Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2013, Vol. 49, No. 4, 934–960
DOI: 10.1214/12-AIHP501

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Perturbed Toeplitz operators and radial determinantal processes

Torsten Ehrhardt[a] and Brian Rider[b]

[a] *Department of Mathematics, University of California, Santa Cruz, CA 95064, USA. E-mail: ehrhardt@math.ucsc.edu*
[b] *Department of Mathematics, University of Colorado at Boulder, Boulder, CO 80309, USA. E-mail: brian.rider@colorado.edu*

Received 13 February 2011; revised 11 February 2012; accepted 10 June 2012

Abstract. We study a class of rotation invariant determinantal ensembles in the complex plane; examples include the eigenvalues of Gaussian random matrices and the roots of certain families of random polynomials. The main result is a criterion for a central limit theorem to hold for angular statistics of the points. The proof exploits an exact formula relating the generating function of such statistics to the determinant of a perturbed Toeplitz matrix.

Résumé. Nous étudions une classe d'ensembles déterminantaux dans le plan complexe invariants par rotation; cette classe comprend les cas des valeurs propres de matrices gaussiennes aléatoires et des zéros de certaines familles de polynomes aléatoires. Le résultat principal est un critère pour l'existence d'un théorème de la limite centrale pour la statistique des angles entre les points. La preuve utilise une formule exacte reliant la fonction génératrice de telles statistiques au déterminant d'une matrice de Toeplitz perturbée.

MSC: 60B20; 60F05; 47B35

Keywords: Random matrices; Determinantal processes; Toeplitz operators; Szegö–Widom limit theorem

1. Introduction

Consider the probability measure on n complex points, $z_1, \ldots, z_n \in \mathbb{C}$, defined by

$$\mathbb{P}_{\mathfrak{m},n}(z_1, \ldots, z_n) = \frac{1}{Z_{\mathfrak{m},n}} \prod_{j<k} |z_j - z_k|^2 \prod_{k=1}^{n} \mathrm{d}\mathfrak{m}(z_k), \tag{1}$$

with a (positive) reference measure $\mathfrak{m}$ on $\mathbb{C}$. This is an instance of a *determinantal ensemble*, so named as the presence of the Vandermonde interaction term $\prod |z_i - z_j|^2$ results in all k-fold ($k \le n$) correlations of the points being given by a determinant of a certain $k \times k$ Gramian. Determinantal ensembles as such were identified in the mathematical physics literature as a model of fermions [18], but also arise naturally in a number of contexts including random matrix theory. For background, [13] and [22] are recommended.

Throughout the paper we restrict to the situation of radially symmetric weights, $\mathrm{d}\mathfrak{m}(z) = \mathrm{d}\mu(r)\,\mathrm{d}\theta$, assuming that $\mathfrak{m}$ does not place positive mass at the origin. The standard examples in this set-up are the following:

Ginibre ensemble. Let M be an $n \times n$ random matrix in which each entry is an independent complex Gaussian of mean zero and mean-square one. Then the n eigenvalues have joint density (1) with $\mathrm{d}\mathfrak{m}(z) = \mathrm{e}^{-|z|^2}\,\mathrm{d}^2 z$ [10].

Circular Unitary Ensemble (CUE). Place Haar measure on n-dimensional unitary group $U(n)$ and consider again the eigenvalues. These points live on the unit circle $\mathbb{T} = \{t \in \mathbb{C}\colon |t| = 1\}$, and it is well known that their joint law is given by (1) in which $\mathfrak{m}$ is arc-length measure.

Truncated Bergman process. Start with the random polynomial $z^n + \sum_{k=0}^{n-1} a_k z^k$ with independent coefficients drawn uniformly from the disk of radius r in $\mathbb{C}$. Condition the roots $z_1, \dots, z_n$ to lie in the unit disk. Then, the $r \to \infty$ limit of the conditional root ensemble is (1) where now $\mathfrak{m}$ is the uniform measure on the disk of radius one. This nice fact may be found in [12]; for an explanation of the name see [19].

Our aim is to identify conditions on μ under which a central limit theorem (CLT) for the quantity

$$X_{f,n} = \sum_{k=1}^{n} f(\arg z_k)$$

holds or not. Certainly the regularity of the test function f matters as well. An enormous industry has grown up around CLT's for linear statistics in determinantal and random matrix ensembles. Despite rather than because of this, there are several reasons for making a special study of such "angular" statistics in the given setting.

