EIGENVALUE STATISTICS FOR CMV MATRICES: FROM POISSON TO CLOCK VIA RANDOM MATRIX ENSEMBLES.

ROWAN KILLIP AND MIHAI STOICIU

ABSTRACT. We study CMV matrices (discrete one-dimensional Dirac-type operators) with random decaying coefficients. Under mild assumptions we identify the local eigenvalue statistics in the natural scaling limit. For rapidly decreasing coefficients, the eigenvalues have rigid spacing (like the numerals on a clock); in the case of slow decrease, the eigenvalues are distributed according to a Poisson process. For a certain critical rate of decay we obtain the beta ensembles of random matrix theory. The temperature β^{-1} appears as the square of the coupling constant.

1. INTRODUCTION

The bulk spectral properties of Schrödinger operators and Jacobi matrices with random decaying potentials are fairly well understood; see [10, 11, 29, 30, 40]. In particular, it has been discovered that there is a change in spectral type at certain critical rate of decay. It is natural to investigate the finer spectral properties; for example, the distribution of eigenvalue spacings. In order to speak of eigenvalue spacings, one first restricts the operator to a finite volume and then studies the behaviour of the rescaled spacings in the infinite volume limit. Such an investigation is the goal of this paper; however we treat neither Schrödinger nor Jacobi operators, but rather CMV matrices:

Definition 1.1. Given an infinite sequence of coefficients $\alpha_0, \alpha_1, \ldots$ in $\mathbb{D}$ (the open unit disk in $\mathbb{C}$), we define 2×2 matrices

$$\Xi_k = \begin{bmatrix} \bar{\alpha}_k & \rho_k \\ \rho_k & -\alpha_k \end{bmatrix},$$

where $\rho_k = \sqrt{1 - |\alpha_k|^2}$. From these, we form block-diagonal matrices

$$\mathcal{L} = \operatorname{diag}\big(\Xi_0, \Xi_2, \Xi_4, \ldots\big) \quad \text{and} \quad \mathcal{M} = \operatorname{diag}\big(\Xi_{-1}, \Xi_1, \Xi_3, \ldots\big),$$

where $\Xi_{-1} = [1]$, the 1×1 identity matrix. The *CMV matrix* associated to the coefficients $\alpha_0, \alpha_1, \ldots$ is $\mathcal{C} = \mathcal{C}(\alpha_0, \alpha_1, \ldots) = \mathcal{L}\mathcal{M}$.

The name CMV matrices is taken from from the initials of the authors of [6], following recent custom; however as discussed in [44, 52], they appeared earlier in several places. A thorough discussion of these operators can be found in [41, 42]. As in these books, will refer to the α_k as Verblunsky coefficients.

CMV matrices are unitary and may naturally be considered a discrete form of one-dimensional Dirac operator. There are several reasons for our choosing to

consider these operators: they are bounded, discrete, and most importantly, possess a rotational symmetry that we will discuss in due course.

The restriction of a CMV matrix to a finite interval is not simply the top left corner of the original matrix—this would result in a non-unitary operator—one needs to add a boundary condition. If one replaces α_{n-1} by $e^{i\eta} \in S^1$ (the unit circle in $\mathbb{C}$), then $\mathcal{C}$ decomposes into a direct sum; specifically, the top left $n \times n$ block decouples from the remainder of the matrix. We define the $n \times n$ truncation of $\mathcal{C}$ with η boundary condition in this way and write

$$\mathcal{C}^{(n)} = \mathcal{C}(\alpha_0, \ldots, \alpha_{n-2}, e^{i\eta}).$$

We wish to consider the model where $e^{i\eta}$ and each α_k is chosen independently (though not identically distributed) with a rotationally invariant distribution, that is,

$$\mathbb{E}\Big\{g(e^{i\eta}) \prod f_k(e^{i\theta_k}\alpha_k)\Big\} = \int_0^{2\pi} g(e^{i\eta}) \tfrac{d\eta}{2\pi} \ \prod \mathbb{E}\{f_k(\alpha_k)\}$$

for any finite collection of $f_k : \mathbb{D} \to \mathbb{C}$ and $\theta_k \in \mathbb{R}$. The reason for considering only rotationally invariant distributions is that in this case the spectral properties are homogeneous under rotations of the circle. For example, for any $\theta \in [0, 2\pi)$,

$$e^{-i\theta}\mathcal{C}(\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, e^{i\eta}) \cong \mathcal{C}(e^{i\theta}\alpha_0, e^{2i\theta}\alpha_1, \ldots, e^{i(n-1)\theta}\alpha_{n-2}, e^{i\eta+in\theta}), \quad (1)$$

which is to say that the two matrices have the same eigenvalues that is, they are unitarily equivalent (cf. [41, p. 81]).

For power-decaying α_k, the bulk spectral properties for this model were worked out in [42, §12.7], following the lines of earlier arguments from the Jacobi and Schrödinger cases. Let us now describe the finer behaviour we wish to investigate.

Associated to the random $n \times n$ matrix $\mathcal{C}^{(n)}$ are n random eigenvalues, which we denote $e^{i\theta_1}, \ldots, e^{i\theta_n}$, with $\theta_j \in (-\pi, \pi]$. By rotation invariance, (1), the average number of eigenvalues with $\theta_j \in [a, b)$ is $\frac{b-a}{2\pi}n$. For this reason it is natural to rescale the eigenvalues by a factor of n. Our goal is to show that under natural assumptions, the *eigenvalue process*, that is, the random set of points $\{n\theta_1, \ldots, n\theta_n\} \subset \mathbb{R}$, converge (in an appropriate sense) to a point process and to understand the limiting process as best we can.

We will prove the convergence of the expected values of certain random variables that are rich enough to characterize the point process:

Definition 1.2. Given a non-negative $f \in C_c^\infty(\mathbb{R})$, we define a random variable on a point process by

$$X_f = \exp\Big(-\sum f(x_j)\Big)$$

where x_j are the points of the process. $\mathbb{E}\{X_f\}$ are known as the Laplace functionals of the process.

One may also characterize a point process in terms of the probability of certain events:

Definition 1.3. Given p disjoint intervals $[a_j, b_j)$, and p non-negative integers $k_1, \ldots, k_p$, we define the event

$$\Omega\big([a_1, b_1), \ldots, [a_p, b_p), k_1, \ldots, k_p\big)$$

as the occurrence of exactly k_j points in the interval $[a_j, b_j)$ for each $1 \le j \le p$.

Convergence of the probabilities of these events is fractionally stronger than that of the random variables X_f; indeed they are equivalent notions if no location is occupied by the limit process with positive probability (cf. [37, §3.5]). We will use the former test for convergence as it is the most convenient for us.

Three processes will arise as limits of the eigenvalue process; we will now define these.

Definition 1.4. The Poisson process with intensity $d\nu$ is characterized by

$$\mathbb{E}\{X_f\} = \exp\left\{\int \left[e^{-f(x)} - 1\right] d\nu(x)\right\}$$

or more simply by

$$\mathbb{P}\left\{\Omega\big([a_1, b_1), \ldots, [a_p, b_p), k_1, \ldots, k_p\big)\right\} = \prod_{j=1}^{p} \frac{\lambda_j^{k_j} e^{-\lambda_j}}{k_j!}$$

where $\lambda_j = \nu([a_j, b_j))$. (This says that the number of points in disjoint intervals are independent Poisson random variables.)

Definition 1.5. The clock (or picket fence) process with spacing h is characterized by

$$\mathbb{E}\{X_f\} = \frac{1}{h} \int_0^h \exp\left\{-\sum_{k \in \mathbb{Z}} f(hk + x)\right\} dx.$$

That is, points occur a random translate of $h\mathbb{Z}$.

The third limiting process we need to describe is considerably more complex. We begin with the $C\beta E_n$, that is, the circular β ensemble with n points. This is a point process on the unit circle. The total number of points is n (non-random) and they are distributed so that

$$\mathbb{E}_n^\beta(f) = \frac{1}{Z_{n,\beta}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(e^{i\theta_1}, \ldots, e^{i\theta_n}) \big|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})\big|^\beta \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi} \qquad (2)$$

for any symmetric function f. The partition function is given by

$$Z_{n,\beta} = \frac{\Gamma(\frac{1}{2}\beta n + 1)}{\left[\Gamma(\frac{1}{2}\beta + 1)\right]^n}; \qquad (3)$$

see [23, 53]. This is a standard family of ensembles in random matrix theory and was introduced by Dyson, [16]. This ensemble also arises as the Gibbs measure for n identical charged particles confined to lie on the circle and interacting via the two-dimensional Coulomb law. For this reason, it is also known as the log-gas.

Just as we discussed earlier for eigenvalues of random CMV matrices, one may take the points of the $C\beta E_n$ and by unwrapping and rescaling obtain a sequence of point processes on the line. It is not obvious that this sequence of point processes converges. One off-shoot of the work presented here is that they do (indeed, we will obtain a new representation of the limiting process). The resulting process may be termed the 'microscopic thermodynamic limit of the one-dimensional log-gas'. The word 'microscopic' here indicates that we are considering the statistical properties of individual particles, rather than bulk, collective, or macroscopic properties; the distinction is comparable to that between inter-particle spacing and density. A more wieldy name would be 'the (limiting) beta ensemble of random matrix theory'. For the remainder of this paper, we will settle for 'β-ensemble':

Definition 1.6. The β-*ensemble* is the point process is characterized by

$$\mathbb{E}\{X_f\} = \lim_{n \to \infty} \mathbb{E}_n^\beta \Big\{ \exp\Big\{ -\sum_j f(n\theta_j) \Big\} \Big\}$$

with the notation of (2). (This limit does indeed exist.)

Convergence of the correlation functions for $\beta \in \mathbb{Q}$ is proved in [35]. (The case $\beta \in \mathbb{Z}$ was treated earlier by Forrester.) To the best of our knowledge, convergence for irrational β was previously open.

It is not difficult to show that $\beta \downarrow 0$ limit of CβE$_n$ converges to a Poisson process as $n \to \infty$ and that the $\beta \uparrow \infty$ limit converges a clock process.

For $\beta \in \{1, 2, 4\}$, the β-ensemble is very well understood; indeed there are explicit formulae for the correlation functions and for the inter-particle spacing distributions. Discussion of this work can scarcely be contained in a single monograph and we have no intention of trying to review it here; the unacquainted should probably begin with [8], [19], or [32].

The content of this paper can be best summarized as

Theorem 1.7. *(i) If* $\mathbb{E}\left\{|\alpha_k|^2\right\} = o(k^{-1})$, *then the eigenvalue process converges to the clock process with spacing* 2π *as* $n \to \infty$.
(ii) If $\mathbb{E}\left\{|\alpha_k|^2\right\} - \frac{2}{\beta(k+1)} = o(k^{-1})$ *and* $\mathbb{E}\left\{\log^2[1 - |\alpha_k|^2]\right\} = o(k^{-1})$, *then the eigenvalue process converges to the* β-ensemble *as* $n \to \infty$.
(iii) If $\mathbb{E}\left\{|\alpha_k|^2\right\} \geq (k+1)^{\epsilon-1}$ *for some* $\epsilon > 0$ *and* $\mathbb{E}\left\{(1 - |\alpha_k|^2)^{-s}\right\} = O(1)$ *for some* $s > 0$, *then the eigenvalue process converges to the Poisson process with intensity* $\frac{1}{2\pi}dx$ *as* $n \to \infty$.

This theorem demonstrates that (at least for this model) the transition from Poisson to Clock statistics passes through a continuum of intermediate processes. This moots the question of the eigenvalue statistics in a (microscopic) neighbourhood of the (conjectured) metal/insulator transition in the multi-dimensional Anderson model. The myriad ways in which one approach this transition point (varying energy and/or coupling constant) potentially allows for a diversity of behaviours. One may also ponder the eigenvalue statistics for less random models, for example, the Almost-Mathieu operator, which is also known to have such a transition. Unfortunately, we have nothing to say on these interesting topics.

As $\log[1 - r^2] = r^2 + O(r^4)$ as $r \downarrow 0$, it is natural to consider replacing the second hypothesis in part (ii) by $\mathbb{E}\left\{|\alpha_k|^4\right\} = o(k^{-1})$ or perhaps discarding it altogether. The first option is conceivable, but would increase the complexity of the proof more than it warrants. This second option is certainly not possible. An α_k lying very close to the unit circle effectively of decouples the CMV matrix into a direct sum. For example, if one considers

$$|\alpha_k| = 1 - e^{-k!} \quad \text{with probability} \sim \tfrac{2}{\beta k},$$

and zero otherwise, then the eigenvalue process becomes a superposition of clock processes with random spacings. This has too little eigenvalue repulsion to be consistent with the β-ensemble, at least for β large. The second hypothesis in part (iii) is surely an artifact of our proof.

The proof of part (i) is the most elementary of the three and is contained in Section 3 using results from Section 2. A discussion of non-random results of this type (with stronger decay assumptions) can be found in [43, §4]; their proofs occupy the subsequent sections of that paper.

Part (iii) of Theorem 1.7 is proved in Section 5 by adapting methods from the Ph.D. thesis of the second author, [47, 48]. This, in turn, is inspired by work of Mochanov, [34], and Minami, [33], on the Anderson model.

Part (ii) is proved by combining two ingredients. The first is an exactly soluble example of our model that was discovered in [27]. The second is an invariance principle that shows that the limiting eigenvalue statistics do not depend on the specifics of the probability distribution of the Verblunsky coefficients.

