le j \le n-1$$

are explicitly computable via the moment generating functions for the Γ distribution.

Our proof is valid for any choice of independent real-valued random variables X_j, Y_j satisfying the following asymptotic moment conditions. X_j and Y_j may also depend on n, in which case the implicit error terms are assumed to be uniform in n.

moment	1st	2nd	4th
	$\mathcal{O}((n-j)^{-3/2})$	$2/\beta + \mathcal{O}((n-j)^{-1})$	$\mathcal{O}(1)$

(40)

Let $u_\ell = u_{\ell,\Lambda}$ $(1 \le \ell \le n)$ be a non-trivial solution of the first $n-1$ components of the eigenvalue equation with a given spectral parameter Λ, i.e.

$$\begin{aligned} &s_\ell u_\ell + X_\ell u_{\ell+1} + (Y_\ell + s_\ell)u_{\ell+2} = \Lambda u_{\ell+1}, \quad 0 \le \ell \le n-2, \\ &u_0 = 0, \qquad u_1 = 1. \end{aligned} \tag{41}$$

Then with $r_\ell = r_{\ell,\Lambda} = u_{\ell+1}/u_\ell$ we have

$$r_{\ell+1} = \left(-\frac{1}{r_\ell} + \frac{\Lambda}{s_\ell} - \frac{X_\ell}{s_\ell}\right)\left(1 + \frac{Y_\ell}{s_\ell}\right)^{-1}, \quad 0 \le \ell \le n-2. \tag{42}$$

This also holds for $r_\ell = 0$ or $r_\ell = \infty$; in fact, the initial value of the recursion is $r_0 = \infty$. If we set $Y_{n-1} = 0$ and define r_n via the $\ell = n-1$ case of (42), then Λ is an eigenvalue if and only if $r_n = 0$.

We will use the point of view and notation introduced in Sect. 4.1. Namely, r takes values in $\partial\mathbb{H} = \mathbb{R} \cup \{\infty\}$, the boundary of the hyperbolic plane. Moreover, the evolution of r can be lifted to the universal cover of $\partial\mathbb{H}$. The extra information there allows us to count eigenvalues, as the following proposition shows. The proposition also summarizes the evolution of r and its lifting $\hat{\varphi} \in \mathbb{R}'$. We note that this is just a discrete analogue of the Sturm-Liouville oscillation theory suitable for our purposes; such analogues are available in the literature. Although we state this proposition in our setting, a trivial modification holds for the eigenvalues of general tridiagonal matrices with positive off-diagonal terms.

Proposition 17 (Wild phase function) *There exist functions* $\hat{\varphi}, \hat{\varphi}^{\odot} : \{0, 1, \ldots, n\} \times \mathbb{R} \to \mathbb{R}$ *satisfying the following*:

(i) $r_{\ell,\Lambda}.\mathbf{U} = e^{i\hat{\varphi}_{\ell,\Lambda}}$.
(ii) $\hat{\varphi}_{0,\Lambda} = \pi$, $\hat{\varphi}^{\odot}_{n,\Lambda} = 0$.
(iii) *For each* $0 < \ell \le n$, $\hat{\varphi}_{\ell,\Lambda}$ *is analytic and strictly increasing in* Λ. *For* $0 \le \ell < n$, $\hat{\varphi}^{\odot}_{\ell,\Lambda}$ *is analytic and strictly decreasing in* Λ.
(iv) *For any* $0 \le \ell \le n$, Λ *is an eigenvalue of* M *if and only if* $\hat{\varphi}_{\ell,\Lambda} - \hat{\varphi}^{\odot}_{\ell,\Lambda} \in 2\pi\mathbb{Z}$.

Proof We consider the following elements of the universal cover UPSL(2, $\mathbb{R}$) of the hyperbolic motion group PSL(2, $\mathbb{R}$):

$$\mathbf{Q}(\pi), \qquad \mathbf{W}_j = \mathbf{A}((1 + Y_j/s_j)^{-1}, -X_j/s_j), \quad 0 \le j \le n-1 \tag{43}$$

where Q corresponds to a rotation in the model $\mathbb{U}$, and A corresponds to an affine map in the model $\mathbb{H}$, as defined in (35)–(36). With this notation, the evolution (42) of r becomes

$$\begin{aligned} \mathbf{R}_{\ell,\Lambda} &= \mathbf{Q}(\pi)\,\mathbf{A}(1, \Lambda/s_\ell)\,\mathbf{W}_\ell, \\ r_{\ell+1} &= r_\ell.\mathbf{R}_{\ell,\Lambda}, \end{aligned} \tag{44}$$

for $0 \le \ell \le n-1$, and Λ is an eigenvalue if and only if $\infty.\mathbf{R}_{0,\Lambda} \cdots \mathbf{R}_{n-1,\Lambda} = 0$. Multiplying this by $(\mathbf{R}_{\ell,\Lambda} \cdots \mathbf{R}_{n-1,\Lambda})^{-1}$ for some $0 \le \ell \le n$ and then moving to the universal cover $\mathbb{R}'$ of $\partial\mathbb{H}$ gives the equivalent characterization $\hat{\varphi}_{\ell,\Lambda} = \hat{\varphi}^{\odot}_{\ell,\Lambda}$ mod 2π, where

$$\hat{\varphi}_{\ell,\Lambda} = \pi_*\mathbf{R}_{0,\Lambda} \cdots \mathbf{R}_{\ell-1,\Lambda}, \qquad \hat{\varphi}^{\odot}_{\ell,\Lambda} = 0_*\mathbf{R}^{-1}_{n-1,\Lambda} \cdots \mathbf{R}^{-1}_{\ell,\Lambda}, \tag{45}$$

which is exactly (iv). Claims (i)–(ii) follow from the definition.

As $\varphi_{0,\Lambda} = \pi$, one readily checks that $\varphi_{1,\Lambda}$ is strictly increasing. Since $(\varphi, \Lambda) \mapsto \varphi_*\mathbf{R}_{\ell,\Lambda}$ are nondecreasing analytic functions in both parameters and so are their compositions, the statement of claim (iii) for $\hat{\varphi}_{\ell,\Lambda}$ now follows. The same proof works for $\hat{\varphi}^{\odot}$. □

Motivated by part (iv) of the proposition, we call $\hat{\varphi}^{\odot}$ the *target phase function*.

4.3 Slowly varying phase evolution for a scaling window

For scaling, we set

$$s(\tau) = s^{(n)}(\tau) = \sqrt{1 - \tau - 1/2n}$$

so that we have $s_\ell = s(\ell/n)\sqrt{n}$. Making s depend on n via the $1/2n$ term helps make the upcoming formulas exact rather than only asymptotic.

The phase function $\hat{\varphi}_\ell$ introduced in the previous section exhibits fast oscillations in ℓ. In this section we will extract a slowly moving component of the phase evolution whose limiting behavior can be identified. The oscillations of $\hat{\varphi}_\ell$ are caused by the macroscopic term $\mathbf{Q}(\pi)\,\mathbf{A}(1, \Lambda/s_\ell)$ of the evolution operator $\mathbf{R}_{\ell,\Lambda}$. The recursion (44) has different behavior depending on whether this macroscopic part is a rotation or not. As we will see later, the continuum limit process comes from the stretch $0 \le \ell < n_0$ where it is a rotation (this is because the corresponding eigenvectors will be localized there). The eigenvalues Λ of interest will be near the scaling window μ_n, so we define the main part of the evolution operator as the macroscopic part of $\mathbf{R}_{\ell,\mu_n}$, that is

$$\mathbf{J}_\ell = \mathbf{Q}(\pi)\,\mathbf{A}(1, \mu_n/s_\ell) = \mathbf{Q}(\pi)\mathbf{A}\left(1, \frac{\mu_n}{\sqrt{n}s(\ell/n)}\right). \tag{46}$$

This is a rotation if it has a fixed point ρ_ℓ in the open upper half plane $\mathbb{H}$, the fixed point equation $\rho_\ell.\mathbf{J}_\ell = \rho_\ell$ turns into

$$\rho_\ell^2 - 2\,\frac{\mu_n/\sqrt{4n}}{s(\ell/n)}\rho_\ell + 1 = 0. \tag{47}$$

Note that $\mu_n/\sqrt{4n}$ is the relative location of the scaling window in the Wigner semicircle supported on $[-1,1]$. Since $s(\tau)$ is decreasing, we have that $\rho_\ell \in \mathbb{H}$ for $\tau < n_0/n$, where $s(n_0/n) = \mu_n/\sqrt{4n}$. This explains the choice of the parameter n_0.

Thus $\rho_\ell = \rho(n_0/n, \ell/n)$, where $\rho(\tau_1,\tau_2)$ is the solution in the closed upper half plane of

$$\rho^2 - 2\,\frac{s(\tau_1)}{s(\tau_2)}\rho + 1 = 0, \quad \text{i.e.} \quad \rho(\tau_1,\tau_2) = \frac{s(\tau_1)}{s(\tau_2)} + i\sqrt{1 - \frac{s(\tau_1)^2}{s(\tau_2)^2}}. \tag{48}$$

More specifically,

$$\rho_\ell = \sqrt{\frac{\mu_n^2/4}{\mu_n^2/4 + n_0 - \ell}} + i\sqrt{\frac{n_0 - \ell}{\mu_n^2/4 + n_0 - \ell}}. \tag{49}$$

Because of our choice of scaling window and the density in the Wigner semicircle law it is natural to choose the scaling (22) by setting

$$\Lambda = \mu_n + \frac{\lambda}{2\sqrt{n_0}}. \tag{50}$$

We recycle the notation $u_{\ell,\lambda}, r_{\ell,\lambda}, \hat{\varphi}_{\ell,\lambda}, \hat{\varphi}^{\odot}_{\ell,\lambda}$ for the quantities $u_{\ell,\Lambda}, r_{\ell,\Lambda}, \hat{\varphi}_{\ell,\Lambda}, \hat{\varphi}^{\odot}_{\ell,\Lambda}$. We separate $\mathbf{J}_\ell$ from the evolution operator $\mathbf{R}$ to get:

$$\mathbf{R}_{\ell,\lambda} = \mathbf{J}_\ell \mathbf{L}_{\ell,\lambda}\mathbf{W}_\ell, \quad \mathbf{L}_{\ell,\lambda} = \mathbf{A}\left(1, \frac{\lambda}{2s(\ell/n)\sqrt{n_0 n}}\right). \tag{51}$$

Note that $\mathbf{L}_{\ell,\lambda}$ and $\mathbf{W}_\ell$ become infinitesimal in the $n \to \infty$ limit while $\mathbf{J}_\ell$ does not. $\mathbf{J}_\ell$ is a hyperbolic rotation, differentiating $z \mapsto z.\mathbf{J}_\ell$ at $z = \rho_\ell$ shows the angle to be $-2\mathrm{Arg}(\rho_\ell) \in [-\pi, 0]$. Let

$$\mathbf{T}_\ell = \mathbf{A}(\mathrm{Im}(\rho_\ell)^{-1}, -\mathrm{Re}(\rho_\ell))$$

correspond to the affine map sending $\rho_\ell \in \mathbb{H}$ to $i \in \mathbb{H}$, then we may write

$$\mathbf{J}_\ell = \mathbf{Q}(-2\mathrm{Arg}(\rho_\ell))^{\mathbf{T}_\ell^{-1}},$$

where $A^B = B^{-1}AB$. Rather than following $\hat{\varphi}$ itself, it will be more convenient to follow a version which is shifted so that the fixed point ρ_ℓ of the rough evolution is shifted to i. Moreover, in order to follow a slowly changing angle, we remove the cumulative effect of the macroscopic rotations $\mathbf{J}_\ell$. Essentially, we study the "difference" between the phase evolution of the random recursion and the version with the noise and λ terms removed. The quantity to follow is

$$\varphi_{\ell,\lambda} = \hat{\varphi}_{\ell,\lambda *}\mathbf{T}_\ell \mathbf{Q}_{\ell-1}, \tag{52}$$

where

$$\mathbf{Q}_\ell = \mathbf{Q}(2\mathrm{Arg}(\rho_0))\ldots\mathbf{Q}(2\mathrm{Arg}(\rho_\ell)), \quad -1 \le \ell \le n_0.$$

Acting on $\mathbb{U}$, $\mathbf{Q}_\ell$ is simply a rotation about 0, more precisely a multiplication by

$$\eta_\ell = \rho_0^2 \rho_1^2 \dots \rho_\ell^2. \tag{53}$$

From (45) and (52) we get that φ evolves by the one-step operator

$$(\mathbf{T}_\ell \mathbf{Q}_{\ell-1})^{-1} \mathbf{R}_{\ell,\lambda} (\mathbf{T}_{\ell+1} \mathbf{Q}_\ell) = (\mathbf{T}_\ell^{-1} \mathbf{L}_\ell \mathbf{W}_\ell \mathbf{T}_{\ell+1})^{\mathbf{Q}_\ell} := (\mathbf{S}_{\ell,\lambda})^{\mathbf{Q}_\ell}.$$

We keep this "conjugated" notation because $\mathbf{S}_{\ell,\lambda}$ corresponds to an affine transformation.

For $\ell \le n_0$ we define the corresponding target phase function

$$\varphi_{\ell,\lambda}^{\odot} = \hat{\varphi}_{\ell,\lambda *}^{\odot} \mathbf{T}_\ell \mathbf{Q}_{\ell-1}. \tag{54}$$

The following summarizes our findings and translates the results of Proposition 17 to this setting. Here and in the sequel we use the difference notation $\Delta x_\ell = x_{\ell+1} - x_\ell$.

Proposition 18 (Slowly varying phase function) *The functions* $\varphi, \varphi^{\odot} : \{0, 1, \dots, \lfloor n_0 \rfloor\} \times \mathbb{R} \to \mathbb{R}$ *satisfy the following for every* $0 < \ell \le n_0$:

(i) $\varphi_{0,\lambda} = \pi$.
(ii) $\varphi_{\ell,\lambda}$ *and* $-\varphi_{\ell,\lambda}^{\odot}$ *are analytic and strictly increasing in* λ, *and are also independent.*
(iii) *With* $\mathbf{S}_{\ell,\lambda} = \mathbf{T}_\ell^{-1} \mathbf{L}_\ell \mathbf{W}_\ell \mathbf{T}_{\ell+1}$, *we have* $\Delta\varphi_{\ell,\lambda} = \mathrm{ash}(\mathbf{S}_{\ell,\lambda}, -1, e^{i\varphi_{\ell,\lambda}} \bar{\eta}_\ell)$.
(iv) $\hat{\varphi}_{\ell,\lambda} = \varphi_{\ell,\lambda *} \mathbf{Q}_{\ell-1}^{-1} \mathbf{T}_\ell^{-1}$.
(v) *For any* $\lambda < \lambda'$ *we have a.s.* $N_{n,\lambda'} - N_{n,\lambda} = \#((\varphi_{\ell,\lambda} - \varphi_{\ell,\lambda}^{\odot}, \varphi_{\ell,\lambda'} - \varphi_{\ell,\lambda'}^{\odot}] \cap 2\pi\mathbb{Z})$.