The conventional wisdom is that choosing f sufficiently smooth produces Gaussian fluctuations with order one variance (i.e., as $n \to \infty$ the un-normalized $X_{f,n} - \mathbb{E}X_{f,n}$ should posses a CLT). This is borne out by a number of results pertaining to ensembles with symmetry and so real, or suitably "one-dimensional," spectra. In the present context in which points inhabit $\mathbb{C}$, [21] proves a result of this type for C^1 statistics of the Ginibre ensemble. See also [1] which represents the most recent refinement of a series of extensions of [21] to smooth statistics for Ginibre-like ensembles in which the Gaussian weight $\mathfrak{m}$ is replaced by a more general $e^{-V(z)}$ (these are the normal matrix models). On the other hand, a smooth function of $\arg z$ is not smooth in the variables $(x, y) = z$. Again back in the Ginibre ensemble, [20] shows that the variance of $X_{f,n}$ is of order $\log n$ (whenever f possesses an L^2-derivative), though is unable to establish a CLT. While there are a number of general results on CLT's for determinantal processes in whatever dimension, notably [23] which employs cumulants, the logarithmic growth in this case is not sufficiently fast for those conclusions to be relevant. We also mention that for any determinantal process on $\mathbb{C}$ with radially symmetric weight, the collections of moduli $|z_1|, |z_2|, \dots$ are independent; this is spelled out nicely in [13]. Hence, CLT's for "radial" statistics in our ensembles may be proved by checking the classical Lindenberg conditions, see [9] and [20] for details in the Ginibre case.

It is likely that the considerations of [21], which entail a refinement of the cumulant method, or those of [1] can be adopted to the matter at hand. Here though we take an operator-theoretic approach, based on the following formula. For any $\varphi \in L^\infty(\mathbb{T})$,

$$\mathbb{E}_{\mathfrak{m},n}\left[\prod_{k=1}^{n} \varphi(\arg z_k)\right] = \det M_{\mu,n}(\varphi), \quad M_{\mu,n}(\varphi) = (\varphi_{k-\ell}\varrho_{k,\ell})_{0 \le k,\ell \le n-1}, \tag{2}$$

where $\varphi_k = \frac{1}{2\pi}\int_0^{2\pi} \varphi(x) e^{-ikx}\,dx$, the kth Fourier coefficient of φ, and

$$\varrho_{k,\ell} = \frac{m_{k+\ell}}{(m_{2k} m_{2\ell})^{1/2}} \quad \text{in which } m_k = \int_0^\infty r^k\,d\mu(r), \tag{3}$$

the kth moment of the half-line measure μ. We identify φ defined on the unit circle $\mathbb{T}$ with the corresponding periodic function defined on $\mathbb{R}$. The brief derivation of (2) can be found in the Appendix.

This provides an explicit formula for the generating function of $X_{f,n}$ by the choice $\varphi = e^{i\lambda f}$. A CLT for $X_{f,n}$ will then follow from sufficiently sharp $n \to \infty$ asymptotics of the determinant on the right-hand side of (2). Of course, if this is to be the strategy we must henceforth assume that $m_k < \infty$ for all k.

In the case of CUE, all $m_k = 1$, and the identity (2) reduces to Weyl's formula relating the Haar average of a class function in $U(n)$ to a standard Toeplitz determinant. The strong Szegö limit theorem and its generalizations to symbols of weaker regularity then imply a variety of CLT's for linear spectral statistics in $U(n)$, see for instance [6,14] and [15] as well as the references therein. For more generic μ, what appears on the right-hand side of (2) is the Hadamard product of (truncated) Toeplitz and Hankel operators. While Hankel determinants arise as naturally as their Toeplitz counterparts in random matrix theory and several applications have prompted investigations of Toeplitz + Hankel forms (see for example [3]), the present problem is the first to our knowledge to motivate an asymptotic study of Toeplitz $\circ$ Hankel matrices. Though, as the title suggests, the analysis more closely follows the Toeplitz framework.

To describe the regularity assumed on the various test functions f, we introduce the function space $F\ell^p(\nu)$, $1 \le p < \infty$ (see [16]), comprised of all $f \in L^1(\mathbb{T})$ such that

$$\|f\|_{F\ell^p(\nu)} := \left(\sum_{n=-\infty}^{\infty} |f_n|^p \nu_n^p\right)^{1/p} < \infty. \tag{4}$$

Here $\nu = \{\nu_n\}_{n=-\infty}^{\infty}$ is a positive weight, and again $\{f_n\}$ are the Fourier coefficients of f. We will in particular deal with the cases $p = 1$ or $p = 2$, and power weights $\nu_n = (1 + |n|)^\sigma$, $\sigma \ge 0$. In the latter case we simply denote the space by $F\ell^p_\sigma$ and write $F\ell^p$ when $\sigma = 0$.