Several results of this genre have been proved, though only for systems which lead to $\beta \in \{1, 2, 4\}$. Such results are reviewed in [9]. To our eyes, the most similar results of this type can be found in [24, 46], which study the Wigner ensemble of Hermitian/Symmetric matrices with independent entries. One may view Theorem 1.7 as an expression of the 'universality' of β-ensembles; though this borders on the vacuous. It is perhaps more startling if viewed as a lack of universality; explaining this point will require a brief detour.

Consider the Dirichlet eigenvalues of the Laplacian in a compact planar domain. It is conjectured that for generic domains whose billiard flow is chaotic, the high energy behaviour of the eigenvalues resembles that of the β-ensemble with $\beta = 1$. The physical intuition behind the appearance of $\beta = 1$ is the fact that this system exhibits time reversal symmetry. The introduction of a magnetic field breaks the time reversal symmetry and is predicted to lead to the β-ensemble with $\beta = 2$, at least for generic domains with chaotic billiard flow. Let us now consider a naive application of these physical principles to our system. For complex Verblunsky coefficients, the CMV matrices fall naturally into the class of systems without time reversal symmetry. This would lead one to believe that the eigenvalue distribution should display a large basin of attraction toward the β-ensemble with $\beta = 2$. However, as Theorem 1.7 shows, no particular value of β is singled out in the case of our model. For a fuller discussion of the matters just described, the reader may begin with [4, 5, 16, 17, 36].

We now describe the exactly soluble model alluded to above.

Definition 1.8. A complex random variable, X, with values in the unit disk, $\mathbb{D}$, is Θ_ν-distributed (for $\nu > 1$) if

$$\mathbb{E}\{f(X)\} = \tfrac{\nu-1}{2\pi} \iint_{\mathbb{D}} f(z)(1 - |z|^2)^{(\nu-3)/2} \, d^2 z \qquad (4)$$

for any continuous $f : \mathbb{D} \to \mathbb{C}$. When $\nu \geq 2$ is an integer, this law has the following geometric interpretation: If v is chosen at random from the unit sphere S^ν in $\mathbb{R}^{\nu+1}$ according to the usual surface measure, then $v_1 + iv_2$ is Θ_ν-distributed.

An elementary computation shows that if $X \sim \Theta_\nu$, then

$$\mathbb{E}\{|X|^2\} = \tfrac{2}{\nu+1}. \qquad (5)$$

The combined content of Theorem 1.2 and Proposition B.2 from [27] is:

Theorem 1.9. *Given $\beta > 0$, let $\alpha_k \sim \Theta_{\beta(k+1)+1}$ be independent random variables. Then the eigenvalues of $\mathcal{C}^{(n)}$ are distributed according to the $C\beta E_n$ ensemble.*

Note that [27] and its precursor, [13], contain similar results for related models on the real line. In these instances, the point process is modelled with a random Jacobi matrix. The bulk spectral properties of one such model are discussed from a random-matrix perspective in [45].

The bulk of the proof of Theorem 1.7(ii) appears in Section 4 although it also relies on material from Section 2.

Next we present our new representation of the β-ensemble. The set of critical values of any non-decreasing function $\psi : \mathbb{R} \to \mathbb{R}$ (i.e., values taken at more than one point) is at most countable; for otherwise one could find uncountably many disjoint open subsets of $\mathbb{R}$. Therefore, for all but countably many $v \in (\inf \psi, \sup \psi)$ we may uniquely define $x = \psi^{-1}(v)$ as the point where ψ passes through (or jumps past) the value v.

Definition 1.10. Given a probability measure on non-decreasing functions $\psi : \mathbb{R} \to \mathbb{R}$, such that $\mathbb{E}\{|\psi(x)|\} < \infty$ for all $x \in \mathbb{R}$, we define an associated point process by

$$\mathbb{E}\{X_f\} = \mathbb{E}\left\{ \int_0^{2\pi} \exp\left[-\sum_{m \in \mathbb{Z}} f \circ \psi^{-1}(2\pi m + \omega) \right] d\omega \right\}, \qquad (6)$$

that is, points are placed at the inverse images of a random translate of $2\pi \mathbb{Z}$.

Remark. The assumption $\mathbb{E}\{|\psi(x)|\} < \infty$, for all $x \in \mathbb{R}$, is a simple (but unnecessarily strong) condition that guarantees that the number of points in any compact set is almost surely finite, as we require for a point process. Indeed, the average number of points in the interval $[a, b)$ is $\mathbb{E}\{\psi(b-) - \psi(a-)\}$.

For example, the picket fence process is associated to $\psi(x) = x$ (non-random), while a Poisson point process arises when ψ is chosen to be (2π times) the Poisson jump process.

The random non-decreasing function that is relevant for us is the relative Prüfer phase. This is described precisely in Section 2; however, we would like to give some of the flavour of this here.

The notion of Prüfer phase originates in the study of Sturm–Liouville problems, see for example, [7, Ch. 8]. Given $q \in L^1$, let us define $u(x; E)$ as the solution of

$$-u''(x) + (q(x) - E)u(x) = 0$$

on $[0, L]$ that obeys $u(0) = 0$ and $u'(0) = 1$. The Prüfer phase $\psi(x; E)$ and amplitude $R(x; E)$ are defined by

$$u(x; E) = R(x; E)\sin(\psi(x; E)), \qquad u'(x; E) = R(x; E)\cos(\psi(x; E))$$

and $\psi(0; E) = 0$. Notice that if we place a self-adjoint boundary condition $u'\sin(\eta) - u\cos(\eta) = 0$ at $x = L$, then the eigenvalues of the Sturm–Liouville problem are the energies where $\psi(L; E) \in \eta + \pi\mathbb{Z}$. Note also that $\psi(L, E)$ is a strictly increasing function of E; this is essentially the Sturm Oscillation Theorem.

All the properties just described carry over to the natural CMV analogue of Prüfer variables (see Section 2 or [42, §10.12]). In the microscopic scaling we consider, the key quantity is the Prüfer phase relative to that at a reference energy. Because our model has rotational symmetry, we can choose the reference energy to be $z = 1$ without loss of generality. In this way, we will find a natural process $k \mapsto \psi_k$, whose values are increasing functions $\psi_k : (-\pi, \pi) \to \mathbb{R}$, with the property that the eigenvalues of $\mathcal{C}^{(n)}$ lie at the angles where ψ_{n-1} matches the random boundary condition η. Actually, that is a lie since ψ is merely the *relative* Prüfer phase; nevertheless, because of the way we choose our boundary condition, it is true that the eigenvalues of $\mathcal{C}^{(n)}$ are related to the random increasing function ψ_{n-1} in the sense of Definition 1.10.

The key to proving both that the $C\beta E_n$ processes converge and part (ii) of Theorem 1.7 is showing that in the natural scaling, $\Psi_n(x) = \psi_{n-1}(\frac{x}{n})$, the relative Prüfer phase process converges as $n \to \infty$. As we will see in Section 4, this is indeed the case; moreover we find an alternate realization of the limiting distribution, which we now describe.

Let $B_1(t)$ and $B_2(t)$ be independent Brownian motions. From Proposition 4.5, it follows that for each x, the stochastic differential equation

$$d\Psi = x\,dt + \tfrac{2}{\sqrt{\beta t}} \operatorname{Im} \left\{ \left[e^{i\Psi} - 1 \right] \left[dB_1(t) + i\,dB_2(t) \right] \right\} \tag{7}$$

has a unique solution $\Psi(t;x)$ that obeys $\mathbb{E}\{\Psi(t;x)\} = xt$ and $x\Psi(t;x) \geq 0$. We will prove that for any finite collection $x_1, \ldots, x_p$ the vector-valued random variable $\Psi_n(x_\mu/n)$ converges in distribution to $\Psi(t = 1; x_\mu)$. (Note that that a single pair of Brownian motions is used in (7) for all values of x.) By Proposition 2.5, this notion of convergence is strong enough to imply that the Laplace functionals of the eigenvalue process converge to those of the point process associated to the random increasing function $\Psi(t = 1, x)$.

While we where typing the last section of this work, two preprints, [18, 38], appeared on the arXiv. These works (independent of our own) discuss matters related to our proof of part (ii) of Theorem 1.7. The latter preprint discusses the edge statistics for the Gaussian β ensembles in the $n \to \infty$ limit. Specifically, they show that in the (naturally re-scaled) limit, the point process can be modelled as the eigenvalues of a half-line Schrödinger operator with stochastic potential. Our representation of the limiting β-ensemble has an analogous interpretation: Equation (7) describes the relative Prüfer phase for the Kreĭn system with 'potential' $A(x) = (\beta x)^{-1/2}[dB_3(t) + i\,dB_4(t)]$, where $B_3(x)$ and $B_4(x)$ are independent Brownian motions. In this way, we may say that the β-ensemble is modelled by the eigenvalues of this Kreĭn system with random self-adjoint boundary condition at $x = 1$. (For a pedagogical account of the theory of Kreĭn systems, we recommend [12].) The remarks in the final section of [18] seem to indicate that the authors were seeking a result of this type.

Subsequent to the completion of this paper, we learned that Valkó and Virág were working contemporaneously (and independently) on a related problem. Their work has now been written up, [51]. Briefly, they study the limiting eigenvalue statistics (in the bulk regime) of a model introduced in [13] for the Gaussian β-ensemble. They obtain a closely related description of the limit process.

Notation. We will write $X \lesssim Y$ to indicate that $X \leq CY$ for some constant C. Typically, this is an absolute numerical constant whose exact value is of no consequence. When it depends some parameters, we will endeavour to point this out.

2. The basic process

The theory of CMV matrices is intimately connected to the theory of orthogonal polynomials on the unit circle, just as Jacobi matrices are inherently tied to polynomials orthogonal with respect to a probability measure on the real line. We will confine ourselves to a review of the material we need; a more systematic treatment can be found in [41].

Let $d\mu$ denote the spectral measure associated to $\mathcal{C}$ (a semi-infinite CMV matrix) and the vector $e_1 = (1, 0, 0, \ldots)$. As $\mathcal{C}$ is unitary and e_1 is a unit vector, this is a probability measure on the unit circle in $\mathbb{C}$. Applying the Gram–Schmidt procedure to $\{1, z, z^2, \ldots\}$ leads to a sequence of monic orthogonal polynomials. We write $\Phi_k(z)$ for the polynomial of degree k.

Two important properties carry over from the Jacobi matrix case: Firstly, the orthogonal polynomials obey a recurrence relation:

$$\begin{aligned}
\Phi_{k+1}(z) &= z\Phi_k(z) - \bar{\alpha}_k \Phi_k^*(z) \\
\Phi_{k+1}^*(z) &= \Phi_k^*(z) - \alpha_k z \Phi_k(z)
\end{aligned} \tag{8}$$

with $\Phi_0^* = \Phi_0 = 1$. Here $\{\alpha_k\}_{k=0}^\infty \subset \mathbb{D}$ are the Verblunsky coefficients appearing in the matrix $\mathcal{C}$, while Φ_k^* denote the reversed polynomials,

$$\Phi_k^*(z) = z^k \overline{\Phi_k(\bar{z}^{-1})}. \tag{9}$$

Secondly, the characteristic polynomial of finite-volume truncations can be expressed in terms of these monic polynomials; specifically,

$$\det[z - \mathcal{C}^{(n)}] = z\Phi_{n-1}(z) - e^{-i\eta}\Phi_{n-1}^*(z). \tag{10}$$

The CMV analogue of the Weyl m-function used in connection with Jacobi matrices is the Carathéodory function,

$$F(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta) = \langle e_1 | (\mathcal{C} + z)(\mathcal{C} - z)^{-1} e_1 \rangle, \tag{11}$$

which will be important in our treatment of part (iii) of Theorem 1.7. The other two parts of that theorem rely on analysis of the relative Prüfer phase, which we will now describe.

Let $B_k(z) = z\Phi_k(z)/\Phi_k^*(z)$. In view of (9), this is a finite Blaschke product; moreover, by (8),

$$B_{k+1}(z) = zB_k(z)\frac{1 - \bar{\alpha}_k \bar{B}_k(z)}{1 - \alpha_k B_k(z)}. \tag{12}$$

Note also that $B_0(z) = z$. It is natural to define the (absolute) Prüfer phase as the argument of $B_k(z)$. By the Cauchy-Riemann equations, the argument of a finite Blaschke product is strictly increasing and by (10), the eigenvalues of $\mathcal{C}^{(n)}$ are precisely the points where $B_{n-1}(z) = e^{i\eta}$.

The relative Prüfer phase is the sequence of random continuous increasing functions $\psi_k : (-\pi, \pi) \to \mathbb{R}$ defined by

$$\exp\{i\psi_k(\theta)\} = \frac{B_k(e^{i\theta})}{B_k(1)} = e^{i\theta}\frac{\Phi_k(e^{i\theta})\Phi_k^*(1)}{\Phi_k^*(e^{i\theta})\Phi_k(1)}, \qquad 0 \le k < \infty, \tag{13}$$

and $\psi_k(0) = 0$. Strictly speaking, this only defines ψ_k modulo $2\pi\mathbb{Z}$; in Proposition 2.2, we will give a precise definition of ψ_k by means of a recurrence relation. This relation is simpler when expressed in terms of a new sequence of random variables γ_k in place of the Verblunsky coefficients. Because we have chosen to study

the rotationally invariant problem, the two systems of coefficients follow the same law:

Lemma 2.1. *Let α_k be statistically independent with each following a rotationally invariant law and let B_k denote the Blaschke products defined via (12). Then, for each n, the random variables*

$$\gamma_k = B_k(1)\alpha_k, \quad k = 0, 1, 2, \ldots, n-2, \quad and \quad e^{-i\omega} = B_{n-1}(1)e^{i\eta}$$

are statistically independent. Moreover, they have the same joint distribution as $\alpha_0, \ldots, \alpha_{n-2}$ and $e^{i\eta}$.