The form

$$\mathbf{S}_{\ell,\lambda} = (\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell} \mathbf{S}_{\ell,0}$$

breaks S into a deterministic λ-dependent part and a random part that does not depend on λ. Let $\varphi_{\ell,\lambda}^* = \varphi_{\ell,\lambda *} (\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell \mathbf{Q}_\ell}$ be the intermediate phase between these two steps. Note that $\varphi_{\ell,0}^* = \varphi_{\ell,0}$, and

$$(\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell} = \mathbf{A}\left(1, \frac{\lambda}{2\sqrt{n_0 n}\sqrt{s(n_0/n)^2 - s(\ell/n)^2}}\right) = \mathbf{A}\left(1, \frac{\lambda}{2\sqrt{n_0(n_0-\ell)}}\right). \tag{55}$$

The relative phase functions

$$\alpha_{\ell,\lambda} = \varphi_{\ell,\lambda} - \varphi_{\ell,0}, \qquad \alpha_{\ell,\lambda}^* = \varphi_{\ell,\lambda}^* - \varphi_{\ell,0}^*$$

are the main tools for counting eigenvalues in intervals.

Proposition 19 (Relative phase function) *The function* $\alpha : \{0, 1, \dots, \lfloor n_0 \rfloor\} \times \mathbb{R} \to \mathbb{R}$ *satisfies*

(i) $\alpha_{0,\lambda} = 0$, $\alpha_{\ell,0} = 0$ *and for each* $\ell > 0$, $\alpha_{\ell,\lambda}$ *is analytic and strictly increasing in* λ.
(ii) $\Delta\alpha_{\ell,\lambda} = \mathrm{ash}((\mathbf{L}_{\lambda,n-\ell})^{\mathbf{T}_\ell}, -1, e^{i\varphi_{\ell,\lambda}} \bar{\eta}_\ell) + \mathrm{ash}(\mathbf{S}_{\ell,0}, e^{i\varphi_{\ell,\lambda}^*} \bar{\eta}_\ell, e^{i\varphi_{\ell,0}} \bar{\eta}_\ell)$

$$= \mathrm{ash}((\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell}, -1, e^{i\varphi_{\ell,\lambda}} \bar{\eta}_\ell) + \mathrm{ash}(\mathbf{S}_{\ell,0}, e^{i\varphi_{\ell,\lambda}^*} \bar{\eta}_\ell, e^{i\varphi_{\ell,\lambda}} \bar{\eta}_\ell)$$
$$+ \mathrm{ash}(\mathbf{S}_{\ell,0}, e^{i\varphi_{\ell,\lambda}} \bar{\eta}_\ell, e^{i\varphi_{\ell,0}} \bar{\eta}_\ell).$$

(iii) *For each* ℓ *and* $\lambda \ge 0$ *we have* $\lfloor \alpha_{\ell,\lambda} \rfloor_{2\pi} \le \lfloor \alpha_{\ell+1,\lambda}^* \rfloor_{2\pi} = \lfloor \alpha_{\ell+1,\lambda} \rfloor_{2\pi}$.

Proof (i)–(ii) are direct consequences of Proposition 18. To check (iii), we note

$$\begin{aligned}\alpha_{\ell,\lambda} &= \varphi_{\ell,\lambda} - \varphi_{\ell,0},\\ \alpha^*_{\ell,\lambda} &= \varphi_{\ell,\lambda *}(\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell \mathbf{Q}_\ell} - \varphi_{\ell,0},\\ \alpha_{\ell+1,\lambda} &= \varphi_{\ell,\lambda *}(\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell \mathbf{Q}_\ell}(\mathbf{S}_{\ell,0})^{\mathbf{Q}_\ell} - \varphi_{\ell,0*}(\mathbf{S}_{\ell,0})^{\mathbf{Q}_\ell}.\end{aligned}$$

Since the map $\mathbf{L}_{\ell,\lambda}$ and its conjugates are monotone in λ, we get $\alpha_{\ell,\lambda} \le \alpha^*_{\ell,\lambda}$. Since $(\mathbf{S}_{\ell,0})^{\mathbf{Q}_\ell}$ is the lifting of a Möbius transformation, it is monotone and 2π-quasiperiodic, whence $\lfloor \alpha^*_{\ell,\lambda} \rfloor_{2\pi} = \lfloor \alpha_{\ell+1,\lambda} \rfloor_{2\pi}$. □

Remark 20 (Translation to the original matrix) Let w_ℓ denote the solution of the discrete eigenvalue equation for the *original matrix* (21). It is given in terms of the diagonal matrix D defined in the beginning of the section and the solution u of recursion (41) as $w_\ell = (Du)_\ell$. The ratios of the consecutive entries of this vector are

$$p_\ell := \frac{w_{\ell+1}}{w_\ell} = \frac{(Du)_{\ell+1}}{(Du)_\ell} = \frac{u_{\ell+1}}{u_\ell}\,\frac{D_{\ell+1,\ell+1}}{D_{\ell,\ell}} = r_\ell\,\frac{\chi_{(n-\ell-1)\beta}}{\sqrt{\beta}\,s_{\ell-1}}.$$

If $\ell \le n_0$ then we may further rewrite this using z_ℓ as

$$p_\ell = \frac{\chi_{(n-\ell-1)\beta}}{\sqrt{\beta}\,s_{\ell+1}}\left((z_\ell.\mathbf{U}^{-1})\bar{\eta}_{\ell-1}\,\mathrm{Im}(\rho_\ell) + \mathrm{Re}(\rho_\ell)\right).$$

4.4 The discrete carousel

Corollary 27 in Sect. 5.2 shows that the appropriate limit of the relative phase function $\alpha_{\ell,\lambda}$ is the stochastic sine equation. In this section we bring the discrete evolution equations in the form that it becomes clear that their limit should be the Brownian carousel.

By (44) the evolution of r_ℓ is governed by a certain discrete process $\hat{\mathbf{G}}_{\ell,\lambda}$ in the hyperbolic automorphism group UPSL(2, $\mathbb{R}$):

$$r_\ell = r_0.\hat{\mathbf{G}}_{\ell,\lambda} = r_0.\mathbf{R}_{0,\lambda}\cdots\mathbf{R}_{\ell-1,\lambda}.$$

This process has rough jumps, but it is a smooth function of the parameter λ. It is therefore natural to expect that the evolution of the automorphism $\hat{\mathbf{G}}_{\ell,\lambda}\hat{\mathbf{G}}^{-1}_{\ell,0}$ will have a continuous scaling limit. In the following, we will rewrite this expression in a form indicating the desired scaling limit.

By (52) the evolution operator $\mathbf{G}_{\ell,\lambda}$ of φ satisfies $\mathbf{G}_{\ell,\lambda} = \hat{\mathbf{G}}_{\ell,\lambda}\mathbf{T}_\ell\mathbf{Q}_{\ell-1}$, and therefore $\mathbf{G}_{\ell,\lambda}\mathbf{G}^{-1}_{\ell,0} = \hat{\mathbf{G}}_{\ell,\lambda}\hat{\mathbf{G}}^{-1}_{\ell,0}$. The evolution of $\varphi_{\ell,\lambda}$ is given by

$$\varphi_{\ell,\lambda} = \varphi_{0,\lambda *}\mathbf{G}_{\ell,\lambda} = \pi_*\mathbf{G}_{\ell,\lambda}, \tag{56}$$

where with $A^B = B^{-1}AB$ we have

$$\begin{aligned}\mathbf{G}_{\ell,\lambda} &= \mathbf{Y}_{0,\lambda}\mathbf{X}_0\,\mathbf{Y}_{1,\lambda}\mathbf{X}_1\,\cdots\,\mathbf{Y}_{\ell-1,\lambda}\mathbf{X}_{\ell-1}\\ &= \mathbf{Y}_{0,\lambda}\,\mathbf{Y}_{1,\lambda}^{\mathbf{G}_1^{-1}}\,\mathbf{Y}_{2,\lambda}^{\mathbf{G}_2^{-1}}\,\cdots\,\mathbf{Y}_{\ell-1,\lambda}^{\mathbf{G}_{\ell-1}^{-1}}\,\mathbf{G}_\ell,\\ \mathbf{G}_\ell = \mathbf{G}_{\ell,0} &= \mathbf{X}_0\mathbf{X}_1\cdots\mathbf{X}_{\ell-1},\end{aligned} \tag{57}$$

and we used the temporary notation $\mathbf{Y}_{\ell,\lambda} = ((\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell})^{\mathbf{Q}_\ell}$, $\mathbf{X}_\ell = (\mathbf{S}_{\ell,0})^{\mathbf{Q}_\ell}$. By definition,

$$\alpha_{\ell,\lambda} = \varphi_{\ell,\lambda *}\mathbf{Q}(-\varphi_{\ell,0}) = \pi_* \mathbf{G}_{\ell,\lambda}\, \mathbf{Q}(-\varphi_{\ell,0}). \tag{58}$$

We introduce the notation

$$\gamma_{\ell,\lambda} := \pi_* \mathbf{G}_{\ell,\lambda}\mathbf{G}_\ell^{-1} = \pi_* \mathbf{Y}_{0,\lambda}\, \mathbf{Y}_{1,\lambda}^{\mathbf{G}_1^{-1}}\, \mathbf{Y}_{2,\lambda}^{\mathbf{G}_2^{-1}} \cdots \mathbf{Y}_{\ell-1,\lambda}^{\mathbf{G}_{\ell-1}^{-1}}, \tag{59}$$

$$B_\ell := 0_\circ \mathbf{G}_\ell^{-1} \in \mathbb{U}. \tag{60}$$

With $\mathcal{T}$ denoting the Möbius transformation defined in (8), we claim that

$$\mathcal{T}(B_\ell, z) = z_\circ \mathbf{G}_\ell\, \mathbf{Q}(-\varphi_{\ell,0})$$

with the choice of $z_0 = -1$. This follows from the fact that $\mathcal{T}(B_\ell, B_\ell) = 0$ by definition and $\mathcal{T}(B_\ell, -1) = 1$ by (56), (57). Hence (58) becomes

$$e^{i\alpha_{\ell,\lambda}} = \mathcal{T}(B_\ell, e^{i\gamma_{\ell,\lambda}})$$

which is the same form as (9) relating the stochastic sine equation to the Brownian carousel ODE.

Remark 21 (Heuristics) Note that $\mathbf{X}_\ell$ is approximately an infinitesimal noise element in $\mathrm{PSL}(2,\mathbb{R})$. $\mathbf{X}_\ell$ acting on $\mathbb{U}$ moves 0 infinitesimally in a random direction. This direction is not necessarily isotropic, but the conjugation by the macroscopic rotation $\mathbf{Q}_\ell$ makes the composition of consecutive $\mathbf{X}_\ell$'s move 0 to an approximately isotropic random direction. Thus B_ℓ in (60) approximates hyperbolic Brownian motion started at 0 run at a time-dependent speed. Similarly, the $\mathbf{Y}_{\ell,\lambda}$ are infinitesimal parallel translations, but because of the conjugation by the macroscopic rotations $\mathbf{Q}_\ell$, their composition approximates rotation about 0. Thus $\gamma_{\ell,\lambda}$ in (59) approximately evolves by rotations about B_ℓ. This is exactly how the Brownian carousel evolves, giving a conceptual explanation of our results. This suggests an alternative way to prove our results via the Brownian carousel ODE (5).

5 The stochastic sine equation as a limit

This section describes the stochastic differential equation limit of the phase function on the first stretch $[0, n_0(1-\varepsilon)]$. In the limit, this stretch completely determines the eigenvalue behavior; this will be proved in Sect. 6.

5.1 Single-step asymptotics

Let $\mathcal{F}_\ell$ denote the σ-field generated by the random variables $X_0, X_1, \ldots, X_{\ell-1}$, and $Y_0, Y_1, \ldots, Y_{\ell-1}$. Let $\mathrm{E}_\ell[\,\cdot\,]$ denote conditional expectation with respect to $\mathcal{F}_\ell$. By definition, the random variables $\hat{\varphi}_{\ell,\lambda}, \varphi_\ell, \alpha_\ell$ are measurable with respect to $\mathcal{F}_\ell$. Moreover, for fixed λ, both $\hat{\varphi}_{\ell,\lambda}$ and $\varphi_{\ell,\lambda}$ are Markov chains adapted to $\mathcal{F}_\ell$.

Throughout this and the subsequent sections we assume that $|\lambda|$ is bounded by a constant $\bar{\lambda}$. By default the notation $\mathcal{O}(x)$ will refer to a deterministic quantity whose absolute value is bounded by $c|x|$, where c depends only on β and $\bar{\lambda}$. As ℓ varies k will denote $n_0 - \ell$.

This section presents the asymptotics for the moments of step $\Delta\varphi_{\ell,\lambda} := \varphi_{\ell+1,\lambda} - \varphi_{\ell,\lambda}$. Recall from Sect. 4.3 that ℓ moves on the interval $[0, n_0]$. The continuum limit of $\varphi_{\ell,\lambda}$ will live on the time interval $[0, 1]$ so we introduce

$$t = \frac{\ell}{n_0} \in [0, 1].$$

We also introduce the rescaling of $s(t)^2$ on this stretch:

$$\hat{s}(t)^2 = \frac{s(t\, n_0/n)^2 - s(n_0/n)^2}{n_0/n} \tag{61}$$

with $\hat{s} \geq 0$. This actually simplifies to

$$\hat{s}(t) = \sqrt{1-t} = \sqrt{k/n_0} \tag{62}$$

in our case. As we will see later, the scaling limit of the evolution of the relative phase function will depend on s and the scaling parameters through $\hat{s}$. The fact that this function only depends on t explains why the point process limits do not depend on the choice of the scaling window in Theorem 1. In Sect. 5.3 we provide a more detailed discussion and further implications. We will keep the notation $\hat{s}$ (instead of writing $\sqrt{1-t}$) to facilitate the treatment of a more general model discussed there.