As for the underlying probability measure μ, a natural criterion arises on the second derivative of the logarithmic moment function.

Moment assumption. *The function*

$$\xi \mapsto m_\xi := \int_0^\infty r^\xi \, \mathrm{d}\mu(r), \quad \xi \ge 0, \tag{5}$$

satisfies one of the following two sets of conditions.

(C1) *or "$\beta > 1$": It holds*

$$(\ln m_\xi)'' = \mathrm{O}\big(\xi^{-\beta}\big), \quad \xi \to \infty \tag{6}$$

with $\beta > 1$.

(C2) *or "$1/2 < \beta \le 1$": It holds*

$$(\ln m_\xi)'' = h_\mu(\xi) + \mathrm{O}\big(\xi^{-\varrho}\big), \quad \xi \to \infty \tag{7}$$

for a differentiable function $h_\mu(\xi) \ge 0$, $\xi > 0$, such that

$$h_\mu(\xi) = \mathrm{O}\big(\xi^{-\beta}\big), \qquad h'_\mu(\xi) = \mathrm{O}\big(\xi^{-\gamma}\big), \quad \xi \to \infty \tag{8}$$

with $1/2 < \beta \le 1$, $\varrho, \gamma > 1$. Additionally,

$$\iota_\mu(x) := \frac{1}{2} \int_1^x h_\mu(\xi) \, \mathrm{d}\xi, \tag{9}$$

tends to infinity as $x \to \infty$.

Notice that since we have already assumed $m_k < \infty$ for all k, m_ξ is infinitely differentiable for positive ξ. The typical behavior we have in mind in both (C1) and (C2) are asymptotics such as

$$(\ln m_\xi)'' = \alpha \xi^{-\beta} + \mathrm{O}\big(\xi^{-\varrho}\big), \quad \xi \to \infty, \tag{10}$$

with $\alpha, \beta > 0$, $\varrho > \max\{1, \beta\}$ (and thus $\gamma = 1 + \beta$ in case (C2)). As examples, we remark that for Ginibre, $(\ln m_\xi)'' = \frac{1}{2}\xi^{-1} + \mathrm{O}(\xi^{-2})$, while both CUE and truncated Bergman satisfy $(\ln m_\xi)'' = \mathrm{O}(\xi^{-2})$. The transition from $\beta \le 1$ to $\beta > 1$ is particularly interesting; Section 2 discusses the moment conditions in greater detail. The restriction to $\beta > 1/2$ is tied to the method in which we show that $M_{\mu,n}$ is a small perturbation of the associated Toeplitz form, in either trace or Hilbert–Schmidt norm, and this breaks down at $\beta = 1/2$. By considering the perturbation in higher Schatten norms it may be possible to push our strategy further.

Theorem 1.1. *Assume the moment condition* (C2), *and let* $\sigma = \max\{1/\beta, 3/(2\gamma)\}$. *Then, for* (*non-constant*) *real-valued* $f \in F\ell^2_\sigma$, *the normalized statistics*

$$X^{\mathrm{scal}}_{f,n} := \frac{X_{f,n} - nf_0}{\sqrt{\iota_\mu(2n)}}$$

converges in law to a mean zero Gaussian with variance $\sum_{k\in\mathbb{Z}} k^2|f_k|^2$ *as* $n \to \infty$.

Staying with (C2) if we assume the particular asymptotics (10), then it holds

$$\iota_\mu(2n) \sim \begin{cases} \dfrac{\alpha \log(2n)}{2}, & \beta = 1, \\ \dfrac{\alpha(2n)^{1-\beta}}{2(1-\beta)}, & 1/2 < \beta < 1. \end{cases}$$

For canonical $\beta = 1$ cases like Ginibre, we have $\sigma = 1$ and hence the assumed regularity on f is optimal. For $\beta < 1$, because the asymptotic variance of $X_{n,f}$ is $\sim n^{1-\beta}$ and the mean is $\sim n$, one may conclude a CLT from [23] (even for $\beta \le 1/2$), though for possibly different classes of f. This highlights what our method can and cannot accomplish.

Next we define the infinite version of the matrix $M_{\mu,n}(a)$ and the related Toeplitz operator,

$$M_\mu(a) = (\varrho_{j,k} a_{j-k}), \qquad T(a) = (a_{j-k}), \quad j, k \ge 0, \tag{11}$$

both viewed as bounded linear operators on $\ell^2 = \ell^2(\mathbb{Z}_+)$, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$.

Theorem 1.2. *Assume the moment condition* (C1), *and let* f *be real-valued and non-constant*.