Proof. One merely needs to make three observations. First, the arguments of the α_k together with $e^{i\eta}$ are all statistically independent of one another and of the moduli of all α_k. Second, these arguments are all uniformly distributed on the circle. Third, if X and Y are independent random variables with values on the unit circle and Y is uniformly distributed, then the product, XY, is independent of X and is uniformly distributed. In particular, this means that despite the fact that $B_k(1)$ is a complicated function of $\alpha_0, \ldots, \alpha_{k-1}$ the argument of γ_k is uniform on the circle and independent of $\alpha_0, \ldots, \alpha_{k-1}$. □

Proposition 2.2. *Let $\psi_k : (-\pi, \pi) \to \mathbb{R}$ be defined by the recurrence*

$$\psi_{k+1} = \psi_k + \theta + 2\,\mathrm{Im}\log\left[\frac{1 - \gamma_k}{1 - \gamma_k e^{i\psi_k}}\right], \qquad \psi_0 = \theta, \tag{14}$$

where the branch of $\log$ is chosen so as to give 0 when $\gamma_k = 0$, that is, by the convergent power series

$$\Upsilon(\psi, \gamma) := 2\,\mathrm{Im}\log\left[\frac{1-\gamma}{1-\gamma e^{i\psi}}\right] = \mathrm{Im}\sum_{\ell=1}^{\infty} \tfrac{2}{\ell}[e^{i\ell\psi} - 1]\gamma^{\ell}. \tag{15}$$

Then (13) holds. Moreover, the eigenvalues of $\mathcal{C}^{(n)}$ are associated to $\psi_{n-1}(\theta)$ in the sense of Definition 1.10.

Proof. The first claim is an elementary computation: from (12),

$$\frac{B_{k+1}(e^{i\theta})}{B_{k+1}(1)} = e^{i\theta}\frac{B_k(e^{i\theta})}{B_k(1)}\frac{1 - \alpha_k B_k(1)}{1 - \bar{\alpha}_k \bar{B}_k(1)}\frac{1 - \bar{\alpha}_k \bar{B}_k(e^{i\theta})}{1 - \alpha_k B_k(e^{i\theta})}$$

and so (13) is equivalent to

$$e^{i\psi_{k+1}} = e^{i\theta + i\psi_k}\frac{1 - \gamma_k}{1 - \bar{\gamma}_k}\frac{1 - \bar{\gamma}_k e^{-i\psi_k}}{1 - \gamma_k e^{i\psi_k}}. \tag{16}$$

To finish the proof, one should note that $\zeta/\bar{\zeta} = \exp\{2i\,\mathrm{Im}\log(\zeta)\}$ for any $\zeta \in \mathbb{C}\setminus\{0\}$.

To prove the second claim, we note that by (10), the eigenvalues of $\mathcal{C}^{(n)}$ are the points $e^{i\theta}$ where $e^{i\eta}B_{n-1}(e^{i\theta}) = 1$ or equivalently, where $\psi_{n-1}(\theta) - \omega \in 2\pi\mathbb{Z}$. Recall that $e^{-i\omega} = B_{n-1}(1)e^{i\eta}$, which is statistically independent of $\psi_{n-1}(\theta)$. □

Warning. Despite the fact that $z = 1$ is an eigenvalue of $\mathcal{C}^{(n)}$ if and only if $\omega = 0$, ψ_{n-1} is not the natural conditioned process. Specifically, let f be a continuous

function on the unit circle, then

$$\mathbb{E}\big\{e^{-\operatorname{tr} f(\mathcal{C}^{(n)})}\big|1 \in \sigma(\mathcal{C}^{(n)})\big\}$$

$$:= \lim_{\epsilon\downarrow 0} \mathbb{E}\big\{e^{-\operatorname{tr} f(\mathcal{C}^{(n)})}\big|\sigma(\mathcal{C}^{(n)}) \cap e^{i[0,\epsilon]} \neq \emptyset\big\}$$

$$= \lim_{\epsilon\downarrow 0} \tfrac{2\pi}{n\epsilon} \mathbb{E}\bigg\{\int_0^{\psi_{n-1}(\epsilon)} \exp\big[-\textstyle\sum_m f \circ \exp\big(i\psi_{n-1}^{-1}(2\pi m + \omega)\big)\big] \tfrac{d\omega}{2\pi}\bigg\}$$

$$= \mathbb{E}\bigg\{\tfrac{1}{n}\psi_{n-1}'(0) \exp\Big[-\textstyle\sum_m f \circ \exp\big(i\psi_{n-1}^{-1}(2\pi m)\big)\Big]\bigg\}.$$

Hence the probability distribution picks up an additional weight, $\frac{1}{n}\psi_{n-1}'(0)$. It is not difficult to derive

$$\psi_{n-1}'(0) = 1 + \sum_{k=0}^{n-2}\prod_{l=0}^{k}\bigg(\operatorname{Re}\frac{1+\gamma_l}{1-\gamma_l}\bigg)$$

from (14).

We will use the following properties of the function Υ:

Proposition 2.3. *Given $0 \leq r < 1$ and $\phi, \psi \in \mathbb{R}$,*

$$\int_0^{2\pi} \Upsilon(\psi, re^{i\theta})\, \tfrac{d\theta}{2\pi} = 0 \tag{17}$$

$$\int_0^{2\pi} \big|\Upsilon(\psi, re^{i\theta})\big|^2\, \tfrac{d\theta}{2\pi} \leq \tfrac{4\pi^2}{3}r^2 \tag{18}$$

$$\bigg|\int_0^{2\pi} \Upsilon(\psi, re^{i\theta})\Upsilon(\phi, re^{i\theta})\, \tfrac{d\theta}{2\pi}\bigg| \leq 2\big(|\psi| + |\phi|\big)\log\big[\tfrac{1}{1-r^2}\big] \tag{19}$$

$$\bigg|2r^2\operatorname{Re}\big\{[e^{i\psi}-1][e^{-i\phi}-1]\big\} - \int_0^{2\pi} \Upsilon(\psi, re^{i\theta})\Upsilon(\phi, re^{i\theta})\, \tfrac{d\theta}{2\pi}\bigg|$$

$$\leq 2\big(|\psi| + |\phi|\big)r^2\log\big[\tfrac{1}{1-r^2}\big]. \tag{20}$$

We also have bounds for the third and fourth moments:

$$\int_0^{2\pi} \big|\Upsilon(\psi, re^{i\theta})\big|^3\, \tfrac{d\theta}{2\pi} \lesssim \min\{|\psi|, \pi\}\log^{3/2}\big[\tfrac{1}{1-r^2}\big] \tag{21}$$

$$\int_0^{2\pi} \bigg|\log\big[\tfrac{1-re^{i\theta}}{1-re^{i\theta+i\psi}}\big]\bigg|^4\, \tfrac{d\theta}{2\pi} \leq 16|\psi|\log^2\big[1-r^2\big]. \tag{22}$$

Proof. Equation (17) follows immediately from the mean value theorem for harmonic functions.

For the second, third, and fourth claims, we apply Parseval's identity to see that

$$\int_0^{2\pi} \Upsilon(\psi, re^{i\theta})\Upsilon(\phi, re^{i\theta})\, \tfrac{d\theta}{2\pi}$$

$$= \sum_{\ell=1}^{\infty} \tfrac{4r^{2\ell}}{\ell^2}\int \operatorname{Im}\big\{[e^{i\ell\psi} - 1]e^{i\ell\theta}\big\}\operatorname{Im}\big\{[e^{i\ell\phi} - 1]e^{i\ell\theta}\big\}\, \tfrac{d\theta}{2\pi}$$

$$= \sum_{\ell=1}^{\infty} \tfrac{2r^{2\ell}}{\ell^2}\operatorname{Re}\big\{[e^{i\ell\psi} - 1][e^{-i\ell\phi} - 1]\big\}.$$

Inequality (18) follows by bounding this by $8r^2 \sum \ell^{-2}$. As $|e^{i\theta} - 1| \leq |\theta|$, we have

$$\left| \mathrm{Re}\{[e^{i\ell\psi} - 1][e^{-i\ell\phi} - 1]\} \right| \leq \ell|\psi| + \ell|\phi|.$$

To deduce (19), we simply substitute this estimate into the identity above and then sum the resulting series. The inequality (20) follows from the identity above by similar reasoning.

We now turn to (22); we will treat (21) last. Applying Parseval's identity to the square of (15) gives

$$\int_0^{2\pi} \left| \log\left[\frac{1-re^{i\theta}}{1-re^{i\theta+i\psi}} \right] \right|^4 \frac{d\theta}{2\pi} = \sum_{q=2}^{\infty} r^{2q} \left| \sum_{\ell=1}^{q-1} \frac{1}{\ell(q-\ell)} [e^{i\ell\psi} - 1][e^{i(q-\ell)\psi} - 1] \right|^2. \tag{23}$$

We may estimate the sum over ℓ in each of two ways: First, one may bound the modulus of each term in square brackets by 2. Second, as the sum is symmetrical under $\ell \mapsto (q - \ell)$, we may bound it by

$$2 \sum_{\ell=1}^{\lfloor q/2 \rfloor} \frac{|e^{i\ell\psi} - 1|}{\ell} \cdot \frac{|e^{i(q-\ell)\psi} - 1|}{(q-\ell)} \leq 2 \sum_{\ell=1}^{\lfloor q/2 \rfloor} |\psi| \cdot \frac{4}{q} \leq 4|\psi|.$$

Using each of these estimates for one factor of the sum in (23) gives

$$\mathrm{LHS}(22) \leq 16|\psi| \sum_{q=2}^{\infty} r^{2q} \sum_{\ell=1}^{q-1} \frac{1}{\ell(q-\ell)} = 16|\psi| \log^2\left[\frac{1}{1-r^2} \right]. \tag{24}$$

When $|\psi| \leq \pi$, we may obtain (21) by interpolating between (22) and (19). For $|\psi| \geq \pi$, we may simply note that the left-hand side depends only of the value of ψ modulo $2\pi\mathbb{Z}$. □

Corollary 2.4. *If the γ_k are statistically independent and follow rotationally invariant laws, then*

$$\mathbb{E}\{\psi_s(\theta)\} = (s+1)\theta, \qquad \mathbb{E}\{|\psi_s(\theta)|\} = (s+1)|\theta|, \tag{25}$$

and for $k \leq m$,

$$\mathbb{E}\left\{ |\psi_m(\tfrac{x}{n}) - \psi_k(\tfrac{x}{n}) - \tfrac{x(m-k)}{n}|^2 \right\} = \sum_{s=k}^{m-1} \mathbb{E}\left\{ |\Upsilon(\psi_s(\tfrac{x}{n}), \gamma_s)|^2 \right\}. \tag{26}$$

Proof. For each s, $\psi_s(\theta)$ is an increasing function, which vanishes at $\theta = 0$. Thus $\psi_s(\theta)$ takes only non-negative values for $\theta > 0$ and non-positive values for $\theta < 0$. For this reason, the second claim follows from the first. As ψ_k is independent of γ_k, the formulae for the mean and the variance follow quickly from the recurrence relation, (14), together with (17). □

As outlined in the introduction, we will deduce convergence of the point processes from that of the relative Prüfer phase. The next result shows how this can be done.

Proposition 2.5. *Suppose that $\Psi_n(x)$, $n \in \mathbb{Z}$, and $\Psi(x)$ are random non-decreasing functions with two properties: (i) for all $x \in \mathbb{R}$,*

$$\sup_n \mathbb{E}\{|\Psi_n(x)|\} < \infty, \tag{27}$$

and (ii) for any finite collection of $x_j \in \mathbb{R}$, the joint distribution of $\Psi_n(x_j)$ converges to that of $\Psi(x_j)$. Then the point processes associated to Ψ_n converge in distribution to that associated to Ψ.