Proposition 18(iii) expresses the difference $\Delta\varphi_{\ell,\lambda} := \varphi_{\ell+1,\lambda} - \varphi_{\ell,\lambda}$ via the angular shift of $\mathbf{S}_{\ell,\lambda}$ and $(\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell}$. Lemma 16, in turn, writes the angular shift in terms of the pre-image of $i = \sqrt{-1}$. In the present case

$$Z_{\ell,\lambda} = i.\mathbf{S}_{\ell,\lambda}^{-1} - i = i.\mathbf{T}_{\ell+1}^{-1}(\mathbf{L}_{\ell,\lambda}\mathbf{W}_\ell)^{-1}\mathbf{T}_\ell - i = v_{\ell,\lambda} + V_\ell, \tag{63}$$

where

$$v_{\ell,\lambda} = -\frac{\lambda}{2n_0\hat{s}(t)} + \frac{\rho_{\ell+1} - \rho_\ell}{\operatorname{Im}\rho_\ell}, \qquad V_\ell = \frac{X_\ell + \rho_{\ell+1}Y_\ell}{\sqrt{n_0}\,\hat{s}(t)}. \tag{64}$$

The random variable V_ℓ is measurable with respect to $\mathcal{F}_{\ell+1}$, but independent of $\mathcal{F}_\ell$.

By Taylor expansion we have the following estimates for the deterministic part of $Z_{\ell,\lambda}$:

$$v_{\ell,\lambda} = \frac{v_\lambda(t)}{n_0} + \mathcal{O}(k^{-2}), \qquad v_\lambda(t) = -\frac{\lambda}{2\hat{s}(t)} + \frac{\frac{d}{dt}\rho(t)}{\operatorname{Im}\rho(t)}, \qquad |v_\lambda(t)| \leq c\frac{n_0}{k}, \tag{65}$$

where we abbreviate $\rho(t) = \rho(n_0/n, tn_0/n) = \rho_\ell$, see (48). The behavior of the random term is governed by

$$\begin{aligned} &\mathrm{E}V_\ell = \mathcal{O}(n-\ell)^{-3/2}k^{-1/2} \qquad \mathrm{E}|V_\ell^2| = \frac{1}{n_0}p(t) + \mathcal{O}(n-\ell)^{-1}k^{-1}, \\ &\mathrm{E}V_\ell^2 = \frac{1}{n_0}q(t) + \mathcal{O}(n-\ell)^{-1/2}k^{-3/2}, \qquad \mathrm{E}|V_\ell|^d = \mathcal{O}(k^{-d/2}), \quad d = 3, 4, \end{aligned} \tag{66}$$

where

$$p(t) = \frac{4}{\beta\hat{s}^2} = \frac{4n_0}{\beta k}, \qquad q(t) = \frac{2(1+\rho^2)}{\beta\hat{s}^2}. \tag{67}$$

Here the error terms come from the moment asymptotics (40), the size of $\hat{s}$, and from the bounds

$$\rho_{\ell+1} - \rho_\ell = \mathcal{O}(n-\ell)^{-1/2}k^{-1/2}, \qquad \frac{d}{dt}\rho - \frac{\rho_{\ell+1}-\rho_\ell}{n_0} = \mathcal{O}(n-\ell)^{-1/2}k^{-3/2}.$$

Proposition 22 (Single-step asymptotics for $\varphi_{\ell,\lambda}$) *For $\ell \le n_0$ with $t = \ell/n_0$ and $k = n_0 - \ell$ we have*

$$\mathrm{E}\big[\Delta\varphi_{\ell,\lambda}\big|\varphi_{\ell,\lambda} = x\big] = \frac{1}{n_0}b_\lambda(t) + \frac{1}{n_0}osc_1 + \mathcal{O}(k^{-3/2}) = \mathcal{O}(k^{-1}), \quad (68)$$

$$\mathrm{E}\big[\Delta\varphi_{\ell,\lambda}\Delta\varphi_{\ell,\lambda'}\big|\varphi_{\ell,\lambda} = x, \varphi_{\ell,\lambda'} = y\big] = \frac{1}{n_0}a(t,x,y) + \frac{1}{n_0}osc_2 + \mathcal{O}(k^{-3/2}), \quad (69)$$

$$\mathrm{E}_\ell|\Delta\varphi_{\ell,\lambda}|^d = \mathcal{O}(k^{-d/2}), \quad d = 2,3,$$

where

$$b_\lambda = \frac{\lambda}{2\hat{s}} - \frac{\mathrm{Re}\,\frac{d}{dt}\rho}{\mathrm{Im}\,\rho} + \frac{\mathrm{Im}(\rho^2)}{2\beta\hat{s}^2}, \qquad a = \frac{2}{\beta\hat{s}^2}\mathrm{Re}\big[e^{i(y-x)}\big] + \frac{3+\mathrm{Re}\,\rho^2}{\beta\hat{s}^2}. \quad (70)$$

The oscillatory terms are

$$osc_1 = \mathrm{Re}\big((-v_\lambda - iq/2)e^{-ix}\eta_\ell\big) + \mathrm{Re}\big(ie^{-2ix}\eta_\ell^2 q\big)/4, \quad (71)$$

$$osc_2 = p\,\mathrm{Re}\big(e^{-ix}\eta_\ell + e^{-iy}\eta_\ell\big)/2 + \mathrm{Re}\big(q(e^{-ix}\eta_\ell + e^{-iy}\eta_\ell + e^{-i(x+y)}\eta_\ell^2)\big)/2.$$

Proof By Proposition 18(iii) the difference $\Delta\varphi_{\ell,\lambda}$ can be written

$$\begin{aligned}\Delta\varphi_{\ell,\lambda} &= \mathrm{ash}(\mathbf{S}_{\ell,\lambda}, -1, z\bar{\eta}) \\ &= \mathrm{Re}\left[-(1+\bar{z}\eta)Z - \frac{i(1+\bar{z}\eta)^2}{4}Z^2\right] + \mathcal{O}(Z^3) \\ &= -\mathrm{Re}\,Z + \frac{\mathrm{Im}\,Z^2}{4} + \eta\text{ terms} + \mathcal{O}(Z^3), \end{aligned} \quad (72)$$

where we used $Z = Z_{\ell,\lambda}$, $\eta = \eta_\ell$ and $z = \exp(i\varphi_{\ell,\lambda})$. The estimate (72) is from the quadratic expansion (37) of the angular shift in Lemma 16. Note that since the second argument of ash is -1, we do not need an upper bound on $|Z|$. We take expectations, the error term becomes

$$\mathcal{O}(\mathrm{E}|Z|^3) = \mathcal{O}(|v_{\ell,\lambda}|^3 + \mathrm{E}|V_\ell|^3) = \mathcal{O}(k^{-3/2}).$$

By (63), (66) we may replace the $\mathrm{E}Z$, $\mathrm{E}|Z|^2$ and $\mathrm{E}Z^2$ terms by $v_\lambda(t)$, $p(t)$ and $q(t)$ while picking up an error term of $\mathcal{O}(k^{-2})$. Significant contributions come only from the non-random terms $v_{\ell,\lambda}$ of Z and the expectation of V_ℓ^2. We are then left with oscillatory terms with η, error terms, and the main term

$$\begin{aligned}-\mathrm{Re}\,v_{\ell,\lambda} + \mathrm{Im}\,\mathrm{E}V_\ell^2/4 &= \frac{1}{n_0}(-\mathrm{Re}\,v_\lambda + \mathrm{Im}\,q) + \mathcal{O}(k^{-3/2}) \\ &= \frac{1}{n_0}\left[\frac{\lambda}{2\hat{s}} - \frac{\mathrm{Re}\,\rho'}{\mathrm{Im}\,\rho} + \frac{\mathrm{Im}(\rho^2)}{2\beta\hat{s}^2}\right] + \mathcal{O}(k^{-3/2}).\end{aligned}$$

The error terms come from the moment bounds (40) of X, Y and from the discrete approximation of the derivative $\operatorname{Re}\rho'$; their exact order is readily computed. This gives (68) with (71). The $\mathcal{O}(k^{-1})$ bound comes from evaluating the continuous functions in the main and oscillatory terms at $t = \ell/n_0$.

For $\mathrm{E}_\ell\left[\Delta\varphi_{\lambda,\ell}\Delta\varphi_{\lambda',\ell}\right]$ one uses the linear approximation of the angular shift to get

$$\Delta\varphi_{\ell,\lambda} = \operatorname{Re}[-(1+\bar{z}\eta)Z_{\ell,\lambda}] + \mathcal{O}(Z_{\ell,\lambda}^2)$$

and similarly for λ'. After multiplying the two estimates and taking expectations, only the noise terms in $Z_{\ell,\lambda}$, $Z_{\ell,\lambda'}$ contribute. Namely, with $V = V_\ell$, we have

$$\begin{aligned}\mathrm{E}_\ell[\Delta\varphi_{\ell,\lambda}\Delta\varphi_{\ell,\lambda'}] &= \frac{1}{4}\mathrm{E}\left[(1+\eta\bar{z})V + (1+\bar{\eta}z)\bar{V}\right]\left[(1+\eta\bar{z}')V + (1+\bar{\eta}z')\bar{V}\right] + \mathcal{O}(k^{-3/2}) \\ &= \frac{1}{2}\operatorname{Re}(1+\bar{z}z')\mathrm{E}|V|^2 + \frac{1}{2}\operatorname{Re}\mathrm{E}V^2 + \eta \text{ terms } + \mathcal{O}(k^{-3/2}),\end{aligned}$$

where we used $z = \exp(i\varphi_{\ell,\lambda})$ and $z' = \exp(i\varphi_{\ell,\lambda'})$. Formula (69) now follows from the asymptotics of $\mathrm{E}|V|^2, \mathrm{E}V^2$. The last claim follows from the third moment asymptotics of X, Y. □

5.2 Continuum limit of the phase evolution

The goal of this section is to show that the first stretch of the phase evolution converges in law to the solution of the SDE (24). Typically, the phase evolves in an oscillatory manner, so we have to take advantage of averaging. Our main tool will be the following proposition, based on Stroock and Varadhan [17] and Ethier and Kurtz [7], which allows for averaging of the discrete evolutions.

Proposition 23 *Fix $T > 0$, and for each $n \geq 1$ consider a Markov chain*

$$(X_\ell^n \in \mathbb{R}^d,\ \ell = 1 \ldots \lfloor nT \rfloor).$$

Let $Y_\ell^n(x)$ be distributed as the increment $X_{\ell+1}^n - x$ given $X_\ell^n = x$. We define

$$b^n(t,x) = nE[Y_{\lfloor nt \rfloor}^n(x)], \qquad a^n(t,x) = nE[Y_{\lfloor nt \rfloor}^n(x)Y_{\lfloor nt \rfloor}^n(x)^{\mathrm{T}}].$$

Suppose that as $n \to \infty$ we have

$$|a^n(t,x) - a^n(t,y)| + |b^n(t,x) - b^n(t,y)| \leq c|x-y| + o(1), \tag{73}$$

$$\sup_{x,\ell} E[|Y_\ell^n(x)|^3] \leq cn^{-3/2}, \tag{74}$$

and that there are functions a, b from $\mathbb{R} \times [0,T]$ to $\mathbb{R}^{d^2}, \mathbb{R}^d$ respectively with bounded first and second derivatives so that

$$\sup_{x,t}\left|\int_0^t a^n(s,x)\,ds - \int_0^t a(s,x)\,ds\right| + \sup_{x,t}\left|\int_0^t b^n(s,x)\,ds - \int_0^t b(s,x)\,ds\right| \to 0. \tag{75}$$

Assume also that the initial conditions converge weakly:

$$X_0^n \overset{d}{\Longrightarrow} X_0.$$

Then $(X^n_{\lfloor nt\rfloor}, 0 \le t \le T)$ *converges in law to the unique solution of the SDE*

$$dX = b\,dt + a\,dB, \qquad X(0) = X_0.$$

We will prove this in Appendix A.3. The next lemma provides the averaging conditions for the above proposition. Recall that $\Delta\varphi_{\ell,\lambda} = \varphi_{\lambda,\ell+1} - \varphi_{\ell,\lambda}$.

Lemma 24 *Fix* λ, λ' *and* $\varepsilon > 0$. *Then for any* $\ell_1 \le n_0(1-\varepsilon)$

$$\begin{aligned} \frac{1}{n_0}\sum_{\ell=0}^{\ell_1-1} \mathrm{E}\big[\Delta\varphi_{\ell,\lambda} \,|\, \varphi_{\ell,\lambda} = x\big] &= \frac{1}{n_0}\sum_{\ell=0}^{\ell_1-1} b_\lambda(t) + \mathcal{O}(\mu_n n_0^{-3/2} + n_0^{-1/2}), \\ \frac{1}{n_0}\sum_{\ell=0}^{\ell_1-1} \mathrm{E}\big[\Delta\varphi_{\ell,\lambda}\Delta\varphi_{\ell,\lambda'} \,|\, \varphi_{\ell,\lambda} = x,\ \varphi_{\ell,\lambda'} = y\big] &= \frac{1}{n_0}\sum_{\ell=0}^{\ell_1-1} a(t,x,y) + \mathcal{O}(\mu_n n_0^{-3/2} + n_0^{-1/2}), \end{aligned} \tag{76}$$

where $t = \ell/n_0$, *the functions* b_λ, a *are defined in* (70), *and the implicit constants in* $\mathcal{O}$ *depend only on* $\varepsilon, \beta, \bar{\lambda}$.

Proof Summing (68) we get (76) with a preliminary error term

$$\frac{1}{n_0}\sum_{\ell=0}^{\ell_1-1} \mathrm{Re}(e_{1,\ell}\,\eta_\ell) + \frac{1}{n_0}\sum_{\ell=0}^{\ell_1-1} \mathrm{Re}(e_{2,\ell}\,\eta_\ell^2) + \mathcal{O}(k_1^{-1/2}),$$

where the first two terms will be denoted ζ_1, ζ_2. Here

$$e_{1,\ell} = (-v_\lambda(t) - iq(t)/2)e^{-ix}, \qquad e_{2,\ell} = iq(t)e^{-2ix}/4$$

and $k_1 = n_0 - \ell_1 > cn_0$, where for this proof c denotes varying constants depending on ε. Using the fact that v_λ, q and their first derivatives are continuous on $[0, 1-\varepsilon]$ we get

$$|e_{i,\ell}| < c, \qquad |e_{i,\ell} - e_{i,\ell+1}| < cn_0^{-1}.$$

Thus by the oscillatory sum Lemma 37,

$$|\zeta_1| \le c\sum_{\ell=0}^{\ell_1-1} (\mu_n/\sqrt{k} + 1)n_0^{-2} + c(\mu_n/\sqrt{k_1} + 1)n_0^{-1} \le c(\mu_n n_0^{-3/2} + n_0^{-1}).$$

Similarly, if we apply the same estimate for the second sum, we get

$$|\zeta_2| \le c\sum_{\ell=0}^{\ell_1-1} (\mu_n/\sqrt{k} + \sqrt{n_0}/\mu_n)n_0^{-2} + c(\mu_n/\sqrt{k_1} + \sqrt{n_0}/\mu_n)n_0^{-1} \le c(\mu_n n_0^{-3/2} + \mu_n^{-1} n_0^{-1/2}).$$

We could also estimate ζ_2 by taking absolute value in each term. Using (66) together with (47) we get

$$|q(t)| = \frac{\mu_n}{\beta\hat{s}(t)^2\sqrt{\mu_n^2/4 + k - 1/2}} \le C\mu_n k^{-3/2} n_0^{-1}$$

which leads to

$$|\zeta_2| \le c \sum_{\ell=0}^{\ell_1-1} \mu_n k^{-3/2} \le c\mu_n n_0^{-1/2}.$$

Using this bound for $\mu_n \le 1$ and the previous one for $\mu_n > 1$ we get the desired estimate (76). The asymptotics of the second sum follow similarly. □

We are now ready to state and prove the continuum limit theorem.