(a) *If* $f \in F\ell^2_{1/\beta}$ *for* $\beta < 2$ *or* $f \in F\ell^2_{1/2} \cap L^\infty(\mathbb{T})$ *for* $\beta \ge 2$, *then*

$$X_{f,n} - nf_0 \Rightarrow \mathcal{Z}$$

as $n \to \infty$ *with a mean-zero random variable* $\mathcal{Z} = \mathcal{Z}(f;\mu)$ *of finite, positive variance*

$$\mathrm{Var}(\mathcal{Z}) = 2\sum_{k=1}^{\infty} k|f_k|^2 + \sum_{j,k=0}^{\infty} \left(1 - \varrho^2_{j,k}\right)|f_{j-k}|^2.$$

(b) *If* $f \in F\ell^1_\sigma$ *or* $f \in F\ell^2_{\sigma+\varepsilon}$, *where* $\sigma = \max\{1, 2/\beta\}$, $\varepsilon > 0$, *then the higher cumulants* c_m *of* $\mathcal{Z}$ *may be described as follows. Introduce the recursion*

$$C_m = M_\mu\left(f^m\right) - \sum_{k=1}^{m-1} \binom{m-1}{k} C_{m-k} M_\mu\left(f^k\right), \quad m \ge 1.$$

Then $c_2(\mathcal{Z}) = \mathrm{Var}(\mathcal{Z}) = \mathrm{trace}\, C_2 + \sum_{k=1}^\infty k|f_k|^2$, *while* $c_m(\mathcal{Z}) = \mathrm{trace}\, C_m$ *for* $m \ge 3$.

For CUE, $\varrho_{k,\ell} \equiv 1$, and one can check that $c_m = 0$ for all $m \ge 3$ and so $\mathcal{Z}$ is Gaussian. That is to say the obvious: Theorem 1.2 reduces to the strong Szegö theorem. In general though it does not appear efficient to compute the cumulants of $\mathcal{Z}$ from the formula above, even in explicit, and seemingly simple examples like truncated Bergman for which $\varrho_{k,\ell} = \frac{2\sqrt{(k+1)(\ell+1)}}{k+\ell+2}$. The more basic problem which remains open is to determine when $\mathcal{Z}$ is Gaussian, i.e., for what μ does c_m vanish for all $m \ge 3$.

We conjecture this is only the case for CUE, or when μ places all its mass on a single point. The intuition stems from the calculation performed below in Proposition 2.2, which shows that if, for example, μ has a "nice" density supported on some $[a, b]$ ($0 \le a < b < \infty$), the normalized counting measure of points will concentrate on $|z| = b$ as $n \to \infty$ while there will remain O(1) points of modulus $\in [a, b - \varepsilon]$ (for whatever $\varepsilon > 0$). A Gaussian noise should result from the O(n) points interacting about $|z| = b$, while the (non-Gaussian) statistics of the phases of the points of modulus $< b$ will not wash in the type of centered (but not scaled) limit considered in Theorem 1.2.

Theorems 1.1 and 1.2 are intimately connected to the following, direct generalization of the Szegö–Widom limit theorem to the determinants of $M_{\mu,n}(a)$.

Theorem 1.3.

(a) *Assume the moment condition* (C2), *let* $\sigma = \max\{1/\beta, 3/(2\gamma)\}$ *and* $B = F\ell^2(\nu)$ *such that* $\nu_m = \nu_{-m}$, ν_m *is increasing* ($m \geq 1$), *and*

$$\nu_m \geq \max\Big\{\big(1+|m|\big)^{\sigma}, \sqrt{1+m^2\iota_\mu\big(2|m|^{2\sigma}\big)}\Big\}, \quad \sup_{m\geq 1}\frac{\nu_{2m}}{\nu_m} < \infty. \tag{12}$$

Let $a \in B$ *and suppose* $T(a)$ *is invertible on* ℓ^2. *Then*

$$\lim_{n\to\infty}\frac{\det M_{\mu,n}(a)}{G[a]^n \exp(\iota_\mu(2n)\Omega[a])} = F[a], \tag{13}$$

with some constant $F[a]$ *and*

$$G[a] = \exp\big([\log a]_0\big), \qquad \Omega[a] = \frac{1}{2}\sum_{k=-\infty}^{\infty} k^2[\log a]_k[\log a]_{-k}. \tag{14}$$