Proof. The probabilistic portion of the proof is quite simple once we have the right deterministic information. To this end, let Ψ and Φ be non-decreasing functions on $\mathbb{R}$ and let $f \in C_c^\infty(\mathbb{R})$ be non-negative. Elementary considerations show that

$$\int_0^{2\pi} \left| \exp\left\{ -\sum_{m \in \mathbb{Z}} f \circ \Phi^{-1}(2\pi m + \omega) \right\} - \exp\left\{ -\sum_{m \in \mathbb{Z}} f \circ \Psi^{-1}(2\pi m + \omega) \right\} \right| d\omega \quad (28)$$

$$\leq \int_0^{2\pi} \sum_{m \in \mathbb{Z}} \left| f \circ \Phi^{-1}(2\pi m + \omega) - f \circ \Psi^{-1}(2\pi m + \omega) \right| d\omega \quad (29)$$

$$\leq \int_{\mathbb{R}} \left| f \circ \Phi^{-1}(\omega) - f \circ \Psi^{-1}(\omega) \right| d\omega \quad (30)$$

Next, we choose $[a, b] \subset \mathbb{R}$ so that $\text{supp}(f) \subset [a, b]$. Then

$$|f(x) - f(y)| \leq |\tilde{x} - \tilde{y}| \cdot \|f'\|_{L^\infty} \quad \text{where} \quad \tilde{x} = \begin{cases} a & : \ x < a \\ x & : \ x \in [a, b] \\ b & : \ x > b \end{cases}$$

and similarly for $\tilde{y}$. Therefore,

$$\int_{\mathbb{R}} |f \circ \Phi^{-1}(\omega) - f \circ \Psi^{-1}(\omega)| \, d\omega \leq \|f'\|_{L^\infty} \int_{\mathbb{R}} |\widetilde{\Phi^{-1}(\omega)} - \widetilde{\Psi^{-1}(\omega)}| \, d\omega \quad (31)$$

$$= \|f'\|_{L^\infty} \int_a^b |\Phi(x) - \Psi(x)| \, dx. \quad (32)$$

The equality follows by realizing that both integrals give the area of the region between the graphs of Φ and Ψ over the interval $[a, b]$ (with vertical lines drawn at any jumps).

Combining our efforts so far, we have bounded the expression in (28) by that in (32). We will make one further reduction. Given an integer $p \geq 1$, let us partition $[a, b]$ into p equal intervals by introducing $x_j = a + j(b - a)/p$ with $0 \leq j \leq p$. As both Φ and Ψ are non-decreasing,

$$\int_{x_j}^{x_{j+1}} |\Phi(x) - \Psi(x)| \, dx \quad (33)$$

$$\leq \tfrac{b-a}{p} \left[\max\{\Phi(x_{j+1}), \Psi(x_{j+1})\} - \min\{\Phi(x_j), \Psi(x_j)\} \right] \quad (34)$$

$$\leq \tfrac{b-a}{p} \left[|\Psi(x_{j+1}) - \Psi(x_j)| + |\Phi(x_{j+1}) - \Psi(x_{j+1})| + |\Phi(x_j) - \Psi(x_j)| \right], \quad (35)$$

which shows that

$$\int_a^b |\Phi(x) - \Psi(x)| \, dx \leq \tfrac{b-a}{p} |\Psi(b) - \Psi(a)| + 2\tfrac{b-a}{p} \sum_{j=0}^p |\Phi(x_j) - \Psi(x_j)|. \quad (36)$$

Let us now begin the probabilistic portion of the argument. We will think of $\Psi_n(x_j)$ and $\Psi(x_j)$ as vector-valued random variables where $0 \leq j \leq p$ indexes the components. By the Skorohod coupling, we can realize the random variables $\Psi(x_j)$ and all $\Psi_n(x_j)$ on a single probability space so that $\Psi_n(x_j) \to \Psi(x_j)$ almost surely. Using (36) and the arbitrariness of $p \geq 1$, we may conclude that

$$\limsup_{n \to \infty} \int_a^b |\Psi_n(x) - \Psi(x)| \, dx = 0 \quad (37)$$

almost surely. Tracing back through the inequalities above, we can deduce convergence of the Laplace functionals from the Dominated Convergence Theorem.

The first condition is merely a heavy-handed way to ensure that Ψ and Ψ_n are associated to honest point processes, that is, so that the number of points in a compact set is almost surely finite (cf. the remark after Definition 1.10). $\qquad \square$

Remark: Once one knows that the distributions of $\Psi_n(x_j)$ converge for any finite set of x_j and that (27) holds, the existence of a limiting $\Psi(x)$, can be deduced from the Kolmogorov extension theorem, rather than being presupposed, as is done in this proposition.

3. Fast decay

Theorem 3.1. *If* $\mathbb{E}\{|\alpha_s|^2\} = o(s^{-1})$ *then* $\mathbb{E}\{|\psi_{n-1}(\frac{x}{n}) - x|\} \to 0$ *for each* $x \in \mathbb{R}$. *This implies that locally, the eigenvalues follow clock statistics.*

Proof. As $\mathbb{E}\{|\gamma_s|^2\} = o(s^{-1})$,

$$\sum_{n \geq s \geq \epsilon n} \mathbb{E}\{|\gamma_s|^2\} = o(1)$$

as $n \to \infty$ for any choice of ϵ. Therefore, there is a sequence k_n that is $o(n)$ and obeys

$$\sum_{s=k_n}^{n} \mathbb{E}\{|\gamma_s|^2\} = o(1)$$

as $n \to \infty$. Combining this with (26) and (18) we have,

$$\mathbb{E}\{|\psi_{n-1}(\tfrac{x}{n}) - \psi_{k_n}(\tfrac{x}{n}) - (n-1-k_n)\tfrac{x}{n}|^2\} \leq \tfrac{4\pi^2}{3} \sum_{s=k_n}^{n-2} \mathbb{E}\{|\gamma_s|^2\} = o(1).$$

Now by the triangle inequality and the Cauchy–Schwarz inequality,

$$\mathbb{E}\{|\psi_{n-1}(\tfrac{x}{n}) - x|\} \leq \mathbb{E}\{|\psi_{k_n}(\tfrac{x}{n}) - (k_n+1)\tfrac{x}{n}|\} + o(1)$$
$$\leq 2(k_n+1)\tfrac{|x|}{n} + o(1),$$

which is $o(1)$ by the way we chose k_n. Note that the second inequality follows from (25).

Convergence of the point processes follows from Proposition 2.5. $\qquad \square$

Remark. If $\mathbb{E}\{|\alpha_k|^2\} \in \ell^1$, then one obtains clock spacing on an almost macroscopic scale. Indeed for any (non-random) sequence x_n with $|x_n| = o(n)$,

$$\mathbb{E}\{|\psi_{n-1}(x_n/n) - x_n|\} \to 0$$

as $n \to \infty$. The proof follows the outline above, but with $k_n \to \infty$ chosen so that $x_n k_n = o(n)$.

4. Critical decay

Given $\vec{x} \in \mathbb{R}^p$ with coordinates x_μ, we define a sequence of $\mathbb{R}^p$-valued processes $\Psi_n(t)$ on $[0, \infty)$ as follows: for all $1 \leq \mu \leq p$,

$$\Psi_{n,\mu}(0) = 0 \tag{38}$$
$$\Psi_{n,\mu}\left(\tfrac{k+1}{n}\right) = \psi_k\left(\tfrac{x_\mu}{n}\right) \qquad \text{for } 0 \leq k < \infty \tag{39}$$

and for intermediate values of t, we define $\Psi_n(t)$ by linear interpolation. The goal of this section is to prove that this sequence of processes converges to a limit that is independent of the specifics of the distribution of the random variables γ_k. The

notion of convergence will be sufficiently strong (cf. Proposition 2.5) to show that the eigenvalue point processes associated to $\mathcal{C}^{(n)}$ converge to the point process associated to $\Psi = \lim_{n \to \infty} \Psi_n$, in the sense of Definition 1.10. We will further give a characterization of Ψ which shows its law depends only on the parameter β. In view of Theorem 1.9, we also obtain equality with (including existence of) the β-ensemble.

Our first job will be to prove the existence of even just a subsequential limit of the processes Ψ_n. This will follow once we prove tightness, which is the role of Proposition 4.4 below.

Consistent with our goal of proving part (ii) of Theorem 1.7, will we assume for the remainder of this section that

$$\mathbb{E}\left\{|\gamma_k|^2\right\} = \tfrac{2}{\beta(k+1)}(1 + \mathfrak{e}_k) \tag{40}$$

$$\mathbb{E}\left\{\log^2\left[\tfrac{1}{1-|\gamma_k|^2}\right]\right\} \leq \tfrac{1}{(k+1)}\mathfrak{e}_k \tag{41}$$

$$\mathbb{E}\left\{\log^{3/2}\left[\tfrac{1}{1-|\gamma_k|^2}\right]\right\} \leq \tfrac{1}{(k+1)}\mathfrak{e}_k \tag{42}$$

for some decreasing sequence of positive numbers $\mathfrak{e}_k$ converging to zero. Note that while (42) does not appear among the assumptions of this part of the Theorem, it is easily deduced from them. Perhaps the simplest argument is to notice that

$$\log^{3/2}\left[\tfrac{1}{1-|\gamma_k|^2}\right] \lesssim \left\{|\gamma_k|^2 \log^2\left[\tfrac{1}{1-|\gamma_k|^2}\right]\right\}^{1/2} + \log^2\left[\tfrac{1}{1-|\gamma_k|^2}\right],$$

uniformly for $\gamma_k \in \mathbb{D}$, and then apply the Cauchy–Schwarz inequality to the first summand after taking expectations. In addition, from (40) and (41) we see that

$$\mathbb{E}\left\{\log\left[\tfrac{1}{1-|\gamma_k|^2}\right]\right\} \leq \mathbb{E}\left\{|\gamma_k|^2 + \log^2\left[\tfrac{1}{1-|\gamma_k|^2}\right]\right\} \lesssim \tfrac{1}{(k+1)}. \tag{43}$$

Here and for the remainder of this section the implicit constants hidden in the $\lesssim$ notation are permitted to depend on β, the properties of the sequence $\mathfrak{e}_k$, and on the values of x under consideration; they will never depend on n.

Before proceeding, let us pause to demonstrate that the exactly soluble model of Theorem 1.9 obeys the hypotheses given above. If $\alpha \sim \Theta_\nu$, then

$$\mathbb{E}\left\{\log^m\left[\tfrac{1}{1-|\alpha|^2}\right]\right\} = \tfrac{\nu-1}{2}\int_0^\infty t^m e^{-(\nu-1)t/2}\, dt = \Gamma(m+1)\left(\tfrac{2}{\nu-1}\right)^m,$$

by making the change of variables $e^{-t} = 1 - |\alpha|^2$.

Lemma 4.1. *For fixed x and $m \geq k$,*

$$\mathbb{E}\left\{|\psi_m(\tfrac{x}{n}) - \psi_k(\tfrac{x}{n}) - \tfrac{x(m-k)}{n}|^2\right\} \lesssim \tfrac{m-k}{n}. \tag{44}$$

In particular, from $k = 0$ we have $\mathbb{E}\left\{|\psi_s(\tfrac{x}{n})|^2\right\} \lesssim \tfrac{s}{n} + (\tfrac{s+1}{n})^2$. Furthermore,

$$\mathbb{P}\left\{\sup_{k \leq s \leq m}|\psi_m(\tfrac{x}{n}) - \psi_k(\tfrac{x}{n}) - \tfrac{x(m-k)}{n}| \geq \epsilon\right\} \lesssim \epsilon^{-4}\left(\tfrac{m-k}{n}\log\left[\tfrac{m+1}{k+1}\right] + \tfrac{m-k}{n}\mathfrak{e}_k\right) \tag{45}$$

Proof. Let k be fixed and define

$$\phi_s = \psi_s(\tfrac{x}{n}) - \psi_k(\tfrac{x}{n}) - \tfrac{x(s-k)}{n}$$

for $s \geq k$. Note that $\phi_{s+1} = \phi_s + \Upsilon(\psi_s, \gamma_s)$. Equation (17) shows that this is a martingale; therefore, using (19), (43) and (25)

$$\mathbb{E}\{|\phi_m|^2\} = \sum_{s=k}^{m-1} \mathbb{E}\{|\Upsilon(\psi_s, \gamma_s)|^2\} \lesssim \sum_{s=k}^{m-1} \mathbb{E}\left\{|\psi_s| \log\left[\frac{1}{1-|\gamma_s|^2}\right]\right\} \lesssim \frac{m-k}{n}.$$

This proves (44); we now consider (45). By (17),

$$\begin{aligned} \mathbb{E}\{|\phi_{s+1}|^4\} = \mathbb{E}\{|\phi_s|^4\} &+ 6\,\mathbb{E}\{|\phi_s|^2|\Upsilon(\psi_s, \gamma_s)|^2\} \\ &+ 4\,\mathbb{E}\{\phi_s \Upsilon(\psi_s, \gamma_s)^3\} + \mathbb{E}\{|\Upsilon(\psi_s, \gamma_s)|^4\}. \end{aligned} \tag{46}$$

Using (18) and (44),

$$\mathbb{E}\{|\phi_s|^2|\Upsilon(\psi_s, \gamma_s)|^2\} \lesssim \mathbb{E}\{|\phi_s|^2|\gamma_s|^2\} \lesssim \frac{s-k}{n}\frac{1}{s+1}$$

while by (21), (25) and (42),

$$\begin{aligned} \mathbb{E}\{|\phi_s||\Upsilon(\psi_s, \gamma_s)|^3\} &\lesssim \mathbb{E}\left\{\left[|\psi_s| + |\psi_k| + (s-k)\frac{x}{n}\right] \log^{3/2}\left[\frac{1}{1-|\gamma_s|^2}\right]\right\} \\ &\lesssim \frac{s+1}{n}\frac{1}{s+1}\mathfrak{e}_s. \end{aligned}$$

Similarly using (22) and (41),

$$\mathbb{E}\{|\Upsilon(\psi_s, \gamma_s)|^4\} \lesssim \frac{s+1}{n}\frac{1}{s+1}\mathfrak{e}_s.$$

By substituting the last three estimates into (46), we obtain

$$\left|\mathbb{E}\{|\phi_{s+1}|^4\} - \mathbb{E}\{|\phi_s|^4\}\right| \lesssim \frac{s-k}{n}\frac{1}{s+1} + \frac{1}{n}\mathfrak{e}_s$$

By summing the preceding estimate over s and using the fact that $\phi_k \equiv 0$, we obtain

$$\mathbb{E}\{|\phi_m|^4\} \lesssim \frac{m-k}{n}\log\left[\frac{m+1}{k+1}\right] + \frac{m-k}{n}\mathfrak{e}_k.$$

Thus (45) follows by the martingale maximal theorem. $\qquad\square$

Proposition 4.2. *For fixed $\vec{x}$, the processes $\Psi_n(t)$ define a tight family of measures on $C([0, \infty); \mathbb{R}^p)$ with the topology of uniform convergence on compact sets.*

Proof. We will prove that for fixed x and $T \geq 1$,

$$\limsup_{n\to\infty} \mathbb{P}\Big(2\epsilon \leq \sup\big\{|\psi_s\big(\tfrac{x}{n}\big) - \psi_t\big(\tfrac{x}{n}\big)| : 0 \leq s < t \leq nT \text{ and } |s-t| \leq n\delta\big\}\Big) = o(1)$$

as $\delta \downarrow 0$. The Ascoli–Arzela theorem shows that this suffices (cf. [49, §1.3]).