Theorem 25 (Continuum limit of the phase function) *Suppose that $n_0/n \to 1/(1+\nu)$ with $\nu \in [0,\infty]$. Then the continuous function $\rho(t) = \rho(n_0/n, tn_0/n)$ (see (47)) converges to a limit for which we use the same notation. Let B and Z be a real and a complex Brownian motion, and for each $\lambda \in \mathbb{R}$ consider the strong solution of*

$$d\varphi_\lambda = \left[\frac{\lambda}{2\hat{s}} - \frac{\operatorname{Re}\rho'}{\operatorname{Im}\rho} + \frac{\operatorname{Im}(\rho^2)}{2\beta\hat{s}^2}\right]dt + \frac{\sqrt{2}\operatorname{Re}(e^{-i\varphi_\lambda}dZ)}{\sqrt{\beta}\,\hat{s}} + \frac{\sqrt{3+\operatorname{Re}\rho^2}}{\sqrt{\beta}\,\hat{s}}dB, \tag{77}$$

$$\varphi_\lambda(0) = \pi.$$

Then we have

$$\varphi_{\lambda,\lfloor n_0 t\rfloor} \overset{d}{\Longrightarrow} \varphi_\lambda(t), \quad \text{as } n \to \infty,$$

where the convergence is in the sense of finite dimensional distributions for λ and in path-space $D[0,1)$ for t.

Remark 26 From (47) we get that the limit of $\rho(n_0/n, tn_0/n)$ as $n_0/n \to 1/(1+\nu)$ is

$$\rho(t) = \frac{\nu}{\nu+1-t} + i\frac{\sqrt{(1-t)(2\nu+1-t)}}{\nu+1-t}$$

with $\rho(t) = 1$ if $\nu = \infty$. Thus (77) can be written as

$$\sqrt{1-t}\,d\varphi_\lambda = \frac{\lambda}{2}dt + \sqrt{\frac{2}{\beta}}\operatorname{Re}(e^{-i\varphi_\lambda}dZ_t) + \left(\frac{1}{\beta} - \frac{1}{2}\right)\frac{\sqrt{\nu}}{\nu+1-t}dt$$
$$+ \sqrt{\frac{2(2\nu+1-t)}{\beta(\nu+1-t)}}dB, \tag{78}$$

where the last two terms are 0 and $2\beta^{-1/2}dB$, respectively when $\nu = \infty$.

Proof of Theorem 25 It suffices to show that for any finite sequence $(\lambda_1, \dots, \lambda_d)$ and for any $T < 1$ the following holds on the time interval $[0, T]$,

$$(\varphi_{\lfloor n_0 t\rfloor,\lambda_1}, \dots, \varphi_{\lfloor n_0 t\rfloor,\lambda_d}) \overset{d}{\Longrightarrow} (\varphi_{\lambda_1}(t), \dots, \varphi_{\lambda_d}(t)).$$

We will use Proposition 23. For $x \in \mathbb{R}^d$ let

$$\underline{\varphi}_\ell = (\varphi_{\ell,\lambda_1}, \dots, \varphi_{\ell,\lambda_d}), \qquad \Delta\underline{\varphi}_\ell = \underline{\varphi}_{\ell+1} - \underline{\varphi}_\ell,$$
$$b_\ell(x) = n_0 \mathrm{E}\left[\Delta\underline{\varphi}_\ell \middle| \underline{\varphi}_\ell = x\right], \qquad a_\ell(x) = n_0 \mathrm{E}\left[(\Delta\underline{\varphi}_\ell)(\Delta\underline{\varphi}_\ell)^T \middle| \underline{\varphi}_\ell = x\right].$$

Recall the estimates (68) and (69). Since $\mu_n^2/(4n_0) \to \nu$, the functions b_λ, a defined in (70) converge uniformly on $[0, T]$ to $\hat{b}_\lambda, \hat{a}$ which are also defined by (70) but in terms of the limit of ρ (recall that $\hat{s}$ is just $\sqrt{1-t}$).

Using this with Lemma 24 we get that

$$\begin{aligned} \sup_{x\in R^d, t\le T} \left| \int_0^t n_0 b_{\lfloor n_0 s\rfloor}(x)ds - \int_0^t \tilde{b}(x,s)ds \right| &\to 0, \\ \sup_{x\in R^d, t\le T} \left| \int_0^t n_0 a_{\lfloor n_0 s\rfloor}(x)ds - \int_0^t \tilde{a}(x,s)ds \right| &\to 0, \end{aligned} \tag{79}$$

where

$$\tilde{b}(x,t) = \big(\hat{b}_{\lambda_1}(t), \dots, \hat{b}_{\lambda_d}(t)\big), \qquad \big(\tilde{a}(x,t)\big)_{j,k} = \hat{a}(t, x_j, x_k).$$

This means that condition (75) in Proposition 23 is satisfied. Because of (70) and the moment bounds we can see that (73) and (74) are also satisfied, thus $(\varphi_{\lambda,\lfloor n_0 t\rfloor}, \dots, \varphi_{\lambda_d,\lfloor n_0 t\rfloor})$ converges weakly to the SDE corresponding to $\tilde{b}(x,t), \tilde{a}(x,t)$. The only thing left is to identify the limiting SDE from the functions $\tilde{b}(x,t), \tilde{a}(x,t)$. This follows easily, by observing that if Z is a complex Gaussian with independent standard real and imaginary parts and $\omega_1, \omega_2 \in \mathbb{C}$ then

$$\begin{aligned} &\mathrm{E}\,\mathrm{Re}(\omega_1 Z)\,\mathrm{Re}(\omega_2 Z) = \mathrm{E}(\omega_1 Z + \overline{\omega_1 Z})(\omega_2 Z + \overline{\omega_2 Z})/4 = (\omega_1\overline{\omega}_2 + \omega_2\overline{\omega}_1)/2 \\ &\quad = \mathrm{Re}(\omega_1\overline{\omega}_2). \end{aligned}$$

□

Theorem 25 leads to the following corollary.

Corollary 27 *Let W_t be complex Brownian motion with standard real and imaginary parts and consider the strong solution of the following one-parameter family of SDEs*

$$\sqrt{1-t}\, d\alpha_\lambda = \frac{\lambda}{2}\, dt + \sqrt{2/\beta}\ \mathrm{Re}((e^{-i\alpha_\lambda} - 1)dW_t).$$

Then

$$\alpha_{\lfloor n_0 t\rfloor,\lambda} \overset{d}{\Longrightarrow} \alpha_\lambda(t), \quad \text{as } n \to \infty, \tag{80}$$

where the convergence is in the sense of finite dimensional distributions for λ and in path-space $D[0,1)$ for t.

Proof If $\mu_n^2/(4n_0)$ converges to a finite or infinite value as $n \to \infty$ then the statement follows immediately with $W_t = e^{i\varphi_0(t)} Z_t$. This implies that for any subsequence of n we can choose a further subsequence along which (80) holds, a characterization of convergence. □

5.3 Why are the limits in different windows the same? Universality and non-universality

This subsection is meant to explain why the continuum limit of the relative phase function does not depend on the choice of the scaling window μ_n. In order to do that, we will discuss a more general model where this is not necessarily true.

The discussion of this section is not an integral part of the proof of the main theorem; the goal is to provide some additional insight for the results.

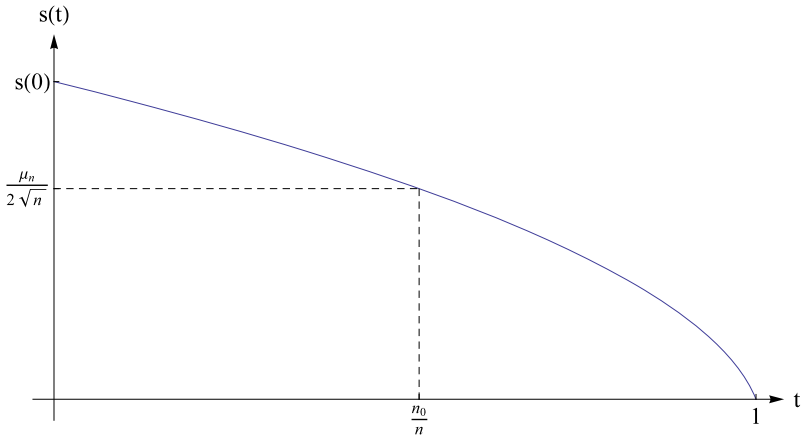

Fig. 2 The definition of the scaling parameter n_0

A more general model The following is a generalization of the model (39). Consider random tridiagonal $n \times n$ matrices with diagonal elements $X_0, X_1, X_2, \ldots$ and off-diagonal elements $s_1, s_2, s_3, \ldots$ and $s_0 + Y_0, s_1 + Y_1, s_2 + Y_2 \ldots$, see (39). The random variables X_i, Y_i are independent with mean approximately zero, variance approximately $2/\beta$ and a bounded fourth moment. The deterministic numbers s_ℓ depend on n and are approximately $\sqrt{n}s(\ell/n)$, where $s(t)$ is a nonnegative, sufficiently smooth decreasing function on $[0, 1]$. In the case of β-ensembles we have $s(t) = \sqrt{1-t}$.

We will try to understand the point process limit of the eigenvalues of these tridiagonal matrices near μ_n where the scaling parameter μ_n will be in the interval $[s(1)\sqrt{n}, s(0)\sqrt{n}]$. It turns out that in this more general setup the arguments of the previous two sections follow through essentially without change. This is the main reason why we expressed everything in terms of $\rho, s, \hat{s}$, instead of using the sometimes more simple explicit values.

We first have to identify the scaling around μ_n so that we have a nontrivial limit, for this we consider (47). We define $n_0 \in [0, n]$ as the unique value for which

$$s(n_0/n) = \mu_n/\sqrt{4n}. \tag{81}$$

Then for $\ell \in [0, n_0]$ the complex number $\rho_\ell \in \mathbb{H}$ defined through (47) (see also (48)) is of unit length and our scaling around μ_n will be given by (50):

$$\Lambda = \mu_n + \lambda/(2\sqrt{n_0}).$$

The subsequent computations, i.e. the introduction of the slowly varying phase function, the single-step asymptotics and the continuum limit of the phase evolution can be carried out the same way as we have done in Sects. 4.2 and 4.3 and Sect. 5. The assumption $n_0^{-1}\mu_n^{2/3} \to 0$ as $n \to \infty$ will ensure that the arising error terms are negligible.

Thus, according to Theorem 25 if n_0/n converges to a constant as $n \to \infty$ then $\varphi_{\lfloor n_0 t \rfloor, \lambda} \xrightarrow{d} \varphi_\lambda(t)$ where $\varphi_\lambda(t)$ is the solution of (77). This gives the following limit for the relative phase function $\alpha_{\lfloor n_0 t \rfloor, \lambda}$:

$$\hat{s}\, d\alpha_\lambda = \frac{\lambda}{2} dt + \sqrt{2/\beta}\, \mathrm{Re}\big((e^{-i\alpha_\lambda} - 1) dW_t\big), \qquad \alpha_\lambda(0) = 0. \tag{82}$$

Here $\hat{s}$ is given by the $n \to \infty$ limit of

$$\hat{s}_n(t) = \sqrt{\frac{s(n_0/nt)^2 - s(n_0/n)^2}{n_0/n}}. \tag{83}$$

There are various ways of interpreting (82), perhaps the most intuitive is via the Brownian carousel which is already apparent in the discrete evolution, see Sect. 4.4.

The fact that the limiting equation depends on s through the function $\hat{s}$ explains two phenomena. First, in the β-ensemble case $s(t) = \sqrt{1-t}$ and thus (83) gives that $\hat{s}(t) = \sqrt{1-t}$ *regardless* of the value of the limit of n_0/n. This shows why all limits will be governed by the same stochastic differential equation, regardless on the choice of the scaling parameter. However, in the more general model, non-universality holds; the limiting stochastic differential equation (82) depends not only on the limit of s but also on the scaling window.

Second, consider a general s, with $s'(0) < 0$, and choose μ_n so that

$$(2\sqrt{n}s(0) - \mu_n)n^{1/6} \to \infty, \quad \text{and} \quad (2\sqrt{n}s(0) - \mu_n)n^{-1/2} \to 0. \tag{84}$$

This means that scaling parameter μ_n is close, but not too close to the edge of the spectrum $2s(0)\sqrt{n}$. Since $\mu_n/\sqrt{4n} \to s(0)$, by (81) we have $n_0/n \to 0$ and

$$n_0 = s'(0)^{-1}\sqrt{n}(\sqrt{n}s(0) - \mu_n/2) + o(n_0).$$

Thus $n_0^{-1}\mu_n^{2/3} \to 0$, so we can apply our previous results. From (83) we get

$$\hat{s}(t) = \lim \hat{s}_n(t) = c\sqrt{1-t}, \quad c = |(s^2)'(0)|^{1/2}$$

which means that the limiting sde (82) is the same as in the β-ensemble case, after a linear rescaling with a new parameter $\tilde{\beta} = \beta c^2$. This means that even for a general choice of s the point process limit of the eigenvalues in the scaling regime (84) is given universally by the Sine_β process.

A similar statement of universality holds for a class of 1-dimensional discrete random Schrödinger operators with tridiagonal matrix representation. Consider the a symmetric tridiagonal matrix with diagonal and off-diagonal terms

$$\tilde{X}_0, \tilde{X}_1, \ldots \qquad \tilde{Z}_0/2 + s_0, \tilde{Z}_1/2 + s_1, \ldots,$$

where $\tilde{X}_i, \tilde{Z}_i$ are independent random variables with mean approximately zero, variance approximately σ^2 and a bounded eighth moment. The deterministic numbers s_ℓ depend on n and are approximately $\sqrt{n}s(\ell/n)$, where $s(t)$ is again a nonnegative, sufficiently smooth decreasing function on $[0, 1]$. This gives the matrix representation of a 1-dimensional discrete random Schrödinger operator.