(b) *Assume the moment condition* (C1), *let* $a \in L^\infty(\mathbb{T}) \cap F\ell^2_{1/2}$ *if* $\beta \geq 2$ *or* $a \in F\ell^2_{1/\beta}$ *if* $1 < \beta < 2$. *Suppose* $T(a)$ *is invertible on* ℓ^2. *Then*

$$\lim_{n\to\infty}\frac{\det M_{\mu,n}(a)}{G[a]^n} = E[a], \tag{15}$$

for a constant $E[a]$. *If further* $a \in F\ell^1_\sigma$ *or* $a \in F\ell^2_{\sigma+\varepsilon}$, $\sigma = \max\{1, 2/\beta\}$, $\varepsilon > 0$, *there is the expression*

$$E[a] = \det\big(T\big(a^{-1}\big)M_\mu(a)\big). \tag{16}$$

The convergences in (13) *and* (15) *is uniform* (*in* a) *on compact subsets of the indicated function spaces.*

The assumption that $T(a)$ is invertible is a natural assumption on the symbol; it is the condition in the (scalar) Szegö–Widom theorem (see [4], Ch. 10, and [24]). One of the general versions of that theorem pertains to symbols drawn from the Krein algebra $\mathrm{K} = L^\infty(\mathbb{T}) \cap F\ell^2_{1/2}$ (which contains discontinuous functions). Hence, at least for $\beta \geq 2$, we achieve the same level of generality.

Except for the Krein algebra K, the various classes of symbols occurring above are Banach algebras continuously embedded in $C(\mathbb{T})$. For those classes, the assumption that $T(a)$ be invertible is equivalent to requiring that a possesses a continuous logarithm $\log a$ on the unit circle $\mathbb{T}$, which then enters the definition of the constant $G[a]$ and $\Omega[a]$. In other words, the continuous function a is nonzero on all of $\mathbb{T}$ and has winding number zero. In case of K, where a can be discontinuous, we must define

$$G[a] = \big[T^{-1}\big(a^{-1}\big)\big]_{00} \tag{17}$$

as the $(0,0)$-entry in the matrix representation of the inverse Toeplitz operator, as is well known in the context of the classical Szegö–Widom theorem.

The quite technical assumptions in (12) can be simplified in special situations such as (10). Then, in case $1/2 < \beta < 1$ we can take $B = F\ell^2_\sigma$, $\sigma = 1/\beta$, while in case $\beta = 1$ we can take $B = F\ell^2(\nu)$, $\nu_m = C(1+|m|)\log^{1/2}(2+|m|)$, which is only slightly stronger than one might expect.

The constant $E[a]$ can be expressed via a well defined operator determinant (16) (see, e.g., [11] for the underlying notions). Its explicit evaluation appears quite hard, although in the CUE case where $M_\mu(a) = T(a)$ its evaluation is classical [24] (see also [4,7,8]). It is exactly this evaluation problem that ties to the identification of $\mathcal{Z}$ in Theorem 1.2. The constant $F[a]$ involves an even more complicated operator determinant.

The theorems above are derived in Sections 6 and 7, as a consequence of a more general result, Theorem 4.4 (Section 4), on the asymptotics of determinants of type (2). Section 3 lays out various preliminaries required for the proof of Theorem 4.4, and also explains how we employ the moment assumption. Section 5 provides detailed asymptotics of a certain trace term occurring in Theorem 4.4 which is tied to the variance of $X_{f,n}$.

Throughout we have considered fixed radial measures μ; any simple scaling of μ in n will not affect any appraisal concerning the angular statistics. There are however examples of interest which fall out of this set-up. Take for instance the so-called spherical ensemble connected to $A^{-1}B$ in which A and B are independent $n \times n$ Ginibre matrices. The resulting eigenvalues form a determinantal process with $\mathrm{d}\mu_n(r) = r(1+r^2)^{-(n+1)}\,\mathrm{d}r$ [17]. Another example is provided by the roots of the degree-n complex polynomial with Mahler measure one, for which $\mathrm{d}\mu_n(r) = r\min(1, r^{-2n-2})\,\mathrm{d}r$ [5]. Our methods could perhaps be adopted to both situations, but we do not pursue this.

2. On the moment condition

Of the key examples, both CUE and truncated Bergman satisfy $(\ln m_\xi)'' = \mathrm{O}(\xi^{-2})$, while the Ginibre ensemble satisfies $(\ln m_\xi)'' = \mathrm{O}(\xi^{-1})$. A few more examples are contained in the following.