It suffices to consider the case $n\delta > 2$. We wish to break the problem into blocks; to this end, let $k_j = \lceil jn\delta \rceil$, that is, $jn\delta$ rounded up to an integer. The zeroth block is easiest: as $\psi_s(\tfrac{x}{n}) - (s+1)\tfrac{x}{n}$ is a martingale, we may apply the weak-type $(1,1)$ bound for the maximal function (a.k.a. Doob's inequality),

$$\begin{aligned} \mathbb{P}\Big\{|\psi_s\big(\tfrac{x}{n}\big)| \geq \epsilon \text{ for some } 0 \leq s \leq k_1\Big\} &\lesssim \epsilon^{-1}\,\mathbb{E}\Big\{|\psi_{k_1}\big(\tfrac{x}{n}\big)| + (k_1 + 1)\tfrac{|x|}{n}\Big\} \\ &\lesssim \delta\epsilon^{-1}. \end{aligned} \tag{47}$$

Naturally this bounds the probability that $|\psi_s - \psi_t| \geq 2\epsilon$ for any pair s, t in the interval $[0, k_1]$.

For $j \geq 1$, we can apply Lemma 4.1 with ϵ half the size and so obtain

$$\mathbb{P}\Big\{\big|\psi_s\big(\tfrac{x}{n}\big) - \psi_t\big(\tfrac{x}{n}\big)\big| \geq \epsilon \text{ for some } k_j \leq s < t \leq k_{j+1}\Big\}$$
$$\lesssim \epsilon^{-4}\Big[\delta \log\big[\tfrac{j+1}{j}\big] + \delta \mathfrak{e}_{k_j} + (\delta x)^4\Big].$$

Because $\mathfrak{e}_s$ decreases to zero,

$$J^{-1} \sum_{j=1}^{J} \mathfrak{e}_{k_j} \leq J^{-1} \sum_{j=1}^{J} \mathfrak{e}_j = o(1)$$

as $J \to \infty$ uniformly in the region $n\delta > 2$. Combining this with $\log[(j+1)/j] \lesssim j^{-1}$, we obtain

$$\limsup_{n\to\infty} \sum_{j=1}^{\lceil T/\delta \rceil} \mathbb{P}\Big\{\big|\psi_s\big(\tfrac{x}{n}\big) - \psi_t\big(\tfrac{x}{n}\big)\big| \geq \epsilon \text{ for some } k_j \leq s < t \leq k_{j+1}\Big\}$$
$$\lesssim \epsilon^{-4}\Big[\delta \log\big[1 + T\delta^{-1}\big] + o(1) + Tx^4\delta^3\Big] = o(1)$$

where $o(1)$ means as $\delta \downarrow 0$ with T and x fixed.

Combining the estimate just proved with (47) gives the claim made in the first paragraph of the proof. Note that the change from ϵ to 2ϵ comes from the fact that the points s and t may lie in adjacent blocks rather than the same one. □

We now begin to examine the properties of any subsequential limit of $\Psi_n(t)$; specifically, we will prove that any such limit is the solution of a certain martingale problem. This approach to studying diffusions was introduced by Stroock and Varadhan and is very convenient for proving convergence; indeed much of this section is modelled on [49, §11.2].

We can no longer treat the different values of x separately. For this reason, it is helpful to introduce some notation for the discrete vector problem. Because the equations are going to become rather long, we need to make it as compact as possible. Let $\vec{\psi}_s^{(n)}$ denote the vector with components

$$\vec{\psi}_{s,\mu}^{(n)} = \psi_s\big(\tfrac{x_\mu}{n}\big), \qquad \mu = 1, 2, \ldots, p.$$

Our first lemma shows that this discrete process almost solves a martingale problem.

Lemma 4.3. *Given $\vec{x}$ and $f \in C_c^\infty(\mathbb{R}^p)$, let*

$$\tilde{M}_n(m,k) = f(\vec{\psi}_m^{(n)}) - \sum_{s=k}^{m-1}\Big[x_\mu f_\mu(\vec{\psi}_s^{(n)}) + \tfrac{2}{\beta(s+1)} f_{\mu\nu}(\vec{\psi}_s^{(n)}) \operatorname{Re}\big\{[e^{i\vec{\psi}_{s,\mu}^{(n)}} - 1][e^{-i\vec{\psi}_{s,\nu}^{(n)}} - 1]\big\}\Big]$$

where subscripts on f denote partial derivatives and μ, ν are summed over whenever repeated. Then for fixed integer $T > 0$,

$$\mathbb{E}\Big\{\sup_{0 \leq k < m \leq Tn} \big|\mathbb{E}\big\{\tilde{M}_n(m,k) - f(\vec{\psi}_k^{(n)})\big|\vec{\psi}_k^{(n)}\big\}\big|\Big\} \to 0$$

as $n \to \infty$.

Proof. The recurrence relation for ψ says

$$\vec{\psi}_{s+1,\mu}^{(n)} = \vec{\psi}_{s,\mu}^{(n)} + \tfrac{x_\mu}{n} + \Upsilon\big(\vec{\psi}_{s,\mu}^{(n)}, \gamma_s\big). \tag{48}$$

Using this we expand $f(\vec{\psi}_{s+1}^{(n)})$ in a Taylor series to third order. By incorporating (17) we find

$$\mathbb{E}\big\{f(\vec{\psi}_{s+1}^{(n)})\big|\vec{\psi}_s^{(n)}\big\} = f(\vec{\psi}_s^{(n)}) + \tfrac{x_\mu}{n}f_\mu(\vec{\psi}_s^{(n)}) + A_s^{\mu\nu}f_{\mu\nu}(\vec{\psi}_s^{(n)}) + b_s \qquad (49)$$

where A_s is the sequence of matrices

$$A_s^{\mu\nu} = \tfrac{x_\mu x_\nu}{2n^2} + \mathbb{E}\big\{\tfrac{1}{2}\Upsilon\big(\vec{\psi}_{s,\mu}^{(n)},\gamma_s\big)\Upsilon\big(\vec{\psi}_{s,\nu}^{(n)},\gamma_s\big)\big|\vec{\psi}_s^{(n)}\big\}$$

and b_s are scalars obeying

$$|b_s| \lesssim \tfrac{1}{n^3} + \sup_\mu \mathbb{E}\big\{\big|\Upsilon\big(\vec{\psi}_{s,\mu}^{(n)},\gamma_s\big)\big|^3\big|\vec{\psi}_s^{(n)}\big\}.$$

(The implicit constant here depends on $\vec{x}$ and the properties of f; this will be the case throughout the proof.) These expressions can be better understood using results from Proposition 2.3. By (20) and the fact that $r^2 \leq \log[\tfrac{1}{1-r^2}]$ for $r \in [0,1)$, we have

$$\bigg|\mathbb{E}\big\{\tfrac{1}{2}\Upsilon\big(\vec{\psi}_{s,\mu}^{(n)},\gamma_s\big)\Upsilon\big(\vec{\psi}_{s,\nu}^{(n)},\gamma_s\big)\big|\vec{\psi}_s^{(n)}\big\} - \mathbb{E}\big\{|\gamma_s|^2\big\}\,\mathrm{Re}\big\{[e^{i\vec{\psi}_{s,\mu}^{(n)}}-1][e^{-i\vec{\psi}_{s,\nu}^{(n)}}-1]\big\}\bigg|$$

$$\lesssim |\vec{\psi}_{s,\mu}^{(n)} + \vec{\psi}_{s,\nu}^{(n)}|\,\mathbb{E}\big\{|\gamma_s|^2\log\big[\tfrac{1}{1-|\gamma_s|^2}\big]\big\}$$

$$\lesssim |\vec{\psi}_{s,\mu}^{(n)} + \vec{\psi}_{s,\nu}^{(n)}|\,\mathbb{E}\big\{\log^2\big[\tfrac{1}{1-|\gamma_s|^2}\big]\big\}.$$

So by invoking (25) and (41), we obtain

$$\sum_{s=0}^{nT}\mathbb{E}\big\{\big|A_s^{\mu\nu} - \tfrac{2}{\beta(s+1)}\mathrm{Re}\big\{[e^{i\vec{\psi}_{s,\mu}^{(n)}}-1][e^{-i\vec{\psi}_{s,\nu}^{(n)}}-1]\big\}\big|\big\} \lesssim \tfrac{T}{n} + \tfrac{1}{n}\sum_{s=0}^{nT}\mathfrak{e}_s = o(1) \quad (50)$$

as $n \to \infty$ with T fixed. Similarly by (21), we have

$$\mathbb{E}\big\{\big|\Upsilon\big(\vec{\psi}_{s,\mu}^{(n)},\gamma_s\big)\big|^3\big|\vec{\psi}_s^{(n)}\big\} \lesssim |\vec{\psi}_{s,\mu}^{(n)}|\,\mathbb{E}\big\{\log^{3/2}\big[\tfrac{1}{1-|\gamma_s|^2}\big]\big\}$$

and so by (42),

$$\sum_{s=0}^{nT}\mathbb{E}\{|b_s|\} = o(1) \qquad (51)$$

as $n \to \infty$.

Combining (49), (50), and (51), we may conclude that

$$\lim_{n\to\infty}\sum_{s=0}^{nT}\mathbb{E}\Big\{\big|\mathbb{E}\big\{\tilde{M}_n(s+1,s) - f(\vec{\psi}_s^{(n)})\big|\vec{\psi}_s^{(n)}\big\}\big|\Big\} = 0, \qquad (52)$$

which is the key estimate for the completion of the proof.

For any $k < m$,

$$\mathbb{E}\big\{\tilde{M}_n(m,k) - f(\vec{\psi}_k^{(n)})\big|\vec{\psi}_k^{(n)}\big\} = \sum_{s=k}^{m-1}\mathbb{E}\big\{\tilde{M}_n(s+1,s) - f(\vec{\psi}_s^{(n)})\big|\vec{\psi}_k^{(n)}\big\}$$

$$= \sum_{s=k}^{m-1}\mathbb{E}\big\{\mathbb{E}\big\{\tilde{M}_n(s+1,s) - f(\vec{\psi}_s^{(n)})\big|\vec{\psi}_s^{(n)}\big\}\big|\vec{\psi}_k^{(n)}\big\},$$

because the process $k \mapsto \vec{\psi}_k^{(n)}$ is Markovian. Therefore,

$$\mathbb{E}\left\{\sup_{k<m}\left|\mathbb{E}\{\tilde{M}_n(m,k) - f(\vec{\psi}_k^{(n)})\,|\,\vec{\psi}_k^{(n)}\}\right|\right\} \le \sum_{s=0}^{nT}\mathbb{E}\left\{\left|\mathbb{E}\{\tilde{M}(s+1,s) - f(\vec{\psi}_s^{(n)})\,|\,\vec{\psi}_s^{(n)}\}\right|\right\},$$

which converges to 0 by (52).   □

Proposition 4.4. *Let $\vec{x}$ be fixed and let $\Psi(t)$ be a subsequential limit of the processes $\Psi_n(t)$. Then*
(a) If $x_\mu > x_\nu$ then $\Psi_\mu - \Psi_\nu$ is a non-negative function with probability one. Similarly, for each μ, $x_\mu\Psi_\mu$ is a non-negative function with probability one.
(b) $\mathbb{E}\{\|\Psi(t)\|^2\} \lesssim t + t^2$ for all $t \in [0,\infty)$.
(c) $\mathbb{E}\{\Psi_\mu(t)\} = x_\mu t$ for all $t \in [0,\infty)$.
(d) For all $t_0 > 0$ and any $f \in C_c^\infty(\mathbb{R}^p)$,

$$
\begin{aligned}
M(t) := f(\Psi(t)) - \int_{t_0}^t x_\mu f_\mu(\Psi(\tau))\,d\tau \\
- 2\int_{t_0}^t f_{\mu\nu}(\Psi(\tau))\,\mathrm{Re}\big\{[e^{i\Psi_\mu(\tau)} - 1][e^{-i\Psi_\nu(\tau)} - 1]\big\}\frac{d\tau}{\beta\tau}
\end{aligned}
\tag{53}
$$

is an $\mathcal{M}_t$-martingale starting at t_0. Here $\mathcal{M}_t$ is the sigma algebra generated by $\Psi(\tau)$ over the interval $t_0 \le \tau \le t$.