The analyze its spectrum, we first conjugate it with a diagonal matrix to transform it into a form similar to (39). Choosing an appropriate diagonal matrix we can transform any off-diagonal pair $(\tilde{Z}_\ell/2 + s_\ell, \tilde{Z}_\ell/2 + s_\ell)$ into $(a_\ell, (\tilde{Y}_\ell/2 + s_\ell)^2/a_\ell)$ with any nonzero a_ℓ, while the diagonal elements stay the same. A simple computation shows that if we set $a_\ell = \sqrt{n}s(\ell/n + 1/(2n)) \simeq \sqrt{n}s((\ell+1)/n)$ then the off-diagonal entries above the diagonal will have mean approximately equal to $\sqrt{n}s(\ell/n)$, variance approximately equal to σ^2 and a bounded fourth moment. Thus the previous results may be applied with $\beta = 2/\sigma^2$. In particular near the edge of the spectrum (but not very near: see (84)) the point process limit of the eigenvalues will be given universally by the Sine_β process.

We would like to note that the point process limit near the edge of the spectrum (i.e. when $(\mu_n - 2\sqrt{n}s(0))n^{1/6}$ converges to a finite constant) one gets the Airy_β process (see [15], Sect. 5). This allows us to complete the proof in the general case with arguments analogous to the following section. To avoid excessive technicalities, we chose to focus on the beta ensemble case in this paper. We plan to treat the more general case in detail in a future work.

6 Asymptotics in the uneventful stretch

Section 5 describes the stochastic differential equation limit of the phase function on the first stretch $[0, n_0(1-\varepsilon)]$. Here we show that in the limit, this stretch completely determines the eigenvalue behavior.

6.1 The uneventful middle stretch

The middle stretch is the discrete time interval $[m_1, m_2]$ with

$$m_1 = \lfloor n_0(1-\varepsilon)\rfloor, \qquad m_2 = \lfloor n - \mu_n^2/4 - \kappa(\mu_n^{2/3} \vee 1)\rfloor \tag{85}$$

for $\varepsilon \in (0,1)$ and $\kappa > 0$. The goal of this section is to prove that if $\alpha_{\ell,\lambda}$ is close to an integer multiple of 2π after time m_1 then it changes little up to time m_2. More precisely, we have

Proposition 28 *There exists a constant $c = c(\bar{\lambda}, \beta)$ so that with $y = n_0^{-1/2}(\mu_n^{1/3} \vee 1)$ we have*

$$\mathrm{E}\big[|\alpha_{\ell_2,\lambda} - \alpha_{\ell_1,\lambda}| \wedge 1 \big| \mathcal{F}_{\ell_1}\big] \le c(\mathrm{dist}(\alpha_{\ell_1,\lambda}, 2\pi\mathbb{Z}) + \sqrt{\varepsilon} + y + \kappa^{-1}), \tag{86}$$

for all $\kappa > 0$, $\varepsilon \in (0,1)$, $\lambda \le |\bar{\lambda}|$ and $m_1 \le \ell_1 \le \ell_2 \le m_2$.

The first step is to estimate $\Delta\alpha_{\ell,\lambda} = \alpha_{\ell+1,\lambda} - \alpha_{\ell,\lambda}$ using the angular shift Lemma 16 with $z = Z_{\ell,\lambda}$ defined in (63). For the finer asymptotics of the lemma, the condition $|z| < 1/3$ is needed. For this, we truncate the original random variables X_ℓ, Y_ℓ. For $m_1 \le \ell \le m_2$, introduce the random variables $\tilde{X}_\ell, \tilde{Y}_\ell$ which agree with X_ℓ, Y_ℓ on the event

$$|X_\ell|, |Y_\ell| \le \frac{1}{10}\sqrt{n_0}\hat{s}(\ell/n_0), \tag{87}$$

and are zero otherwise; this event depends on n via $\hat{s}$. By Markov's inequality and the fourth moment assumption (40) for X_ℓ, Y_ℓ, this event has probability at least $1 - c(n_0 - \ell)^{-2}$. Summing this for $\ell \le m_2$ shows that the total probability that the truncation has an effect is at most $c\kappa^{-1}$. This can be absorbed in the error term κ^{-1} in (86), so it suffices to prove Proposition 28 for the truncated random variables.

To keep the notation under control, we will drop the tildes and instead modify the assumptions on X_ℓ, Y_ℓ. Namely, denoting $k = n_0 - \ell$ we assume the bounds (87) and the modified moment asymptotics

moment	1st	2nd	4th
	$\mathcal{O}(k^{-3/2})$	$2/\beta + \mathcal{O}(k^{-1})$	$\mathcal{O}(1)$

which follow from the original ones (40) and our choice of truncation. With p, q defined in (67), this changes the moment asymptotics of V_ℓ (64) the following way:

$$\begin{array}{c|c|c|c} \mathrm{E}V_\ell & \mathrm{E}V_\ell^2 & \mathrm{E}|V_\ell^2| & \mathrm{E}|V_\ell|^4 \\ \hline \mathcal{O}(k^{-2}) & \frac{1}{n_0}q(t)+\mathcal{O}(k^{-2}) & \frac{1}{n_0}p(t)+\mathcal{O}(k^{-2}) & \mathcal{O}(k^{-2}) \end{array}. \tag{88}$$

Proposition 29 (Single-step asymptotics for $\alpha_{\ell,\lambda}$) *There exists $k^* = k^*(\beta, \bar{\lambda})$ so that for every $\ell \le n_0 - k^*$ and $|\lambda| < \bar{\lambda}$ we have the following.*

$$\begin{aligned} \mathrm{E}_\ell\left[\Delta\alpha_{\ell,\lambda}\right] &= \frac{1}{n_0}\mathrm{Re}\big[(e^{-i\varphi_{\ell,\lambda}} - e^{-i\varphi_{\ell,0}})\eta_\ell(-v_\lambda - iq/2)\big] \\ &\quad + \frac{1}{n_0}\mathrm{Re}\left[\frac{iq}{4}(e^{-2i\varphi_{\ell,\lambda}} - e^{-2i\varphi_{\ell,0}})\eta_\ell^2\right] + \mathcal{O}(n_0^{-1/2}k^{-1/2} + k^{-3/2}\hat{\alpha}_{\ell,\lambda}) \quad (89) \\ &= \mathcal{O}(k^{-1}\hat{\alpha}_{\ell,\lambda} + n_0^{-1/2}k^{-1/2}), \quad (90) \\ \mathrm{E}_\ell\left[(\Delta\alpha_{\ell,\lambda})^2\right] &= \mathcal{O}(k^{-1}\hat{\alpha}_{\ell,\lambda} + n_0^{-1}k^{-1}), \quad (91) \\ \mathrm{E}_\ell|\Delta\varphi_{\ell,\lambda}\Delta\alpha_{\ell,\lambda}| &= \mathcal{O}(\hat{\alpha}_{\ell,\lambda}k^{-1}). \quad (92) \end{aligned}$$

The functions $v_\lambda = v_\lambda(\ell/n_0)$, $q = q(\ell/n_0)$ are defined in (65), (67), and $\hat{\alpha}_{\ell,\lambda}$ denotes the distance between $\alpha_{\ell,\lambda}$ and the set $2\pi\mathbb{Z}$.

Proof By choosing a large enough $k^* \ge 1$ we can assume that for $\ell \le n - k^*$

$$\frac{\bar{\lambda}}{k^2} \le \frac{1}{10}, \qquad |v_{\ell,\lambda}| \le \frac{1}{10}$$

which together with (87) guarantees that the random variable defined in (63) satisfies $|Z_{\ell,\lambda}| \le 1/3$. The proof of the proposition relies on the evolution rule, Proposition 19(ii),

$$\Delta\alpha_{\ell,\lambda} = \mathrm{ash}((\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell}, -1, e^{i\varphi_{\ell,\lambda}}\bar{\eta}_\ell) + \mathrm{ash}(\mathbf{S}_{\ell,0}, e^{i\varphi^*_{\ell,\lambda}}\bar{\eta}_\ell, e^{i\varphi_{\ell,\lambda}}\bar{\eta}_\ell) + \mathrm{ash}(\mathbf{S}_{\ell,0}, e^{i\varphi_{\ell,\lambda}}\bar{\eta}_\ell, e^{i\varphi_{\ell,0}}\bar{\eta}_\ell)$$

whose terms we denote $\zeta_1, \zeta_2, \zeta_3$. First we show that ζ_1, ζ_2 are small. By the definition (55) of L we have

$$\left|i.\big((\mathbf{L}_{\ell,\lambda})^{\mathbf{T}_\ell}\big)^{-1} - i\right| = \left|\frac{1}{n_0}\frac{\lambda}{2\hat{s}}\right| = \left|\frac{\lambda}{\sqrt{kn_0}}\right| \le \frac{1}{10}.$$

This estimate with the third bound of Lemma 16 gives

$$\zeta_1 = \varphi^*_{\ell,\lambda} - \varphi_{\ell,\lambda} = \mathcal{O}(n_0^{-1/2}k^{-1/2}). \tag{93}$$

Applying again the third bound of Lemma 16 with $|Z_{\ell,0}| \le 1/3$ and (93) gives

$$\zeta_2 = \mathcal{O}(\varphi^*_{\ell,\lambda} - \varphi_{\ell,\lambda}) = \mathcal{O}(n_0^{-1/2}k^{-1/2}).$$

For ζ_3 we use the first estimate of Lemma 16 and note that in our case $|v - w|$ equals

$$|e^{i\varphi_{\ell,\lambda}} - e^{i\varphi_{\ell,0}}| = |e^{i\alpha_{\ell,\lambda}} - 1| \le \hat{\alpha}_{\ell,\lambda}. \tag{94}$$

Thus with $Z = Z_{\ell,0}$ we have

$$\zeta_3 = -\mathrm{Re}\big[(e^{-i\varphi_{\ell,\lambda}} - e^{-i\varphi_{\ell,0}})\eta_\ell(Z + iZ^2/2)\big] + \mathrm{Re}\big[i(e^{-2i\varphi_{\ell,\lambda}} - e^{-2i\varphi_{\ell,0}})\eta_\ell Z^2/4\big] + \mathcal{O}(\hat{\alpha}_{\ell,\lambda}Z^3). \tag{95}$$

Since Z is independent of $\mathcal{F}_\ell$ and $\hat{\alpha}_{\ell,\lambda} \in \mathcal{F}_\ell$, the error term becomes $\mathcal{O}(\hat{\alpha}_{\ell,\lambda}k^{-3/2})$ after taking conditional expectation. The definition (63) of $Z_{\ell,\lambda}$ and the moment bounds (88) imply that replacing $\mathrm{E}Z$ and $\mathrm{E}Z^2$ by $v_\lambda(\ell/n_0)$ and $q(\ell/n_0)$ gives error terms of order $\mathcal{O}(k^{-2})$. Because of (94) and

$$|e^{-2i\varphi_{\ell,\lambda}} - e^{2i\varphi_{\ell,0}}| = |e^{2i\alpha_{\ell,\lambda}} - 1| \le 2\hat{\alpha}_{\ell,\lambda} \tag{96}$$

we get (89). Using the explicit form of v_λ and q and (94), (96) again, we obtain (90). The other estimates follow similarly from the first-order version of (95) and Proposition 22. □

The following lemma relies on the careful use of single-step bounds and oscillatory sum estimates. We postpone the proof till Sect. 6.2.

Lemma 30 *Recall the definition of m_1, m_2 from (85). There exist c_0, c_1 depending on $\bar{\lambda}, \beta$ so that with $y = n_0^{-1/2}(\mu_n^{1/3} \vee 1)$ we have*

$$|\mathrm{E}[\alpha_{\ell_2,\lambda} - \alpha_{\ell_1,\lambda}|\mathcal{F}_{\ell_1}]| \le c_1(y + \sqrt{\varepsilon}) + \frac{\mathrm{E}[\hat{\alpha}_{\ell_2-1}|\mathcal{F}_{\ell_1}]}{2} + \sum_{\ell=\ell_1}^{\ell_2-2} b_\ell \mathrm{E}[\hat{\alpha}_\ell|\mathcal{F}_{\ell_1}],$$

$$0 \le b_\ell \le c_1\left(k^{-3/2} + \mu_n k^{-5/2} + k^{-3/2}\mu_n \mathbf{1}_{k \ge \mu_n^2/4}\right),$$

whenever $\kappa > c_0$, $|\lambda| < \bar{\lambda}$, and $m_1 \le \ell_1 \le \ell_2 \le m_2$. Here $k = n_0 - \ell$.

The last ingredient needed for the proof of Proposition 28 is the following deterministic Gronwall-type estimate.

Lemma 31 (Gronwall estimate) *Suppose that for positive numbers x_ℓ, b_ℓ, c, integers $\ell_1 < \ell_2$ and $\ell = \ell_1 + 1, \ell_1 + 2, \dots, \ell_2$ we have*

$$x_\ell \le \frac{x_{\ell-1}}{2} + c + \sum_{j=\ell_1}^{\ell-1} b_j x_j. \tag{97}$$

Then

$$x_{\ell_2} \le 2(x_{\ell_1} + c)\exp\left(3\sum_{j=\ell_1}^{\ell_2-1} b_j\right).$$

Proof We can assume $\ell_1 = 0$. Let $y_\ell = x_\ell - x_{\ell-1}/2$, so that we have

$$x_\ell = y_\ell + \frac{y_{\ell-1}}{2} + \frac{y_{\ell-2}}{4} + \cdots + \frac{y_1}{2^{\ell-1}} + \frac{x_0}{2^\ell}. \tag{98}$$

Then (97) and the positivity of x_0 and b_j gives

$$y_\ell \le c + x_0 s + \sum_{j=1}^{\ell-1} b_j\left(y_\ell + \frac{y_{\ell-1}}{2} + \frac{y_{\ell-2}}{4} + \cdots + \frac{y_1}{2^{\ell-1}}\right), \tag{99}$$

where $s = b_0 + \cdots + b_{\ell_2 - 1}$. Taking positive parts in (99), and then summation by parts yields

$$y_\ell^+ \le (c + x_0 s) + \sum_{j=1}^{\ell-1} \tilde{b}_j y_j^+ \tag{100}$$

with $\tilde{b}_j = b_j + b_{j-1}/2 + \cdots + b_{\ell_2-1}/2^{\ell_2-j-1}$. Let ℓ_3 be so that $1 \le \ell_3 \le \ell_2$. We now multiply the inequality (100) by $\tilde{b}_\ell(1+\tilde{b}_{\ell+1})\cdots(1+\tilde{b}_{\ell_3-1})$ and sum it for $1 \le \ell \le \ell_3 - 1$. We add (100) again with $\ell = \ell_3$. After cancellations, we get

$$y_{\ell_3}^+ \le (c + x_0 s) \prod_{j=1}^{\ell_3-1} (1 + \tilde{b}_j) \le (c + x_0 s) \exp\left(\sum_{j=1}^{\ell_2-1} \tilde{b}_j \right) \le (c + x_0 s) e^{2s}.$$

Applying this inequality for all the y terms in (98) with $\ell = \ell_2$ we get

$$x_{\ell_2} \le 2(c + x_0 s) e^2 s + x_0 \le 2(x_0 + c) e^{3s}. \qquad \square$$

Proof of Proposition 28 For this proof, let $a = \alpha_{\ell_1,\lambda}$, and define $a_\diamond, a^\diamond \in 2\pi\mathbb{Z}$ so that $[a_\diamond, a^\diamond)$ is an interval of length 2π containing a. We condition on the σ-field $\mathcal{F}_{\ell_1}$, so in this proof E denotes the corresponding conditional expectation. We also drop the index λ from α. We will show that there exists c_0 so that if $\kappa > c_0$, then with the quantifiers of the proposition

$$\mathrm{E}|\alpha_{\ell_2} - a_\diamond| \le c_1((a - a_\diamond) + \sqrt{\varepsilon} + y), \tag{101}$$

$$\mathrm{E}|\alpha_{\ell_2} - a^\diamond| \le c_1((a^\diamond - a) + \sqrt{\varepsilon} + y). \tag{102}$$

The claim of the proposition follows from this by an application of the triangle inequality to the stronger bound. The additional condition $\kappa > c_0$ is treated via the error term $1/\kappa$.