Proposition 2.1. *Consider positive measures on $\mathbb{R}_+$ with density $\mathrm{d}\mu(r) = \mu(r)\,\mathrm{d}r$ and corresponding moment function $m_\xi = \int_0^\infty r^\xi \mu(r)\,\mathrm{d}r$.*

- (i) *If $\mu(r)$ is supported on a finite interval $[a, b]$, and is "regular" at b as in $\mu(r) = c(b-r)^{\alpha-1}$ for $r \in (b-\delta, b]$ and $\alpha > 0$, then $(\ln m_\xi)'' = \alpha\xi^{-2} + \mathrm{O}(\xi^{-3})$.*
- (ii) *If $\mu(r) = p(r)\mathrm{e}^{-cr^\alpha}$ for polynomials p and $\alpha > 0$, then $(\ln m_\xi)'' = \alpha\xi^{-1} + \mathrm{O}(\xi^{-2})$.*
- (iii) *If $\mu(r) = \mathrm{e}^{-c(\ln(\mathrm{e}+r))^q}$ for $q > 1$, then $(\ln m_\xi)'' = \alpha\xi^{(2-q)/(q-1)} + \mathrm{O}(\xi^{(3-2q)/(q-1)})$ upon choosing $c = \alpha^{1-q}(q^{1/(1-q)} - q^{q/(1-q)})$.*

Proof. We start with explicit instances of cases (i) and (ii). For (i), there is no loss in assuming that $[a, b] = [0, 1]$ and we consider further $\mu^{(\mathrm{i})}(r) = (1-r)^{\alpha-1}1_{[0,1]}$. For case (ii), consider a simple polynomial term $\mu^{(\mathrm{ii})}(r) = r^p\mathrm{e}^{-r^\alpha}$. Then we have,

$$\ln m_\xi^{(\mathrm{i})} = \ln\Gamma(\xi+1) - \ln\Gamma(\xi+\alpha+1) + \ln\Gamma(\alpha)$$

and

$$\ln m_\xi^{(\mathrm{ii})} = \ln\Gamma\big((\xi+p+1)/\alpha\big) - \ln\alpha.$$

From this point the verifications may be completed using that $\frac{\mathrm{d}^2}{\mathrm{d}z^2}\ln\Gamma(z) = z^{-1} + (1/2)z^{-2} + \mathrm{O}(z^{-3})$ for large real values of z.

More generally, for case (i) we write

$$(\ln m_\xi)'' = \frac{\langle(\ln r)^2 r^\xi\rangle_\mu}{\langle r^\xi\rangle_\mu} - \frac{\langle(\ln r) r^\xi\rangle_\mu^2}{\langle r^\xi\rangle_\mu^2}$$

and note that Laplace asymptotic considerations yield: for $d = 0, 1, 2$, $\langle(\log r)^d r^\xi\rangle_\mu = \langle(\log r)^d r^\xi\rangle_{\mu^{(\mathrm{i})}} + \mathrm{O}(\mathrm{e}^{-C_\delta\xi})$, which is more than enough to show that one has the same asymptotics for any such μ as for $\mu^{(\mathrm{i})}$. That (ii) extends to more general polynomials $p(r)$ is self-evident.

For case (iii) we only mention that it is most convenient to consider the asymptotically equivalent object $m_\xi = \int_0^\infty \mathrm{e}^{\xi r - cr^q}\,\mathrm{d}r$ (after an obvious change of variable) for which the leading order arises from a neighborhood of the stationary point $r^* = (\xi/cq)^{1/(q-1)}$. The details are straightforward. □

The above is intended to be illustrative; no attempt to optimize the regularity conditions on μ has been made. We also mention here without proof that the measure $\mathrm{d}\mu(r) = \mathrm{e}^{-\mathrm{e}^r}\,\mathrm{d}r$ produces a moment sequence for which there is the not strictly polynomial decay $(\ln m_\xi)'' = \mathrm{O}(\frac{1}{\xi \ln \xi})$. Further, by Fourier inversion, one may produce measures for which $(\log m_\xi)''$ is exactly $\alpha(1+\xi)^{-\beta}$ for $0 < \beta \le 2$, $\beta \neq 1$, $\alpha > 0$.

Moment condition and the mean measure. *Our condition(s) on the moment sequence also dictate the limit shape of the mean measure of the points. This object is given by*

$$\mathrm{d}\Lambda_n(z) = \left(\frac{1}{n} \sum_{k=1}^{n-1} \frac{|z|^{2k}}{2\pi m_{2k}} \right) \mathrm{d}\mathfrak{m}(z);$$

as the name suggests $\mathbb{E}_{\mathfrak{m},n}[\#\text{ points in } A] = n \int_A \mathrm{d}\Lambda_n(z)$ *for (measurable)* $A \subseteq \mathbb{C}$, *see again* [13]. *We provide one description of the shift from a "*$\beta = 1$*" setting, resulting in an extended limit support, to a "*$\beta > 1$*" setting for which the limit support is degenerate. This is in line with the conjecture discussed after Theorem* 1.2.