Proof. To avoid unsightly subscripts in what follows, n is restricted throughout to lie in the subsequence that converges rather than be an arbitrary positive integer.

(a) The subset K of functions with $\Psi_\mu - \Psi_\nu \ge 0$ is closed in $C([0,\infty);\mathbb{R}^p)$; therefore, $\mathbb{P}(\Psi \in K) \ge \limsup \mathbb{P}(\Psi_n \in K) = 1$. The same argument applies for the second statement.

(b) By taking $k = 0$ in (44) from Lemma 4.1,

$$\mathbb{E}\big\{|\Psi_{n,\mu}(\tfrac{s+1}{n})|^2\big\} = \mathbb{E}\big\{\big|\psi_s(\tfrac{x_\mu}{n})\big|^2\big\} \lesssim \tfrac{s}{n} + \big(\tfrac{s+1}{n}\big)^2.$$

One now simply sums over μ and takes $n \to \infty$.

(c) By (25), we know that $\mathbb{E}\{\Psi_{n,\mu}(t)\} = x_\mu t$. Moreover by hypothesis, we know that $\mathbb{E}\{f(\Psi_n(t))\} \to \mathbb{E}\{f(\Psi(t))\}$ for any bounded continuous function f. The result now follows from (b) by a simple approximation argument (cf. [15, §4.5]).

(d) It suffices to show that

$$\mathbb{E}\left\{G\big(\Psi|_{[t_0,t_1]}\big)\big[M(t) - M(t_1)\big]\right\} = 0$$

for any bounded continuous function G from $C([t_0,t_1];\mathbb{R}^p)$ to $[0,1]$ and any $t > t_1 \ge t_0$. As this random variable is a continuous function of Ψ, in the $C([t_0,t];\mathbb{R}^p)$ topology, we need only prove that

$$\lim_{n\to\infty} \mathbb{E}\left\{G\big(\Psi_n|_{[t_0,t_1]}\big)\big[M_n(t) - M_n(t_1)\big]\right\} = 0, \tag{54}$$

where $M_n(t)$ is the analogue of (53) with Ψ_n replacing Ψ. This will follow from Lemma 4.3 once we relate it back to the discrete process. We will use the notation

$$k = \lceil nt_1\rceil - 1 \quad \text{and} \quad m = \lfloor nt\rfloor - 1.$$

Note that $\Psi_n(t) = (1 - \theta)\vec{\psi}_m^{(n)} + \theta\vec{\psi}_{m+1}^{(n)}$ for some $\theta \in [0,1]$, while $\Psi_n(t_1)$ is a combination of $\vec{\psi}_k^{(n)}$ and $\vec{\psi}_{k-1}^{(n)}$.

The standard estimate for the trapezoid rule, [3, §15.19], tells us that

$$\left| \tfrac{1}{2}\big[g(a) + g(b)\big] - \int_0^1 g\big((1-u)a + ub\big)\,du \right| \le \tfrac{1}{12} \|b - a\|^2 \|g_{\mu\nu}\|_{L^\infty}$$

for any $a, b \in \mathbb{R}^p$ and any $g \in C_c^\infty(\mathbb{R}^p)$. Consequently,

$$\mathbb{E}\left\{ \left| \int_{t_1}^t g\big(\Psi_n(t)\big)\,dt - \frac{1}{n} \sum_{s=k}^m g\big(\vec{\psi}_s^{(n)}\big) \right| \right\} \tag{55}$$

$$\lesssim \frac{\|g\|_{L^\infty}}{n} + \frac{\|g_{\mu\nu}\|_{L^\infty}}{n} \sum_{s=k}^m \mathbb{E}\big\{ \|\vec{\psi}_{s+1}^{(n)} - \vec{\psi}_s^{(n)}\|^2 \big\} \tag{56}$$

$$\lesssim \frac{\|g\|_{L^\infty}}{n} + \frac{\|g_{\mu\nu}\|_{L^\infty}}{n}. \tag{57}$$

(Here and for the remainder of the proof, implicit constants are permitted to depend on t and t_1.) The first term in (56) comes from the ends of the interval. To pass to (57), we used (44):

$$\mathbb{E}\big\{ \|\vec{\psi}_{s+1}^{(n)} - \vec{\psi}_s^{(n)}\|^2 \big\} \lesssim \tfrac{1}{n}.$$

From the inequality immediately above we have

$$\mathbb{E}\left\{ \big|\Psi_n(t) - \vec{\psi}_m^{(n)}\big|^2 \right\} + \mathbb{E}\left\{ \big|\Psi_n(t_1) - \vec{\psi}_k^{(n)}\big|^2 \right\} \lesssim \tfrac{1}{n}.$$

Combining this with the estimate above for (55), gives

$$\mathbb{E}\big\{ \big|M_n(t) - M_n(t_1) - \tilde{M}_n(m,k) + f(\vec{\psi}_k^{(n)})\big| \big\} \to 0 \tag{58}$$

as $n \to \infty$. In this way, we have reduced the proof to showing that

$$\lim_{n\to\infty} \mathbb{E}\big\{ G\big(\Psi_n|_{[t_0,t_1]}\big)\big[\tilde{M}_n(m,k) - f(\vec{\psi}_k^{(n)})\big] \big\} = 0.$$

Now $G\big(\Psi_n|_{[t_0,t_1]}\big)$ is a measurable function of $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$, while $\tilde{M}(m,k)$ is a function of $\vec{\psi}_k^{(n)}, \gamma_k, \ldots, \gamma_{m-1}$. Thus the statement above follows from

$$\lim_{n\to\infty} \mathbb{E}\big\{ \tilde{M}_n(m,k) - f(\vec{\psi}_k^{(n)}) \big| \vec{\psi}_k^{(n)} \big\} = 0,$$

which is a consequence of Lemma 4.3. $\qquad\square$

Naturally, the key to proving part (ii) of Theorem 1.7 is to show that (some subset of) these properties uniquely determines the process $\Psi(t)$. For this not only shows that the limit exists without passing to a subsequence, but also that the limit depends only on the assumptions (40) and (41), not on the specifics of the distribution.

As described in [49, §8.1], for $t_0 > 0$, any solution of the martingale problem (53) on $[t_0, \infty)$ gives rise to a solution of the system of stochastic differential equations

$$d\Psi_\mu = x_\mu\,dt + \tfrac{2}{\sqrt{\beta t}} \operatorname{Im}\big\{ \big[e^{i\Psi_\mu} - 1\big]\big[dB_1(t) + idB_2(t)\big] \big\} \tag{59}$$

with independent Brownian motions B_1 and B_2. The restriction to $[t_0, \infty)$ is necessary because we only know this result for SDEs with bounded coefficients. As the same probability distribution gives a solution of (59) for all $t_0 > 0$, we can speak of this as a solution of (59) on the *open* interval $(0, \infty)$.

Proposition 4.5. *For each $\vec{x} \in \mathbb{R}^p$, there is exactly one non-anticipating solution of* (59) *on* $(0,\infty)$ *that obeys* $x_\mu \Psi_\mu(t) \geq 0$ *for all* $t > 0$ *and*

$$\mathbb{E}\{\Psi_\mu(t)\} = x_\mu t. \tag{60}$$

Moreover, it is a strong solution, that is, Ψ is measurable with respect to the filtration generated by the Brownian motions (cf. [26, §5.2]).

Proof. Existence of a solution follows from Proposition 4.4 and [49, §8.1] or [26, §5.4.B]. To prove the other claims, we will construct a sequence of processes $\Phi_n(t)$ on the same probability space so that Φ_n is a measurable (and non-anticipating) with respect to the filtration generated by the Brownian motions and so that

$$\sup\{|\Phi_n(t) - \Psi(t)| : n^{-1} < t < n\} \to 0 \tag{61}$$

in probability as $n \to \infty$.

By the standard existence and uniqueness theory for SDE with Lipshitz coefficients (e.g. [49, §5.1] or [26, §5.2.B]), there is a unique solution $\Phi_n(t)$ of (59) on the interval $[n^{-1}, \infty)$ with initial data $\Phi_{n,\mu}(n^{-1}) = x_\mu n^{-1}$; moreover, Φ_n has the measurability properties just described and $\mathbb{E}\{\Phi_{n,\mu}(t)\} = x_\mu t$.

The process $\Psi(t) - \Phi_n(t)$ is a martingale for $t \in [n^{-1}, \infty)$. By considering separately the cases when $\Psi(n^{-1}) - \Phi_n(n^{-1})$ is non-negative or non-positive and using the fact that solutions of a SDE cannot cross (cf. [26, p. 293]), it follows that

$$\mathbb{E}\{|\Psi_\mu(n) - \Phi_{n,\mu}(n)|\} = \mathbb{E}\{|\Psi_\mu(n^{-1}) - \Phi_{n,\mu}(n^{-1})|\}.$$

Using the martingale property again and (60) we obtain

$$\mathbb{P}\left\{\sup_{n^{-1}<t<n} |\Psi(t) - \Psi_n(t)| > \epsilon\right\} \lesssim \frac{1}{\epsilon} \mathbb{E}\{|\Psi_\mu(n^{-1})| + |\Phi_{n,\mu}(n^{-1})|\} \lesssim \frac{1}{n\epsilon},$$

which proves the claim. $\qquad\square$

Remark. Elementary manipulations of (59) allow one to deduce the following two properties: (i) For $\lambda > 0$, $\Psi(\lambda t; x)$ and $\Psi(t; \lambda x)$ follow the same law. (ii) For any fixed $a \in \mathbb{R}$, $\Psi(x,t) - \Psi(a,t)$ and $\Psi(x - a, t)$ have the same distribution.

Theorem 4.6. *For any system of parameters obeying* $\mathbb{E}\left\{|\alpha_k|^2\right\} = \frac{2}{\beta(k+1)} + o(k^{-1})$ *and* $\mathbb{E}\left\{\log^2[1 - |\alpha_k|^2]\right\} = o(k^{-1})$ *the eigenvalue process converges in distribution; moreover, the limit depends only on β.*

Proof. By Proposition 4.5, there is a unique process with the properties described in Proposition 4.4. Thus all subsequential limits of $\Psi_n(t)$ are identical and so the limit exists without passing to a subsequence. Moreover, the interpretation given in Proposition 4.5 shows that the limiting distributions are consistent under extension of the vector $\vec{x}$ (in the sense of Kolmogorov's Theorem) and so may be regarded as the marginals of a probability measure defining a non-decreasing-function-valued process, $t \mapsto \Psi(x,t)$. More importantly, Proposition 4.5 shows that the distribution of the limit, $\Psi(x,t)$, depends only on β, not on the specifics of the distribution of the Verblunsky coefficients, $\{\alpha_k\}$.

Tracing our definitions back, we see that $\psi_{n-1}(x/n)$ converges in distribution to $\Psi(x, t = 1)$ at least for any finite collection points $x \in \mathbb{R}$. Therefore, by Proposition 2.5, the eigenvalue process converges to that associated to $\Psi(x, t = 1)$ in the sense of Definition 1.10. In view of the exactly soluble example described in Theorem 1.7, we feel justified in calling this the β-ensemble. $\qquad\square$

5. Slow decay

In this section we prove part (iii) of Theorem 1.7. That is, we prove that in the $n \to \infty$ limit, the re-scaled eigenvalues behave as a Poisson point process under the assumptions that the Verblunsky coefficients are independent, have rotationally invariant laws,

$$\mathbb{E}\big\{(1 - |\alpha_k|^2)^{-s_0}\big\} \lesssim 1 \tag{62}$$

for some $s_0 > 0$, and

$$\mathbb{E}\big\{|\alpha_k|^2\big\} \gtrsim (1 + k)^{\epsilon - 1}. \tag{63}$$

As in the previous sections, the constants implicit in the $\lesssim$ notation are permitted to depend on the constants in these inequalities, s_0 and ϵ included.

For the most part, one may apply the methods of [47, 48], which prove this result when the Verblunsky coefficients α_k are identically distributed (and so non-decaying). The approach taken there is, in turn, inspired by works of Molchanov, [34], and Minami, [33], on eigenvalue statistics for the Anderson model in the localized regime. To give full details of this argument would consume considerable space, much of which would more or less reproduce earlier published work with only a few minor changes. For this reason, we will focus on the one part of the argument (see the discussion after Lemma 5.1) where the analysis in [47, 48] is not applicable.

We believe that under the assumptions set out above, $\psi_{n-1}(\frac{x}{n})$ converges to 2π times a Poisson jump process. Further, it appears that in the natural scaling limit with x fixed, the process $\psi_k(\frac{x}{n})$ also converges to (2π times) a Poisson jump process (cf. [28]). It would be nice to deduce these facts directly from the recurrence equation; we have not done this, but hope that someone will pursue these matters. This approach to the problem would be closer to the original approach to the one dimensional Anderson model, [34], although estimates on the size of the eigenfunctions are used even there.

The basis of both Mochanov's and Minami's work is a limit law for a Poisson process—the analogue of the convergence of the binomial distribution to a Poisson distribution. Specifically, suppose that for each n we have independent point processes $d\eta_{j,n}$ for $1 \le j \le n$ and let us form the union, $d\xi_n := \sum_j d\eta_{j,n}$. Then $d\xi_n$ converges to the Poisson process with intensity $d\nu$ provided

(a) $d\eta_{j,n}$ is a null array, that is, $\sup_j \mathbb{P}\{\eta_{j,n}(I) > 0\} \to 0$ as $n \to \infty$,
(b) $\sum_j \mathbb{P}\{\eta_{j,n}(I) \ge 2\} \to 0$ as $n \to \infty$, and
(c) $\sum_j \mathbb{P}\{\eta_{j,n}(I) = 1\} \to \nu(I)$ as $n \to \infty$,

for each compact set I. This result is Corollary 7.5 in [25] and can be found in most books on point processes.