Lemma 30 provides the bound

$$|\mathrm{E}\alpha_\ell - a_\diamond| \le (a - a_\diamond) + c(y + \sqrt{\varepsilon}) + \mathrm{E}\hat{\alpha}_{\ell-1}/2 + \sum_{j=\ell_1}^{\ell-2} b_j \mathrm{E}\hat{\alpha}_j.$$

Note that α never goes below an integer multiple of 2π that it passes (Proposition 19(iii)), so $\alpha_\ell \ge a_\diamond$ for all $\ell \ge \ell_1$. This means that for $\ell \ge \ell_1$ we have $\hat{\alpha}_\ell \le \alpha_\ell - a_\diamond$ and with $x_\ell = \mathrm{E}|\alpha_\ell - a_\diamond|$ we have the bound

$$x_\ell \le (a - a_\diamond) + c(y + \sqrt{\varepsilon}) + x_{\ell-1}/2 + \sum_{j=\ell_1}^{\ell-2} b_j x_j. \tag{103}$$

According to Lemma 30 we can bound the sum of the coefficients b_ℓ as

$$\sum_{\ell=\ell_1}^{\ell_2-2} b_\ell \le c(k_2^{-1/2} + \mu_n k_2^{-3/2} + 1) \le c'$$

which means that (101) follows via the Gronwall-type estimate of Lemma 31.

Next, we consider the first time $T \ge \ell_1$ so that $\alpha_T - a^\diamond \ge 0$. Proposition 19(ii) breaks one step of the evolution of α into two parts, from α_ℓ to $\alpha_{\ell+1}^*$ and from $\alpha_{\ell+1}^*$ to $\alpha_{\ell+1}$. It

shows that α can only pass an integer multiple of 2π in the first part. Since the first part is non-random, even the time $T-1$ (and not just T) is a stopping time adapted to our filtration. The overshoot can be estimated in two steps. By (93), and the fact that $k > c_0$ we have

$$\mathrm{E}\big[(\alpha_T^* - a^\diamond)\mathbf{1}(T \le \ell_2)\big] \le c n_0^{-1/2}. \tag{104}$$

By the expected increment bound (90) and the strong Markov property applied at $T-1$ we have

$$\mathrm{E}\big[(\alpha_T - \alpha_T^*)\mathbf{1}(T \le \ell_2)\big] \le c n_0^{-1/2}. \tag{105}$$

This gives

$$\begin{aligned}\mathrm{E}(\alpha_{\ell_2} - a^\diamond)^+ &= \mathrm{E}\big[\mathbf{1}(T \le \ell_2)\mathrm{E}[\alpha_{\ell_2} - a^\diamond \big| \mathcal{F}_T]\big] \\ &\le c_1(\mathrm{E}[(\alpha_T - a^\diamond)\mathbf{1}(T \le \ell_2)] + \sqrt{\varepsilon} + y) \\ &\le c_1'(\sqrt{\varepsilon} + y),\end{aligned} \tag{106}$$

where the first inequality uses (101) and the strong Markov property, and the second uses (104), (105). To prove (102) first note that Lemma 30 also gives

$$|\mathrm{E}\alpha_\ell - a^\diamond| \le (a^\diamond - a) + c(y + \sqrt{\varepsilon}) + \mathrm{E}\hat{\alpha}_{\ell-1}/2 + \sum_{j=\ell_1}^{\ell-2} b_j \mathrm{E}\hat{\alpha}_j.$$

Then by (106) and the identity $|a| = -a + 2a^+$ we get

$$\begin{aligned}\mathrm{E}|\alpha_\ell - a^\diamond| &\le |\mathrm{E}\alpha_\ell - a^\diamond| + 2\mathrm{E}(\alpha_\ell - a^\diamond)^+ \\ &\le (a^\diamond - a) + c(y + \sqrt{\varepsilon}) + \mathrm{E}\hat{\alpha}_{\ell-1}/2 + \sum_{j=\ell_1}^{\ell-2} b_j \mathrm{E}\hat{\alpha}_j.\end{aligned}$$

Since $\hat{\alpha}_\ell \le |\alpha_\ell - a^\diamond|$, the inequality (103) follows with $x_\ell = \mathrm{E}|\alpha_\ell - a^\diamond|$, and the Gronwall-type estimate in Lemma 31 implies (102). □

6.2 Bounds for oscillations in the middle stretch

This section presents the proof of Lemma 30, isolated as the most technical ingredient of the proof in the previous section. We start with a bound on the mixed differences.

Lemma 32 *There exists an absolute constant c so that for $\ell \le n - k^*$ (with k^* as in Proposition 29) we have*

$$|\mathrm{E}_\ell[\Delta e^{i\varphi_{\ell,\lambda}} - \Delta e^{i\varphi_{\ell,0}}]| \le c k^{-1}\hat{\alpha}_\ell + c n_0^{-1/2} k^{-1/2}$$

and the same inequality holds with $e^{2i\varphi}$ replacing $e^{i\varphi}$ on the left-hand side.

Proof The left-hand side equals

$$\begin{aligned}&\big|e^{i\varphi_{\ell,0}}\mathrm{E}_\ell\big[(e^{i\alpha_{\ell,\lambda}} - 1)(e^{i\Delta\varphi_{\ell,\lambda}} - 1) + (e^{i\Delta\alpha_{\ell,\lambda}} - 1)(e^{i\Delta\varphi_{\ell,0}} - 1) + (e^{i\Delta\alpha_{\ell,\lambda}} - 1)\big]\big| \\ &\le |e^{i\alpha_{\ell,\lambda}} - 1||\mathrm{E}_\ell[e^{i\Delta\varphi_{\ell,\lambda}} - 1]| + \mathrm{E}_\ell|\Delta\alpha_{\ell,\lambda}\Delta\varphi_{\ell,0}| + |\mathrm{E}_\ell[e^{i\Delta\alpha_{\ell,\lambda}} - 1]|.\end{aligned}$$

The statement now follows from (68), Proposition 29, the bounds (94), (96) and the bound

$$|\mathrm{E}[e^{iX}-1]| \le |\mathrm{E}[e^{iX}-iX-1]| + |\mathrm{E}(iX)| \le \mathrm{E}|e^{iX}-iX-1| + |\mathrm{E}X| \le \mathrm{E}X^2 + |\mathrm{E}X|.$$

The inequality involving $e^{2i\varphi}$ can be proved the same way. □

Now we are ready to prove Lemma 30.

Proof of Lemma 30 We will drop λ in $\alpha_{\ell,\lambda}$, and condition on the σ-field $\mathcal{F}_{\ell_1}$. Let E denote the conditional expectation with respect to this σ-field and let $x_\ell = \mathrm{E}\hat{\alpha}_\ell$. We have

$$\left|\mathrm{E}\left[\alpha_{\ell_2}-\alpha_{\ell_1}\right]\right| \le \left|\sum_{\ell=\ell_1}^{\ell_2-1} \mathrm{E}\left[\mathrm{E}\left(\Delta\alpha_\ell \middle| \mathcal{F}_\ell\right)\right]\right|. \tag{107}$$

Let

$$g_{1,\ell} = \frac{1}{n_0}(-v_\lambda - iq/2)\mathrm{E}(e^{-i\varphi_{\ell,\lambda}} - e^{-i\varphi_{\ell,0}}), \qquad g_{2,\ell} = \frac{1}{n_0}\frac{iq}{4}\mathrm{E}(e^{-2i\varphi_{\ell,\lambda}} - e^{-2i\varphi_{\ell,0}}).$$

By the single-step asymptotics (89) the right-hand side of (107) can be bounded by

$$\left|\sum_{\ell=\ell_1}^{\ell_2-1}\mathrm{Re}(g_{1,\ell}\eta_\ell)\right| + \left|\sum_{\ell=\ell_1}^{\ell^*}\mathrm{Re}(g_{2,\ell}\eta_\ell^2)\right| + \left|\sum_{\ell=\ell^*+1}^{\ell_2-1}\mathrm{Re}(g_{2,\ell}\eta_\ell^2)\right| + c\sum_{\ell=\ell_1}^{\ell_2-1}k^{-3/2}x_\ell + c\sum_{\ell=\ell_1}^{\ell_2-1}n_0^{-1/2}k^{-1/2},$$

with the usual notation $k = n_0 - \ell$. We call the terms $\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5$. Note that ζ_2, ζ_3 come from a single sum cut in two parts at $\ell^* = n_0 - \lfloor \mu_n^2/4 \rfloor$, and one part may be empty. Clearly, we have $\zeta_5 \le c\sqrt{\varepsilon}$, and ζ_4 is already in the desired form. By (94) and the bounds (65), (67) on q, v we have

$$|g_{1,\ell}| \le \frac{c}{n_0}|v_\lambda + iq/2|x_\ell \le \frac{c}{k}x_\ell.$$

Lemma 32 with $t_+ = (\ell+1)/n_0$ gives

$$\begin{aligned}|\Delta g_{1,\ell}| &\le \frac{c}{n_0}\Bigg(\left|v_\lambda(t_+) + \frac{iq(t_+)}{2}\right||\mathrm{E}[\Delta e^{i\varphi_{\ell,\lambda}} - \Delta e^{i\varphi_{\ell,0}}]| \\ &\quad + (|\Delta_\ell v_\lambda| + |\Delta_\ell q|)\,\mathrm{E}|e^{-i\varphi_{\ell,\lambda}} - e^{-i\varphi_{\ell,0}}|\Bigg) \\ &\le ck^{-2}x_\ell + cn_0^{-1/2}k^{-3/2},\end{aligned}$$

where we used the notation $\Delta_\ell f = f((\ell+1)/n_0) - f(\ell/n_0)$. The oscillatory sum Lemma 37 gives

$$\begin{aligned}\zeta_1 &\le c(\mu_n(n_0-\ell_2)^{-3/2} + (n_0-\ell_2)^{-1})x_{\ell_2-1} + c\sum_{\ell=\ell_1}^{\ell_2-2}(x_\ell k^{-2} + n_0^{-1/2}k^{-3/2})(\mu_n k^{-1/2}+1) \\ &\le \frac{x_{\ell_2-1}}{6} + c(\mu_n^{1/3}\vee 1)n_0^{-1/2} + c\sum_{\ell=\ell_1}^{\ell_2-2}x_\ell(k^{-2}+\mu_n k^{-5/2}),\end{aligned}$$

where the coefficient $1/6$ is achieved by choosing a large enough c_0. We continue

$$\zeta_2 \le \sum_{\ell=\ell_1}^{\ell^*} |g_{2,\ell}| \le \frac{c}{n_0}|q| \sum_{\ell=\ell_1}^{\ell^*} x_\ell \le c' \sum_{\ell=\ell_1}^{\ell^*} x_\ell \mu_n k^{-3/2}.$$

The term ζ_3 is handled by Lemma 37 with $g_j = g_{2,j}$. Standard bounds on q, q' and Lemma 32 give

$$|g_\ell| \le ck^{-1}x_\ell, \qquad |g_\ell - g_{\ell+1}| \le ck^{-2}x_\ell + cn_0^{-1/2}k^{-3/2},$$

hence from Lemma 37 we get

$$\begin{aligned}\zeta_3 &\le c(\mu_n(n_0-\ell_1)^{-1/2}+1)k^{-1}x_{\ell_2-1} + c\sum_{\ell=\ell^*+1}^{\ell_2-2}(x_\ell k^{-2} + n_0^{-1/2}k^{-3/2})(\mu_n k^{-1/2}+1)\\ &\le \frac{x_{\ell_2-1}}{6} + c(\mu_n^{1/3}\vee 1)n_0^{-1/2} + c\sum_{\ell=\ell^*+1}^{\ell_2-2} x_\ell(k^{-2}+\mu_n k^{-5/2})\end{aligned}$$

if c_0 is chosen sufficiently large. The claim follows. □

6.3 Why does the right boundary condition disappear?

The goal of this section is to show that the phase evolution picks up sufficient randomness that will neutralize the right boundary condition of the discrete equations.

Proposition 33 *Let $m = \lfloor n - \mu_n^2/4 - \kappa(\mu_n^{2/3}\vee 1)\rfloor$ and suppose that $\kappa \to \infty$ with $n_0^{-1}\kappa(\mu_n^{2/3}\vee 1) \to 0$. Then $\varphi_{m,0}$ modulo 2π converges in distribution to Uniform$(0, 2\pi)$.*

Proof We will show that given $\varepsilon > 0$, every subsequence of indices has a further subsequence along which $\varphi_{m,0}$ modulo 2π is eventually ε-close to uniform distribution. So we first pick an integer $\tau = \tau(\varepsilon)$ and show that along a suitable subsequence, the conditional distribution given $\mathcal{F}_{m-\tau\xi}$ of $\varphi_{m,0} - \varphi_{m-\tau\xi,0}$ converges to Gaussian with variance tending to ∞ with τ. Here the scaling factor is $\xi = \lfloor \kappa(\mu_n^{2/3}\vee 1)\rfloor$. Since a constant plus a Gaussian with large variance is close to uniform modulo 2π, the claim follows if we let τ go to ∞.