Proposition 2.2. *For all sufficiently large* ξ *let the moment sequence* $m_\xi = \int_0^\infty r^\xi \,\mathrm{d}\mu(r)$ *satisfy*

$$(\ln m_\xi)'' = \frac{\alpha}{\xi+1} + \varepsilon(\xi) \tag{18}$$

with $\alpha \ge 0$ *and* $\varepsilon \in L^1(\mathbb{R}_+)$. *Then there exists a rescaling of* $\mathbb{P}_{\mathfrak{m},n}$ *so that* $\mathrm{d}\Lambda_n$ *converges weakly to either: a weighted circular law with density* $\frac{1}{2\pi\alpha}|z|^{1/\alpha-2}$ *on* $|z| \le 1$ *when* $\alpha > 0$, *or to the uniform measure on* $|z| = 1$ *when* $\alpha = 0$.

Note, ε is necessarily nonnegative when $\alpha = 0$. And of course, when $\alpha = 1/2$ the advertised limit is the standard circular law (see e.g. [2]).

Proof. Choose $q \gg 1$ so that (18) is in effect for $s \ge q$, and then integrate the equality twice: first over $q \le s \le t$, and then in t from k to $k+\ell$ to find

$$\ln\left(\frac{m_{k+\ell}}{m_k}\right) = \alpha\ell \log k + c\ell + \mathrm{o}(1). \tag{19}$$

(Here $c = (\ln m)'(q) + \alpha \ln m(q) - \int_q^\infty \varepsilon(s)\,\mathrm{d}s$, and the $\mathrm{o}(1)$ holds in k — we view ℓ as fixed.) Next compute the ℓth absolute moment in the mean measure: which here can be considered as supported on $\mathbb{R}_+$,

$$\int_{\mathbb{C}} |z|^\ell \,\mathrm{d}\Lambda_n(z) = \frac{1}{n} \sum_{k=0}^{n-1} \frac{m_{2k+\ell}}{m_{2k}} = \frac{1}{n} \left(2^\alpha \mathrm{e}^c\right)^\ell \sum_{k=1}^{n-1} k^{\alpha\ell}\left(1+\mathrm{o}(1)\right). \tag{20}$$

Neglecting the multiplicative errors, in the case $\alpha = 0$ the sum (20) converges to $\mathrm{e}^{c\ell}$ for any ℓ, unambiguously the moment sequence defined by placing unit mass at the place $\mathrm{e}^c \in \mathbb{R}_+$. When $\alpha > 0$, we rescale $\mathbb{P}_{\mathfrak{m},n}$ by sending $\{z_i\}_{1 \le i \le n} \mapsto \{n^{-\alpha} z_i\}_{1 \le i \le n}$. Then, the sum converges to $(2^\alpha \mathrm{e}^c)^\ell \frac{\mathrm{e}^{c\ell}}{\alpha\ell+1}$ as $n \to \infty$. Matching constants in $\int_0^b t^\ell \,\mathrm{d}(t/b)^{p+1} = \frac{p+1}{p+\ell+1} b^\ell$ identifies (uniquely) the scaled $\alpha > 0$ moment sequence with that of the measure with density $(p+1)t^p/b^{p+1}$ on $[0,b]$ where $b = 2^\alpha \mathrm{e}^c$ and $p = \frac{1-\alpha}{\alpha}$. Thus the limiting mean measure is also identified. In either case, $\alpha > 0$ or $\alpha = 0$, an additional rescaling will place the outer edge of the support at one. □

3. Hilbert–Schmidt and trace class conditions

Our results hinge on being able to consider $M_\mu(a)$ as a suitable compact perturbation of the Toeplitz operator $T(a)$ (see (11)). Here we will establish sufficient conditions on a and μ such that

$$K_\mu(a) = M_\mu(a) - T(a) = \left((\varrho_{j,k} - 1)a_{j-k}\right)), \quad j,k \ge 0,$$

is Hilbert–Schmidt or trace class operator. We refer to [11] for general information about these notions. Since $T(a)$ is bounded on ℓ^2 whenever $a \in L^\infty(\mathbb{T})$, under the appropriate conditions $M_\mu(a)$ is then also bounded. While it might be interesting to ask for necessary and sufficient conditions for the boundedness of $M_\mu(a)$ and the compactness of $K_\mu(a)$, we think it is a non-trivial issue, which we will not pursue here.