This limit theorem appears here in the following manner. Let us break the $n \times n$ CMV matrix $\mathcal{C}^{(n)}$ into $\sim \log[n]$ blocks of equal length $N \sim n/\log[n]$ by introducing independent random unitary boundary conditions (just as we did in the Introduction in order to obtain an $n \times n$ matrix from the originally semi-infinite one). In this way, we obtain the direct sum of statistically independent matrices, so the eigenvalue process is the union of $\sim n/\log[n]$ independent point processes. This leaves us with two jobs: to check the hypotheses of the theorem above (this is quick) and to show that the introduction of these boundary conditions has not significantly affected the eigenvalue statistics (this is anything but quick).

By choosing the boundary conditions uniformly from the circle, the distribution of the eigenvalues of each $N \times N$ block is rotationally invariant. This resolves the hypothesis (a) above and will suffice to justify (c) once we prove (b). To verify (b) we invoke the following result, which is Theorem 6.3 in [48]:

Lemma 5.1. *Given any CMV matrix of size $N \times N$ with Verblunsky coefficients that are independent and follow rotationally invariant laws, one has*

$$\mathbb{P}(\Omega) \le \tfrac{1}{2} L^2 N^2, \tag{64}$$

where Ω is the event that two or more eigenvalues belong to a fixed arc of length $2\pi L$. (Note that the expected number of eigenvalues in this arc is LN.)

This settles the first of our two jobs. The second is to explain why the introduction of boundary conditions does not completely rearrange the eigenvalues. The reason is simple: the eigenfunctions are very strongly localized with sharply decaying tails. Proving this fact is not so simple; however, there is a wealth of work on the Anderson model (i.e., for related operators with non-decaying random potentials) that provides guidance. We will be studying fractional moments of the Green's function, which is an idea of [1]; while our approach to estimating the transfer matrix is reminiscent of [31, 39]. The argument culminates in Corollary 5.6.

The steps leading to Corollary 5.6 are the one place where we cannot mimic the treatment of the i.i.d. case in [47, 48]. In short, the method used there is not sufficiently quantitative to be effective in our setting. More precisely, it relies on positivity of the Lyapunov exponent, which is defined as an infinite volume limit (without a rate of convergence). Recall that the Lyapunov exponent $\gamma \ge 0$ governs the exponential behaviour of formal eigenfunctions of $\mathcal{C}$. This is a serious difficulty here since the effective Lyapunov exponent associated to $\mathcal{C}^{(n)}$ decays to zero as $n \to \infty$. The saving grace for us, however, will be that it converges to zero no faster than $cn^{\epsilon-1}$; see (86).

For most of the remainder of this section, that is, until we reach the statement of Corollary 5.6, we will be dealing primarily with the full semi-infinite CMV matrix, $\mathcal{C}$.

The Szegő recurrence for the orthonormal polynomials can be encoded with a matrix:

$$A_k = \frac{1}{\sqrt{1 - |\alpha_k|^2}} \begin{bmatrix} z & -\bar{\alpha}_k \\ -\alpha_k z & 1 \end{bmatrix}.$$

For $l \ge k$, we define the transfer matrices

$$T(l, k) = A_{l-1} A_{l-2} \cdots A_{k+1} A_k,$$

with $T(k, k) = \text{Id}$. Notice that both A_k and $T(l, k)$ depend on $z \in \mathbb{C}$; however we suppress this in the notation, since invariably z will be fixed.

Lemma 5.2. *Let α_k follow a rotationally invariant law, then*

$$\mathbb{E}\{\|A_k v\|^{-1}\} \le \|v\|^{-1} \mathbb{E}\{1 - \tfrac{1}{4}|\alpha_k|^2\} \tag{65}$$

for any non-zero vector $v \in \mathbb{C}^2$ and any $|z| = 1$.

Proof. Clearly it suffices to prove this result when $v = (x, y)$ is a unit vector. In this case, a little computation shows

$$\|A_k v\|^2 = \frac{1 - 4\,\text{Re}(x\bar{y}z\alpha_k) + |\alpha_k|^2}{1 - |\alpha_k|^2}.$$

The lemma now degenerates to the estimation of an integral. We will use following notation: $r = |\alpha_k|$, $R = 2|x\bar{y}|$, $\theta = \arg(x\bar{y}\alpha_k)$. Note that $R, r \in [0, 1]$, in the former case, because $|x|^2 + |y|^2 = 1$.

$$\int_0^{2\pi} \frac{\sqrt{1-r^2}}{\sqrt{1-2Rr\cos(\theta)+r^2}} \frac{d\theta}{2\pi} \leq \frac{2}{\pi} \int_0^{2\pi} \frac{\sqrt{1-r^2}}{\sqrt{1-2r\cos(\theta)+r^2}} \frac{d\theta}{2\pi} \tag{66}$$

$$= \frac{2}{\pi} \frac{\sqrt{1-r^2}}{1+r} \mathbf{K}\left(\frac{2\sqrt{r}}{1+r}\right) \leq 1 - \tfrac{1}{4}r^2, \tag{67}$$

where $\mathbf{K}(k)$ is the complete elliptic integral of modulus k. The first inequality here follows from the fact that this integral is an increasing function of R; to see this differentiate and compare the resulting integrand at θ and $\theta + \pi$. To obtain the second inequality with the factor $\frac{1}{4}$ takes some trouble; however, the fact that it is true with some small coefficient follows easily by studying the behaviour as $r \downarrow 0$. This simpler fact is actually sufficient for our purposes. $\qquad \square$

By applying this inductively, we obtain results for the transfer matrices:

Lemma 5.3. *Suppose the Verblunsky coefficients are independent and follow rotationally invariant laws. Given $0 \leq s \leq 1$ and $|z| = 1$,*

$$\mathbb{E}\left\{\|T(l,k)\|^{-s}\right\} \leq \exp\left[-\tfrac{s}{4} \sum_{j=k}^{l-1} \mathbb{E}\{|\alpha_j|^2\}\right] \tag{68}$$

for any $k \leq l$. More generally,

$$\mathbb{E}\left\{\frac{\|T(l,k)\|^s}{\|T(r,k)\|^s}\right\} \leq \exp\left[-\tfrac{s}{4} \sum_{j=k}^{r-1} \mathbb{E}\{|\alpha_j|^2\}\right] \tag{69}$$

for $k \leq l \leq r$.

Proof. Let us first remark that (68) is the special case of (69) when $k = l$. We will prove a slightly more general result:

$$\mathbb{E}\left\{\|T(r,l)v\|^{-s}\|v\|^s\right\} \leq \exp\left[-\tfrac{s}{4} \sum_{j=l}^{r-1} \mathbb{E}\{|\alpha_j|^2\}\right] \tag{70}$$

where v is a random unit vector that is statistically independent of α_j for $l \leq j \leq r-1$. Note that the result for general s will follow from $s = 1$ by Jensen's inequality.

Clearly (70) is well formulated for proof by induction. When $r = l$ one has equality. Assuming the result for $T(r, l)$, we have

$$\mathbb{E}\left\{\|T(r+1,l)v\|^{-1}\|v\|^s\right\} = \mathbb{E}\left\{\|A_r T(r,l)v\|^{-1}\|v\|^s\right\} \tag{71}$$

$$\leq \mathbb{E}\left\{1 - \tfrac{1}{4}|\alpha_k|^2\right\} \mathbb{E}\left\{\|T(r,l)v\|^{-1}\|v\|^s\right\} \tag{72}$$

by Lemma 5.2. To finish the proof, notice that $1 - \frac{x}{4} \leq e^{-x/4}$ for $x \in [0, \infty)$. $\quad \square$

Proposition 5.4. *Fix $0 < s \leq \min\{\tfrac{1}{2}s_0, \tfrac{1}{4}\}$, where s_0 is as in (63). With probability one, there exists a (random) unit vector v so that $T(l,k)v \to 0$ as $l \to \infty$; moreover, there exists $c > 0$ so that*

$$\mathbb{E}\left\{\|T(l,k)v\|^s\right\} \lesssim (k+1)\exp\left[-c(l^\epsilon - k^\epsilon)\right] \tag{73}$$

uniformly for $0 \leq k \leq l$ and $|z| = 1$.

Proof. We simply take expectations in the proof of the Osceledec Theorem. This theorem is covered in many textbooks, for example, [42, §10.5], which can be consulted for more detail on what follows.

Let P_l denote the orthogonal projection onto the contracting subspace of $T(l, k) \in SL(2, \mathbb{C})$; that is, the subspace where $\|Tu\| = \|T\|^{-1}\|u\|$. (This is one dimensional when $\|T\| > 1$.) Then

$$\|P_{l+1} - P_l\| \leq \frac{\|A_l\|^2}{\|T(l, k)\|^2}. \tag{74}$$

This is a ratio of independent random variables. As $\|A_l\|^2 \leq 2(1 - |\alpha_l|^2)^{-1}$, it follows from (62) that $\mathbb{E}\{\|A_l\|^{2s}\} \lesssim 1$. (This implicit constant depends on s as do many more below.) Therefore, by applying Lemma 5.3 and (63) we see that there exists $c > 0$ so that

$$\mathbb{E}\{\|P_{l+1} - P_l\|^s\} \lesssim \exp\left[-\frac{s}{2}\sum_{j=k}^{l-1}\mathbb{E}\{|\alpha_j|^2\}\right] \lesssim \exp\left[-c(l^\epsilon - k^\epsilon)\right]$$

for any $s \leq \min\{s_0, \frac{1}{2}\}$. This implies that $P_{l+1} - P_l$ is almost surely summable and so $P_\infty := \lim_{l \to \infty} P_l$ exists.

Let v be a unit vector from the range of P_∞. We must now study $\|T(l, k)v\|$. Using only elementary facts, we have

$$\mathbb{E}\{\|T(l, k)v\|^s\} = \mathbb{E}\{\|T(l, k)P_\infty\|^s\} \tag{75}$$

$$\leq \mathbb{E}\{\|T(l, k)P_l\|^s\} + \sum_{r=l}^{\infty}\mathbb{E}\{\|T(l, k)\|^s\|P_{r+1} - P_r\|^s\} \tag{76}$$

$$\leq \mathbb{E}\{\|T(l, k)\|^{-s}\} + \sum_{r=l}^{\infty}\mathbb{E}\{\|T(l, k)\|^s\|P_{r+1} - P_r\|^s\}. \tag{77}$$

By Lemma 5.3 and (63),

$$\mathbb{E}\{\|T(l, k)\|^{-s}\} \lesssim \exp\left[-\frac{s}{4}\sum_{j=k}^{l-1}\mathbb{E}\{|\alpha_j|^2\}\right] \lesssim \exp\left[-c(l^\epsilon - k^\epsilon)\right]$$

for some $c > 0$. Next we deal with the sum: using (74), the Cauchy–Schwarz inequality, and the independence of the Verblunsky coefficients, we have

$$\mathbb{E}\{\|T(l, k)\|^s\|P_{r+1} - P_r\|^s\}^2 \tag{78}$$

$$\leq \mathbb{E}\{\|A_r\|^{4s}\}\,\mathbb{E}\{\|T(r, k)\|^{-2s}\}\,\mathbb{E}\left\{\frac{\|T(l, k)\|^{2s}}{\|T(r, k)\|^{2s}}\right\}. \tag{79}$$

The first factor is uniformly bounded by (62). The second and third factors can be estimated via Lemma 5.3. In this way, we obtain

$$\mathbb{E}\{\|T(l, k)\|^s\|P_{r+1} - P_r\|^s\} \lesssim \exp\left[-c(2r^\epsilon - l^\epsilon - k^\epsilon)\right]$$

for some $c > 0$. This is clearly summable in r; indeed

$$\sum_{r=l}^{\infty}\exp\left[-2cr^\epsilon\right] \lesssim (l+1)^{1-\epsilon}\exp\left[-2cl^\epsilon\right].$$

Consequently,

$$\sum_{r=l}^{\infty} \mathbb{E}\big\{\|T(l,k)\|^s \|P_{r+1} - P_r\|^s\big\} \lesssim (l+1)\exp\big[-cs(l^\epsilon - k^\epsilon)\big].$$

Collecting all the estimates we have derived, we obtain the existence of a $c > 0$ so that

$$\mathbb{E}\big\{\|T(l,k)v\|^s\big\} \lesssim (l+1)\exp\big[-c(l^\epsilon - k^\epsilon)\big].$$

By reducing c a little bit, one may replace the polynomial factor by $(k+1)$ as written in the statement. $\qquad\square$

It is presumably possible to obtain much more precise estimates for the decay of $\|T(l,k)v\|$ using the methods from [29] (cf. [42, §12.7]). While their techniques are strong, they pursue only the almost sure behaviour, which is why we dragged the reader through the mud here.