To show the distributional convergence, we apply the SDE limit Theorem 25 to the evolution of φ from time $m - \tau\xi$ on. To adapt to the setup of the theorem we introduce the new parameters

$$\check{n} = n - m + \tau\xi, \qquad \mu_{\check{n}} = \mu_n, \qquad \check{n}_0 = \check{n} - \mu_{\check{n}}^2/4 - 1/2, \qquad \check{\varphi}_{\ell,\lambda} = \varphi_{\ell+m-\tau\xi,\lambda}.$$

By assumption, we have $\check{n}_0\mu_{\check{n}}^{-2/3} \to \infty$. We pass to a subsequence so that $\check{n}_0/\check{n}$ has a limit $1/(1+\nu) \in [0,\infty]$, so Theorem 25 (trivially modified to allow general initial conditions) applies. The result is that $\check{\varphi}_{\lfloor t\check{n}_0\rfloor,0}$ has an SDE limit given by (78) with $\lambda = 0$. Thus $\varphi_{m,0} - \varphi_{m-\tau\xi,0}$ converges to a normal random variable which does not depend on the initial value $\check{\varphi}(0)$. Its variance is given by integrating the sum of the squares of the independent diffusion coefficients on the corresponding scaled time interval:

$$\int_0^{1-(1+\tau)^{-1}} \frac{6\nu+2-2t}{\beta(1-t)(\nu+1-t)}dt \ge \frac{2}{\beta}\log(\tau+1),$$

which goes to ∞ with τ, as required. □

6.4 The uneventful ending

This section is about the last part of the recursion, from

$$m_2 = \lfloor n - \mu_n^2/4 - \kappa(\mu_n^{2/3} \vee 1) \rfloor$$

to n where $\kappa > 0$ is a constant. The goal is to show that nothing interesting happens on this stretch. More precisely, we show

Lemma 34 *For every $\lambda \in \mathbb{R}$ and $\kappa > 0$ as $n \to \infty$ we have $\varphi^{\odot}_{m_2,\lambda} - \varphi^{\odot}_{m_2,0} \to 0$ in probability.*

Fix κ and λ. We will show the convergence by showing that any subsequence has a further subsequence with the desired limit. Because of this, we may assume that the limit of μ_n exists. We will consider two cases: $\lim \mu_n < \infty$ and $\lim \mu_n = \infty$.

Proof of Lemma 34 in the case when $\lim \mu_n$ *is finite* In this case we can assume that $n - m_2$ is eventually equal to some integer ξ. Also, ρ_{m_2} converges to a unit complex number ρ with $\operatorname{Im} \rho > 0$. By (45) we have

$$\varphi^{\odot}_{n-\xi,\lambda} * \mathbf{Q}^{-1}_{n-\xi-1} = 0 * \mathbf{R}^{-1}_{n-1} \cdots \mathbf{R}^{-1}_{n-\xi}, \tag{108}$$

where $\mathbf{R}^{-1}_{n-j,\lambda} = \mathbf{W}^{-1}_{n-j} \tilde{\mathbf{L}}^{-1}_{n-j,\lambda} \mathbf{Q}(\pi)^{-1}$. Consider the components of the product on the right-hand side of (108). The elements $\tilde{\mathbf{L}}_{n-j,\lambda}$ are deterministic (see (51)) and as functions on $\mathbb{R}'$—the lifted unit circle—they converge uniformly to non-degenerate limits that do not depend on λ. (Here we also used $s_{n-j} = \sqrt{j - 1/2}$.) In the same sense, we also have $\mathbf{T}_{n-\xi} \to \mathbf{A}(\operatorname{Im}(\rho)^{-1}, -\operatorname{Re} \rho)$. Because of the moment bounds (40) we may find a subsequence along which $X_{n-1}, \ldots, X_{n-\xi}$ and $Y_{n-1}, \ldots, Y_{n-\xi}$ all converge. Then (using the definition (43)) it follows that the random elements $\mathbf{W}_{n-j}$ converge as functions for $j = 1, \ldots, \xi$.

Since all of these limits are non-degenerate and the dependence on λ disappears, we have

$$\left|\varphi^{\odot}_{n-\xi,\lambda} - \varphi^{\odot}_{n-\xi,0}\right| = \left|\varphi^{\odot}_{n-\xi,\lambda} * \mathbf{Q}^{-1}_{n-\xi-1} - \varphi^{\odot}_{n-\xi,0} * \mathbf{Q}^{-1}_{n-\xi-1}\right| \to 0. \qquad \square$$

Remark 35 For the second case, we review some of the results of Ramírez, Rider, and Virág [15], henceforth denoted RRV, about the eigenvalues of the stochastic Airy operator. The paper considers the eigenvalue process Λ_n of the random matrix M (see (21)) under the edge scaling $n^{1/6}(\Lambda_n - 2\sqrt{n})$. By Theorem 1.1 of RRV, the limit is a point process Ξ given by the eigenvalues of the *stochastic Airy operator*, the random Schrödinger operator

$$\mathcal{H}_\beta = -\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} b'_x$$

on the positive half-line. Here b' is white noise and the initial condition for the eigenfunction f is $f(0) = 0$, $f'(0) = 1$. By RRV, Proposition 3.5 and the discussion preceding RRV, Proposition 3.7,

$$\Xi \text{ is a.s. simple, and for every } x \in \mathbb{R}, \text{ we have } \mathrm{P}(x \in \Xi) = 0. \tag{109}$$

The proof is based on the observation that after appropriate rescaling the matrix M acts on vectors as a discrete approximation of $\mathcal{H}_\beta$. The initial condition $f(0) = 0$, $f'(0) = 1$ comes from the fact that the discrete eigenvalue equation for an eigenvalue $\Lambda = 2\sqrt{n} + n^{-1/6}\nu$ is

equivalent to a three-term recursion for the vector entries $w_{\ell,\nu}$ (c.f. (41) and Remark 20) with the initial condition $w_{0,\nu}=0$ and $w_{1,\nu}\neq 0$.

By RRV, Remark 3.8, the results of RRV extend to solutions of the same three-term recursion with more general initial conditions. We say that a value of ν is an eigenvalue for a family of recursions parameterized by ν if the corresponding recursion reaches 0 in its last step. Suppose that for given $\zeta\in[-\infty,\infty]$ the initial condition for the three-term recursion equation satisfies

$$\frac{w_{0,\nu}}{n^{1/3}(w_{1,\nu}-w_{0,\nu})}=n^{-1/3}(p_n-1)^{-1}\xrightarrow{P}\zeta,$$

where $p_n:=w_{1,\nu}/w_{0,\nu}$ does not depend on ν. Here the factor $n^{1/3}$ is the spatial scaling for the problem (RRV, Sect. 5). Then the eigenvalues of this family of recursions converge to those of the stochastic Airy operator with initial condition $f(0)/f'(0)=\zeta$. The corresponding point process Ξ_ζ will also satisfy (109), see RRV, Remark 3.8.

Now we are ready to complete the proof of Lemma 34.

Proof of Lemma 34 in the case when $\lim\mu_n=\infty$ Without loss of generality we assume $\lambda>0$. Fix a $\theta\in\mathbb{R}'$ and let B denote the event that

$$x_*\mathbf{Q}_{m_2-1}\not\equiv\theta \bmod 2\pi \quad \text{for } x\in[\varphi^{\odot}_{m_2,\lambda},\varphi^{\odot}_{m_2,0}]. \tag{110}$$

It suffices to show that $\mathrm{P}(B)\to 1$. Indeed, by considering a subdivision of the unit circle into arcs of length at most ε at points $e^{i\theta_j}$, if the event (110) holds for each θ_j then

$$\left|\varphi^{\odot}_{m_2,\lambda}{}_*\mathbf{Q}^{-1}_{m_2-1}-\varphi^{\odot}_{m_2,0}{}_*\mathbf{Q}^{-1}_{m_2-1}\right|=\left|\varphi^{\odot}_{m_2,\lambda}-\varphi^{\odot}_{m_2,0}\right|$$

cannot be greater than ε. Taking $\varepsilon\to 0$ completes the proof.

Equation (54) translates B to an event about $\hat\varphi^{\odot}_{\ell,\lambda}$. More specifically, by Proposition 17 B is the event that the one-parameter family of recursions parameterized by ν

$$\hat\varphi_{\ell+1,\nu}=\hat\varphi_{\ell,\nu}{}_*\mathbf{R}_{\ell,\nu},\qquad \ell\geq m_2$$

with initial condition

$$\hat\varphi_{m_2,\nu}=\theta_*\mathbf{T}^{-1}_{m_2} \tag{111}$$

does not have an eigenvalue in the interval $[0,\lambda]$. This recursion is determined by the bottom right $n_2\times n_2$ submatrix of $M(n)^{D(n)}$ (39), where $n_2=n-m_2$. Thus the recursion is in fact the discrete eigenvalue equation for $M(n_2)^{D(n_2)}$ with a generalized initial condition. This can be transformed back to the discrete eigenvalue equation for $M(n_2)$ with the corresponding initial condition. Let $u=\mathbf{U}^{-1}(e^{i\theta})\in\mathbb{R}$ be the point corresponding to $\theta\in\mathbb{R}'$. Then (111) translates to the initial condition

$$r_{m_2,\nu}=u.\mathbf{T}^{-1}_{m_2}=\mathrm{Im}(\rho_{m_2})u+\mathrm{Re}(\rho_{m_2}),$$

for the eigenvalue equation of M^D (see (42)) and by Remark 20 to the initial condition

$$p_{m_2,\nu}=r_{m_2,\nu}\frac{\chi_{(n-m_2-1)\beta}}{\sqrt{\beta}s_{m_2+1}}=\frac{\chi_{(n-m_2-1)\beta}}{\sqrt{\beta(n-m_2-1/2)}}(\mathrm{Im}(\rho_{m_2})u+\mathrm{Re}(\rho_{m_2})) \tag{112}$$

for the eigenvalue equation of M. As $\mu_n \to \infty$, we have

$$n_2 \to \infty, \qquad \mu_n^2/4 = n_2 - \kappa n_2^{1/3} + o(n_2^{1/3}), \quad \text{and} \quad \rho_{m_2} = 1 + i\sqrt{\kappa} n_2^{-1/3} + o(n_2^{-1/3}). \tag{113}$$

Since $\ell^{-1/2}\chi_\ell$ converges to 1 in probability as $\ell \to \infty$, (112) and (113) imply

$$n_2^{-1/3}\left(p_{m_2,\nu} - 1\right)^{-1} \xrightarrow{P} \kappa^{-1/2} u^{-1} =: \zeta.$$

This means that the limit of $\mathrm{P}(B)$ can be related to the limit point process Ξ_ζ. The interval $[0, \lambda]$ corresponds to $2\mu_n + [0, \lambda n_0^{-1/2}/2]$ in our scaling (50). In the edge scaling corresponding to n_2, the length of the remaining stretch, this turns into the interval

$$n_2^{1/6}(2\mu_n - 2n_2^{1/2}) + [0, \lambda n_0^{-1/2}/2] \to -\kappa + [0, 0],$$

where the convergence follows from (113).

For $\delta > 0$ let B_δ be the event that the discrete eigenvalue equation for $M(n_2)$ with initial condition (112) does not have an eigenvalue in the interval

$$2n_2^{1/2} + n_2^{-1/6}(-\kappa - \delta, -\kappa + \delta).$$

By Remark 35, for any fixed δ we have

$$\limsup_{n\to\infty} \mathrm{P}(B) \le \limsup_{n\to\infty} \mathrm{P}(B_\delta) \le \mathrm{P}(\Xi_\zeta \text{ doesn't have a point in } [\kappa - \delta, \kappa + \delta]).$$

Since this holds for all $\delta > 0$, the fact (109) gives $\lim \mathrm{P}(B) = 1$, as required. □

Acknowledgements This research is supported by the Sloan and Connaught grants, the NSERC discovery grant program, and the Canada Research Chair program (Virág). Valkó is partially supported by the Hungarian Scientific Research Fund grant K60708. We thank Yuval Peres for comments simplifying the proof of Proposition 23, and also Mu Cai, Laure Dumaz, Peter Forrester and Brian Sutton for helpful comments. We are indebted to the anonymous referees for their extensive comments and suggestions.

Appendix A: Tools

A.1 Angular shift bounds

The objective of this section is to prove Lemma 16, which relies on Fact 15.

Proof of Fact 15 The general form of such a transformation is given by $w_\circ T = e^{i\alpha}(w - \sigma)/(1 - \bar{\sigma} w)$, where σ is the pre-image of 0. We may assume $\alpha = 0$ since post-composing T with a rotation does not change the quantities in question. Using the definition of $\mathrm{ash}(\mathbf{T}, v, w)$ and the fact that $|w| = |v| = 1$ we have

$$\mathrm{ash}(\mathbf{T}, v, w) = \mathrm{Arg}_{[0,2\pi)}\left(\frac{w - \sigma}{v - \sigma}\,\frac{\bar{v} - \bar{\sigma}}{\bar{w} - \bar{\sigma}}\,\frac{v}{w}\right) - \mathrm{Arg}_{[0,2\pi)}(w/v).$$

The additivity of Arg mod 2π proves (34) mod 2π. By definition, ash is continuous in T and so also in σ. Since $|\sigma| < 1$, the right-hand side of (34) is continuous in σ. As equality holds for $\sigma = 0$, the proof is complete. □

Proof of Lemma 16 Recall that $r.\mathbf{U} = (i-r)/(i+r)$ maps the upper half plane to the unit disk, sending i to 0. By Fact 15 we have

$$\operatorname{ash}(\mathbf{T}, v, w) = 2\operatorname{Arg}\left(\frac{1-((i+z).\mathbf{U})\,\bar{w}}{1-((i+z).\mathbf{U})\,\bar{v}}\right) = 2\operatorname{Arg}(1+x), \quad x = \frac{z(\bar{w}-\bar{v})}{2i+z(1+\bar{v})}.$$

If $|z| \le 1/3$ then we have $|x| \le 1/2$ so we can write $\operatorname{ash}(\mathbf{T}, v, w) = \operatorname{Re} h_{v,w}(z)$ with

$$h_{v,w}(z) = \frac{2}{i}\log\left(1+\frac{z(\bar{w}-\bar{v})}{2i+z(1+\bar{v})}\right) = (\bar{w}-\bar{v})\left(-z-\frac{i(2+\bar{v}+\bar{w})}{4}z^2+\eta_{v,w}(z)\right).$$

Here we use the standard branch of the logarithm defined outside the negative real axis. The second equality is Taylor expansion in z. To bound the error term, we write

$$h'''_{v,w}(z) = \frac{(\bar{w}-\bar{v})\,p(z,\bar{v},\bar{w})}{(2i+z(1+\bar{v}))^3(2i+z(1+\bar{w}))^3},$$

where p is some (explicitly computable) polynomial, so the Taylor error term satisfies

$$|\eta_{v,w}(z)| \le \frac{|z^3|}{3!}\sup_{|z|\le 1/3,\,|v|=|w|=1}\frac{|h'''_{v,w}(z)|}{|w-v|} < c|z|^3.$$

This proves the quadratic approximation of the angular shift for $|z| \le 1/3$, and the other two estimates of (37) follow easily.