In order to proceed, we need to upgrade Proposition 5.4 to an estimate on the Green's function (i.e., the matrix elements of the resolvent),

$$G_{kl}(z) := (\mathcal{C} - z)_{kl}^{-1},$$

for $|z| = 1$. From [41, §4.4], we learn the following:

$$G_{kl}(z) = \begin{cases} (2z)^{-1}\chi_l(z)p_k(z) & k > l \text{ or } k = l \text{ is odd} \\ (2z)^{-1}x_k(z)\pi_l(z) & l > k \text{ or } k = l \text{ is even.} \end{cases} \tag{80}$$

Here

$$\begin{bmatrix} p_{2l} \\ \pi_{2l} \end{bmatrix} = z^{-l}T(2l+1,0)\begin{bmatrix} F(z)+1 \\ F(z)-1 \end{bmatrix}, \quad \begin{bmatrix} z^{-1}\pi_{2l-1} \\ p_{2l-1} \end{bmatrix} = z^{-l}T(2l,0)\begin{bmatrix} F(z)+1 \\ F(z)-1 \end{bmatrix},$$

$$\begin{bmatrix} x_{2l} \\ \chi_{2l} \end{bmatrix} = z^{-l}T(2l+1,0)\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ and } \begin{bmatrix} z^{-1}\chi_{2l-1} \\ x_{2l-1} \end{bmatrix} = z^{-l}T(2l,0)\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where $F(z)$ is the Carathéodory function given in (11). Note that $x_k(z)$ is the complex conjugate of $\chi_k(1/\bar{z})$; in particular, these two sequences have the same moduli when z lies on the unit circle.

For $|z| = 1$, $F(z)$ is purely imaginary with probability one because $\mathcal{C}$ has purely point spectrum with probability one. (For the latter fact, see [42, Theorem 12.7.5] or combine the Simon–Wolff criterion, [42, Theorem 10.2.1], with Proposition 5.4.) This implies that $p_k(z)$ and $\pi_k(z)$ are complex conjugates for such z.

The way we have introduced p_k and π_k belies their true significance; they are the Weyl solutions—they are proportional to the decaying solutions (whenever such solutions exist). In this way we find that for $|z| = 1$ and $k \le l$,

$$\left|\frac{p_l(z)}{p_k(z)}\right| = \left|\frac{\pi_l(z)}{\pi_k(z)}\right| = \|T(l,k)v\| \tag{81}$$

with probability one, provided v is chosen as in Proposition 5.4.

Corollary 5.5. *Fix $0 < s < 1$. There exists $c > 0$ so that*

$$\mathbb{E}\big\{\big|G_{kl}(z)\big|^s + \big|G_{lk}(z)\big|^s\big\} \lesssim (k+1)\exp\big[-c(l^\epsilon - k^\epsilon)\big] \tag{82}$$

uniformly for $0 \le k \le l$ and $|z| = 1$. Moreover, the same estimate holds uniformly in $|z| \le 1$ for a reduced value of c.

Proof. By employing Kolmogorov's bound on Cauchy integrals, [14, Theorem 4.2], as in [47, Lemma 4.1.1] or [48, Lemma 3.1], we see that

$$\mathbb{E}\big\{\big|G_{kl}(z)\big|^s\big\} \leq 4\sec(s\pi/2), \tag{83}$$

for any choice of k, l and $|z| \leq 1$. The key idea is that rotational invariance, (1), allows us to control the expectation on the left-hand side by integration in z.

The inequality (83) shows us that it is sufficient to prove the result for s small; specifically, Hölder's inequality implies

$$\mathbb{E}\big\{\big|G_{kl}(z)\big|^s\big\} \leq \mathbb{E}\big\{\big|G_{kl}(z)\big|^{s_1}\big\}^{\frac{1-s}{1+s-2s_1}} \mathbb{E}\big\{\big|G_{kl}(z)\big|^{\frac{1+s}{2}}\big\}^{\frac{2s-2s_1}{1+s-2s_1}},$$

which allows us to deduce the result with power $s > s_1$ from that with power s_1.

Let us begin with $|z| = 1$. From the discussion preceding this corollary, we have

$$\left|\frac{G_{kl}(z)}{G_{kk}(z)}\right| = \left|\frac{G_{lk}(z)}{G_{kk}(z)}\right| = \|T(l,k)v\| \tag{84}$$

and so by the Cauchy–Schwartz inequality,

$$\mathbb{E}\big\{\big|G_{kl}(z)\big|^s + \big|G_{lk}(z)\big|^s\big\} \leq 2\,\mathbb{E}\big\{\|T(l,k)v\|^{2s}\big\}\,\mathbb{E}\big\{|G_{kk}(z)|^{2s}\big\}. \tag{85}$$

The first factor here was estimated in Proposition 5.4, while the second can be bounded using (83). This proves (82).

To extend the result to $z \in \mathbb{D}$, one may apply the argument from Theorem 4.2 in [2]. □

Corollary 5.6. *Fix* $0 < s < 1$. *There exists* $c > 0$ *and* $A > 0$ *so that*

$$\mathbb{E}\big\{\big|(\mathcal{C}^{(n)} - z)_{kl}^{-1}\big|^s\big\} \leq A\,n\exp\big[-cn^{\epsilon-1}|l-k|\big] \tag{86}$$

for all $|z| \leq 1$ *and* n.

Proof. This result for finite matrices follows from the previous corollary by applying the resolvent identity as in [47, Lemma 4.2.3] or [48, Lemma 3.4]. To make the final result look cleaner, we employed the estimate

$$|l^\epsilon - k^\epsilon| \geq \epsilon n^{\epsilon-1}|k-l|, \quad \forall\, k, l \in [0,n],$$

which follows from the lower bound on the derivative of $x \mapsto x^\epsilon$. □

With Corollary 5.6 as input, one may now simply repeat the analysis in Section 5 of [48] to see that breaking $\mathcal{C}^{(n)}$ into blocks does not significantly alter the eigenvalue process. Indeed, only one change to the argument is required: the constant D_0 in that paper needs to be replaced with $cn^{\epsilon-1}$. With this second job completed, no obstacles remain to proving Theorem 1.7(iii) in the manner outlined at the beginning of this section.

Incidentally, Corollary 5.6 is also sufficient to prove Poisson statistics for the locations of the eigenfunctions as in [28].

REFERENCES

[1] M. Aizenman and S. Molchanov, Localization at large disorder and at extreme energies: an elementary derivation. *Comm. Math. Phys.* **157** (1993), 245–278.

[2] M. Aizenman, J. Schenker, R. Friedrich, and D. Hundertmark, Finite-volume fractional moment criteria for Anderson localization. *Commun. Math. Phys.* **224** (2001), 219–253.

[3] T. M. Apostol, *Calculus. Vol. II. 2nd Ed.* John Wiley & Sons, 1969.

[4] M. V. Berry, Quantum Chaology. *Proc. Roy. Soc. Lond. A* **413** (1987), 183–198.

[5] O. Bohigas, M. Giannoni, and C. Schmit, Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.* **52** (1984), 1–4.

[6] M. J. Cantero, L. Moral, and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.* **362** (2003), 29–56.

[7] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations.* McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.

[8] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach.* New York Univ., Courant Inst. Math. Sci., New York, 1999.

[9] P. Deift, Universality for mathematical and physical systems. *International Congress of Mathematicians. Vol. I,* 125–152, Eur. Math. Soc., Zürich, 2007.

[10] F. Delyon, Appearance of purely singular continuous spectrum in a class of random Schrödinger operators. *J. Statist. Phys.* **40** (1985) 621–630.

[11] F. Delyon, B. Simon, and B. Souillard, From power pure point to continuous spectrum in disordered systems. *Ann. Inst. H. Poincaré* **42** (1985) 283–309.

[12] S. Denisov, *Continuous Analogs of Polynomials Orthogonal on the Unit Circle. Krein Systems.* Lecture notes. Available from the author's web page.

[13] I. Dumitriu and A. Edelman, Matrix models for beta ensembles. *J. Math. Phys.* **43** (2002), 5830–5847.

[14] P. L. Duren, *Theory of H^p spaces.* Academic Press, New York-London 1970.

[15] R. Durrett, *Probability: theory and examples.* Second edition. Duxbury Press, Belmont, CA, 1996.

[16] F. J. Dyson, Statistical theory of the energy levels of complex systems. I, II, and III. *J. Math. Phys.* **3** (1962), 140–156, 157–165, and 166–175.

[17] F. J. Dyson, The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics. *J. Math. Phys.* **3** (1962), 1199–1215.

[18] A. Edelman and B. D. Sutton, From Random Matrices to Stochastic Operators. Preprint `math-ph/0607038`.

[19] P. J. Forrester, *Log-gases and Random matrices.* Available from the author's web page.

[20] P. J. Forrester and E. M. Rains, Interpretations of some parameter dependent generalizations of classical matrix ensembles. Preprint `math-ph/0211042`.

[21] Ya. L. Geronimus, On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathéodory and Schur functions. *Mat. Sbornik* **15(57)** (1944), 99–130.

[22] Ya. L. Geronimus, *Orthogonal Polynomials.* Consultants Bureau, New York 1961.

[23] I. J. Good, Short proof of a conjecture by Dyson. *J. Math. Phys.* **11** (1970), 1884.

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

[25] O. Kallenberg, *Random measures.* Academic Press. 4th. ed. 1986.

[26] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus.* Second edition. Graduate Texts in Mathematics, 113. Springer-Verlag, New York, 1991.

[27] R. Killip and I. Nenciu, Matrix models for circular ensembles. *Int. Math. Res. Not.* **2004**, 2665–2701.

[28] R. Killip and F. Nakano, Eigenfunction statistics in the localized Anderson model. *Ann. Henri Poincaré* **8** (2007), 27–36.

[29] A. Kiselev, Y. Last, and B. Simon, Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators. *Comm. Math. Phys.* **194** (1998), 1–45.

[30] S. Kotani and N. Ushiroya, One-dimensional Schrödinger operators with random decaying potentials. *Comm. Math. Phys.* **115** (1988), 247–266.

[31] H. Kunz and B. Souillard, Sur le spectre des opérateurs aux différences finies aléatoires. *Comm. Math. Phys.* **78** (1980/81), 201–246.

[32] M. L. Mehta, *Random matrices. Third Edition.* Pure and Applied Mathematics (Amsterdam), 142. Elsevier/Academic Press, Amsterdam, 2004.

[33] N. Minami, Local fluctuation of the spectrum of a multidimensional Anderson tight-binding model. *Comm. Math. Phys.* **177** (1996), 709–725.

[34] S. Molchanov, The local structure of the spectrum of the one-dimensional Schrödinger operator. *Comm. Math. Phys.* **78** (1981), 429–446.

[35] A. Okounkov, On n-point correlations in the log-gas at rational temperature. Preprint `hep-th/9702001`.

[36] C. E. Porter, Fluctuations of Quantal Spectra. *Statistical Theories of Spectra: Fluctuations,* 2–87, Academic Press, 1965.

[37] S. I. Resnick, *Extreme values, regular variation, and point processes.* Applied Probability. A Series of the Applied Probability Trust, 4. Springer-Verlag, New York, 1987.

[38] J. A. Ramírez, B. Rider, and B. Virág, Beta ensembles, stochastic Airy spectrum, and a diffusion. Preprint `math.PR/0607331`.

[39] C. Shubin, R. Vakilian, and T. Wolff, Some harmonic analysis questions suggested by Anderson-Bernoulli models. *Geom. Funct. Anal.* **8** (1998), 932–964.

[40] B. Simon, Some Jacobi matrices with decaying potential and dense point spectrum. *Comm. Math. Phys.* **87** (1982), 253–258.

[41] B. Simon, *Orthogonal Polynomials on the Unit Circle, Vol. 1.* American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, Rhode Island, 2004.

[42] B. Simon, *Orthogonal Polynomials on the Unit Circle, Vol. 2.* American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, Rhode Island, 2004.

[43] B. Simon, Fine structure of the zeros of orthogonal polynomials, I. A tale of two pictures. To appear in Proc. Constructive Functions Tech-04.

[44] B. Simon, CMV matrices: five years after. To appear in the Proceedings of the W.D. Evans 65th Birthday Conference.

[45] U. Smilansky and P. J. Forrester, Random discrete Schrödinger operators from Random Matrix Theory. Preprint `math-ph/0507036`.

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

[47] M. Stoiciu, *Zeros of random orthogonal polynomials on the unit circle.* Ph.D. thesis, Caltech, 2005. Available online from the Caltech library.

[48] M. Stoiciu, The statistical distribution of the zeros of random paraorthogonal polynomials on the unit circle. *J. Approx. Theory* **139** (2006), 29–64.

[49] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes.* Grundlehren der Mathematischen Wissenschaften, 233. Springer-Verlag, Berlin-New York, 1979.

[50] G. Szegő, *Orthogonal Polynomials.* American Mathematical Society Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, Rhode Island, 1975.

[51] B. Valkó and B. Virág, Continuum limits of random matrices and the Brownian carousel. Preprint `arXiv:0712.2000`.

[52] D. Watkins, Some perspectives on the eigenvalue problem. *SIAM Rev.* **35** (1993), 430–471.

[53] K. Wilson, Proof of a conjecture by Dyson. *J. Math. Phys.* **3** (1962), 1040–1043.

ROWAN KILLIP, UCLA MATHEMATICS DEPARTMENT, BOX 951555, LOS ANGELES, CA 90095
E-mail address: `killip@math.ucla.edu`

MIHAI STOICIU, DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267
E-mail address: `mstoiciu@williams.edu`