For the case $|z| > 1/3$, $v = -1$ we use the fact that $|\operatorname{Arg}(1+x)| \le \pi|x|$ to conclude that

$$|\operatorname{ash}(\mathbf{T}, v, w)| \le 4\pi\left|\frac{z(\bar{w}-\bar{v})}{i+z(1+\bar{v})/2}\right| = 4\pi\,|z(\bar{w}-\bar{v})| \le 4\pi 3^{d-1}\left|z^d(\bar{w}-\bar{v})\right|$$

for any $d \ge 1$. Using $|z| > 1/3$ we get that the main terms on the right-hand side of (37) may also be bounded by $c_d|w-v||z|^d$ and from this we get upper bounds of (38) as well. □

A.2 Oscillatory sums

Recall from the definition (53) that η_ℓ is a unit complex number with a rapidly oscillating angle. Lemma 37 below will show that this oscillation has an averaging effect in sums. In order to prove that we first need the following harmonic analysis lemma.

Lemma 36 *Suppose that $2\pi > \theta_1 > \theta_2 > \cdots > \theta_m > 0$ and let $s_\ell = \sum_{j=1}^{\ell}\theta_j$. Then*

$$\max_{1\le\ell\le m}\left|\sum_{j=1}^{\ell} e^{is_j}\right| \le c(\theta_m^{-1} + (2\pi-\theta_1)^{-1}).$$

Proof We first consider the case when $\theta_1 \le \pi$. Using second order interpolation we can construct a differentiable function $s(x)$ on $[1, m]$ with $s(\ell) = s_\ell$ for $1 \le \ell \le m$ for which the derivative $s'(x)$ is monotone decreasing derivative with $-\pi \le s'(x) \le -\theta_1/2$.

Our proof is based on the following lemmas of van der Corput (see [9] for the first and [16], Chap. VIII, Proposition 2 for the second):

(i) If $s(x)$ has a monotone derivative with $|s'(x)| \le \pi$ for $x \in [a, b]$ (with $a, b \in \mathbb{Z}$) then the difference of $\sum_{\ell=a}^{b} e^{is(l)}$ and $\int_a^b e^{is(x)}dx$ is at most 3.

(ii) If $s'(x)$ is monotone and $|s'(x)| > p$ on an interval $[a, b]$ then $|\int_a^b e^{is(x)}dx| \le 3p^{-1}$.

Since for our function $\pi > |s'(x)| > \theta_m/2$ for $x \in [1, m]$ we may apply these lemmas to get the bound $|\sum_{j=1}^{\ell} e^{is_j}| \le c\theta_m^{-1}$.

Consider now the case $2\pi > \theta_1 > \pi$. Let ℓ^* be the largest index with $\theta_{\ell^*} > \pi$, then

$$\left|\sum_{j=1}^{\ell} \exp\left[i\sum_{u=1}^{j}\theta_u\right]\right| \le \left|\sum_{j=1}^{\ell\wedge\ell^*} \exp\left[i\sum_{u=1}^{j}\theta_u\right]\right| + \left|\sum_{j=\ell^*+1}^{\ell} \exp\left[i\sum_{u=\ell^*+1}^{j}\theta_u\right]\right|. \tag{114}$$

The second sum can be bounded by $c\theta_m^{-1}$ using the first half of our proof. To bound the first sum, note that

$$\left|\sum_{j=1}^{\ell} \exp\left[i\sum_{u=1}^{j}\theta_u\right]\right| = \left|\sum_{j=1}^{\ell} \exp\left[i\sum_{u=1}^{j}\tilde{\theta}_u\right]\right|, \quad \tilde{\theta}_u = 2\pi - \theta_{\ell+1-u}$$

and for $\ell \le \ell^*$ we have

$$\pi > 2\pi - \theta_{\ell^*} \ge \tilde{\theta}_1 > \tilde{\theta}_2 > \cdots > \tilde{\theta}_\ell \ge 2\pi - \theta_1 > 0.$$

Thus the first half of the proof can be applied again to get the bound $c(2\pi - \theta_1)^{-1}$. □

The following lemma describes the averaging effects of the oscillating unit complex numbers η_ℓ.

Lemma 37 *Let $g_\ell \in \mathbb{C}$ for $\ell \in \mathbb{N}$ and $\ell_0 < \ell_1 \le n_0$. Then*

$$\left|\mathrm{Re}\sum_{\ell=\ell_0}^{\ell_1} g_\ell\eta_\ell\right| \le c\big(\mu_n k_1^{-1/2} + 1\big)|g_{\ell_1}| + c\sum_{\ell=\ell_0}^{\ell_1-1}\big(\mu_n k^{-1/2} + 1\big)|g_{\ell+1} - g_\ell|,$$

$$\left|\mathrm{Re}\sum_{\ell=\ell_0}^{\ell_1} g_\ell\eta_j^2\right| \le c\big(\mu_n k_1^{-1/2} + \mu_n^{-1}k_0^{1/2}\big)|g_{\ell_1}| + c\sum_{\ell=\ell_0}^{\ell_1-1}\big(\mu_n k^{-1/2} + \mu_n^{-1}k_0^{1/2}\big)|g_{\ell+1} - g_\ell|.$$

(*We used the shorthanded notation $k = n_0 - \ell$, $k_1 = n_0 - \ell_1$ and $k_0 = n_0 - \ell_0$.*)

Proof For $d = 1, 2$ we introduce $F_{d,j} = \sum_{m=\ell_0}^{j} \eta_m^d$ with $F_{d,\ell_0-1} = 0$. By partial summation

$$\sum_{j=\ell_0}^{\ell_1} g_j\eta_j = F_{d,\ell_1}g_{\ell_1} + \sum_{j=\ell_0}^{\ell_1-1} F_{d,j}(g_j - g_{j+1}). \tag{115}$$

From (49) we get the estimates

$$\mathrm{Arg}\,\rho_\ell \le \mu_n^{-1}k^{1/2}, \quad \text{and} \quad \pi/2 - \mathrm{Arg}\,\rho_\ell \le \mu_n k^{-1/2}.$$

Together with (53) this means that we can use Lemma 36 to get

$$|F_{1,\ell}| \le c\big(\mu_n k_1^{-1/2} + 1\big), \qquad |F_{2,\ell}| \le c\big(\mu_n k_1^{-1/2} + \mu_n^{-1}k_0^{1/2}\big).$$

This with (115) implies the lemma. □

A.3 A limit theorem for random difference equations

Proof of Proposition 23 Let $\|\cdot\|_\infty$ denote supremum norm on $[0,T]$. For a two-parameter function f and $x\in\mathbb{R}$ let $\mathcal{I}$ denote the integral $\mathcal{I}_{f,x}(t)=\int_0^t f(s,x)ds$. We recycle this notation for a function $X:[0,T]\to\mathbb{R}$ to write $\mathcal{I}_{f,X}(t)=\int_0^t f(s,X(s))ds$.

The proof of this proposition is based on Theorem 7.4.1 of Ethier and Kurtz [7], as well as Corollary 7.4.2 and its proof. (See also [17].) It states that if the limiting SDE has unique distribution (i.e. the martingale problem is well-posed) and also

$$\begin{aligned}\|\mathcal{I}_{b^n,X^n}-\mathcal{I}_{b,X^n}\|_\infty &\xrightarrow{P} 0,\\ \|\mathcal{I}_{a^n,X^n}-\mathcal{I}_{a,X^n}\|_\infty &\xrightarrow{P} 0,\end{aligned}\tag{116}$$

$$\text{for every } \varepsilon>0 \quad \sup_{x,\ell}\mathrm{P}(|Y^n_\ell(x)|\ge\varepsilon)\longrightarrow 0,\tag{117}$$

then $X^n\overset{d}{\Longrightarrow}X$. The theorem there only deals with the case of time-independent coefficients, but adding time as an extra coordinate extends the results to the general case.

Because of our assumptions on a and b the well-posedness of the martingale problem follows from Theorem 5.3.7 of Ethier and Kurtz [7] (see especially the remarks following the proof), and even pathwise uniqueness holds. Condition (117) follows from the uniform third absolute moment bounds (74) and Markov's inequality. Thus we only need to show (116) as well as the analogous statement for a, for which the proof is identical. We do this by bounding the successive uniform-norm distances between

$$\mathcal{I}_{b^n,X^n},\quad \mathcal{I}_{b^n,X^{n,L}},\quad \mathcal{I}_{b,X^{n,L}},\quad \mathcal{I}_{b,X^n},$$

where $X^{n,L}_\ell=X^n_{K\lfloor \ell/K\rfloor}$ with $K=\lceil nT/L\rceil$, and $X^{n,L}(t)=X^{n,L}_{\lfloor nt\rfloor}$. In words, we divide $[0,\lfloor nT\rfloor]$ into L roughly equal intervals and then set $X^{n,L}_\ell$ to be constant on each interval and equal to the first value of X^n_ℓ occurring there.

If a function f takes countably many values f_i, then for any h we have

$$\|\mathcal{I}_{h,f}\|_\infty\le\sum_i\|\mathcal{I}_{h,f_i}\|_\infty.$$

Since $X^{n,L}$ takes at most L values, we have

$$\|\mathcal{I}_{b^n,X^{n,L}}-\mathcal{I}_{b,X^{n,L}}\|_\infty=\|\mathcal{I}_{b^n-b,X^{n,L}}\|_\infty\le L\sup_x\|\mathcal{I}_{b^n-b,x}\|_\infty=Lo(1)$$

by (75) where $o(1)$ is uniform in L and refers to $n\to\infty$. From (73), the other terms satisfy

$$\begin{aligned}\|\mathcal{I}_{b^n,X^{n,L}}-\mathcal{I}_{b^n,X^n}\|_\infty &\le T\|b^n(\cdot,X^{n,L}(\cdot))-b^n(\cdot,X^n(\cdot))\|_\infty\\ &\le cT\|X^n-X^{n,L}\|_\infty+o(1).\end{aligned}$$

The same holds with b replacing b^n. It now suffices to show that

$$\mathrm{E}\|X^{n,L}-X^n\|_\infty=\mathrm{E}\sup_\ell|X^{n,L}_\ell-X^n_\ell|\le f(L)\tag{118}$$

uniformly in n where $f(L) \to 0$ as $L \to \infty$. The left-hand side of (118) is bounded by

$$\mathrm{E}\sup_{\ell}\left|X_\ell^n - \frac{1}{n}\sum_{k=\lfloor \ell/K\rfloor K}^{\ell-1} b_n(X_\ell^n) - X_\ell^{n,L}\right| + \mathrm{E}\sup_{\ell}\left|\frac{1}{n}\sum_{k=\lfloor \ell/K\rfloor K}^{\ell} b(X_\ell)\right|$$

and the second quantity is bounded by $T\sup_{\ell,x}|b_\ell^n(x)|/L$. The first quantity can be written as $\mathrm{E}M^*$ where

$$M^* = \max_{i=0,\dots,L-1} M_i^*, \qquad M_i^* = \max_{\ell=0,\dots,K-1} |M_{i,\ell}|,$$

$$M_{i,\ell} = X_{iK+\ell} - X_{iK} - \frac{1}{n}\sum_{k=0}^{\ell-1} b^n(X_{iK+k}).$$

Note that for each i, $M_{i,\ell}$ is a martingale. For any martingale with $M_0 = 0$ we have

$$\mathrm{E}\max_{k\le n}|M_k|^3 \le c\mathrm{E}\left|\sum_{k\le n}\mathrm{E}[(M_k - M_{k-1})^2|\mathcal{F}_{k-1}]\right|^{3/2} \le cn^{3/2}\max_{k\le n}\mathrm{E}[|M_k - M_{k-1}|^3|\mathcal{F}_{k-1}].$$

The first step is the Burkholder-Davis-Gundy inequality (see [11], Theorem 26.12) and the second step follows from Jensen's inequality. Therefore (74) implies

$$\mathrm{E}[|M_i^*|^3|\mathcal{F}_{iL}] \le c(n/L)^{3/2}n^{-3/2} = cL^{-3/2},$$

which gives the desired conclusion

$$(\mathrm{E}M^*)^3 \le \mathrm{E}(M^*)^3 \le \mathrm{E}\sum_{i=0}^{L-1}(M_i^*)^3 \le cL^{-1/2}.$$

Letting first $n \to \infty$ and then $L \to \infty$ gives (118) and (116). □

References

1. Deift, P.A.: Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach. Courant Lecture Notes in Mathematics. AMS, New York (1999)
2. Deift, P., Its, A., Zhou, X.: A Riemann-Hilbert approach to asymptotic problems arising in the theory of random matrices and also in the theory of integrable statistical mechanics. Ann. Math. **146**, 149–235 (1997)
3. Deift, P., Kriecherbauer, T., McLaughlin, K.T.-R., Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. Commun. Pure Appl. Math. **52**(12), 1491–1552 (1999)
4. Dumitriu, I., Edelman, A.: Matrix models for beta ensembles. J. Math. Phys. **43**(11), 5830–5847 (2002)
5. Dyson, F.: Statistical theory of energy levels of complex systems II. J. Math. Phys. **3**, 157–165 (1962)
6. Edelman, A., Sutton, B.D.: From random matrices to stochastic operators. J. Stat. Phys. **127**(6), 1121–1165 (2007)
7. Ethier, S.N., Kurtz, T.G.: Markov Processes. Wiley, New York (1986)
8. Forrester, P.: Log-Gases and Random Matrices (2008) (book in preparation). www.ms.unimelb.edu.au/~matpjf/matpjf.html
9. Hille, E.: Note on a power series considered by Hardy and Littlewood. J. Lond. Math. Soc. **4**(15), 176–183 (1929). doi:10.1112/jlms/s1-4.15.176
10. Jimbo, M., Miwa, T., Môri, Y., Sato, M.: Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent. Physica D **1**, 80–158 (1980)

11. Kallenberg, O.: Foundations of Modern Probability. Springer, New York (2002)
12. Killip, R.: Gaussian fluctuations for β ensembles. Int. Math. Res. Not. (2008). Art. ID mn007, 19 pp.
13. Killip, R., Stoiciu, M.: Eigenvalue statistics for CMV matrices: From Poisson to clock via random matrix ensembles. Duke Math. J. **146**(3), 361–399 (2009)
14. Mehta, M.L.: Random Matrices, 3rd edn. Elsevier/Academic, Amsterdam (2004)
15. Ramírez, J., Rider, B., Virág, B.: Beta ensembles, stochastic Airy spectrum, and a diffusion (2007). arXiv:math/0607331
16. Stein, E.M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton (1993)
17. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Classics in Mathematics. Springer, Berlin (1979)
18. Sutton, B.D.: The stochastic operator approach to random matrix theory. Ph.D. thesis, MIT, Department of Mathematics (2005)
19. Trotter, H.F.: Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegő. Adv. Math. **54**(1), 67–82 (1984)
20. Virág, B.: Scaling limits of random matrices. Plenary lecture, 31st Conference on Stochastic Processes and their Applications, Paris, July 17–21, 2006
21. Widom, H.: The asymptotics of a continuous analogue of orthogonal polynomials. J. Approx. Theory **77**, 51–64 (1996)