The compactness properties of $K_\mu(a)$ rely mainly on the "shape" of $\varrho_{j,k}$ near the diagonal. An application of Hölder's inequality shows that $0 < \varrho_{j,k} \le 1$. More detailed information on $\varrho_{j,k}$ is provided by the following technical lemma, for which we use the set of indices,

$$\mathcal{I}_\delta = \{(j,k) \in \mathbb{Z}_+ \times \mathbb{Z}_+ : |j-k|^\delta < (j+k)/2\}, \tag{21}$$

always assuming $\delta \ge 1$ ($\mathbb{Z}_+ = \{0, 1, \ldots\}$). The factor $1/2$ in $\mathcal{I}_\delta$ is only for technical convenience. In particular, $(j,k) \in \mathcal{I}_\delta$ implies $j, k \ge 1$.

Part (a) of the following lemma will be used at several places, while the more elaborate part (b) is used only in Lemma 5.2 and under the assumption $\varrho, \gamma > 1 \ge \beta > 1/2$ (see the moment condition (C2)), although the statement remains true in general. Notice that for $\varrho \le \beta$, part (b) reduces to part (a) since one can put $h_\mu = 0$. Throughout what follows we will utilize the notation $a \vee b := \max\{a, b\}$.

Lemma 3.1.

(a) *Let* $\beta > 0$, $\delta \ge 1$, $\beta\delta \ge 2$, *and assume that the measure* μ *satisfies the condition*

$$(\ln m_\xi)'' = \mathrm{O}(\xi^{-\beta}), \quad \xi \to \infty.$$

Then, for $(j,k) \in \mathcal{I}_\delta$ *with* $\Delta = j-k$, $\sigma = j+k$, *we have the uniform estimate*

$$\varrho_{j,k} = 1 + \mathrm{O}\left(\frac{\Delta^2}{\sigma^\beta}\right). \tag{22}$$

(b) *Let* $\beta, \gamma, \varrho > 0$, $\delta \ge 1$, $\beta\delta \ge 2$, $\gamma\delta \ge 3$, $\varrho\delta \ge 2$, *and assume that there exists a differentiable function* $h_\mu(\xi) \ge 0$ *such that*

$$(\ln m_\xi)'' = h_\mu(\xi) + \mathrm{O}(\xi^{-\varrho}), \quad \xi \to \infty,$$

and

$$h_\mu(\xi) = \mathrm{O}(\xi^{-\beta}), \qquad h'_\mu(\xi) = \mathrm{O}(\xi^{-\gamma}), \quad \xi \to \infty.$$

Then, for $(j,k) \in \mathcal{I}_\delta$ *with* $\Delta = j-k$, $\sigma = j+k$, *we have the uniform estimate*

$$\varrho_{j,k} = 1 - \frac{\Delta^2}{2} h_\mu(\sigma) + \mathrm{O}\left(\frac{\Delta^4}{\sigma^{2\beta}} \vee \frac{|\Delta|^3}{\sigma^\gamma} \vee \frac{\Delta^2}{\sigma^\varrho}\right). \tag{23}$$

Proof. We can assume without loss of generality that $\Delta > 0$. Then

$$\ln \varrho_{j,k} = \ln m_\sigma - \frac{\ln m_{\sigma+\Delta} + \ln m_{\sigma-\Delta}}{2} = -\frac{\Delta^2}{2}(\ln m_\eta)'', \quad \eta \in (\sigma - \Delta, \sigma + \Delta),$$

after applying the mean-value theorem twice. We can write $\eta = \sigma(1+\tau)$, where the error term τ is estimated by $|\tau| \le |\Delta|/\sigma \le |\Delta|^\delta/\sigma \le 1/2$ using $\delta \ge 1$.

In case (a) we can conclude that

$$\ln \varrho_{j,k} = \mathrm{O}\left(\frac{\Delta^2}{\eta^\beta}\right) = \mathrm{O}\left(\frac{\Delta^2}{\sigma^\beta}\right).$$

Because $\beta\delta \ge 2$ we get $\Delta^2 \le \sigma^{2/\delta} \le \sigma^\beta$. Hence the above term is bounded and exponentiating yields the assertion. In case (b) we first obtain

$$\ln \varrho_{j,k} = -\frac{\Delta^2}{2} h_\mu(\eta) + \mathrm{O}\left(\frac{\Delta^2}{\eta^\varrho}\right).$$

Now we apply once more the mean value theorem to obtain the estimate

$$\ln \varrho_{j,k} = -\frac{\Delta^2}{2} h_\mu(\sigma) + \mathrm{O}\biggl(\frac{|\Delta|^3}{\sigma^\gamma} \vee \frac{\Delta^2}{\sigma^\varrho}\biggr).$$

Notice that, as above, $\eta = \sigma(1+\tau)$ with $|\tau| \le 1/2$. All these terms are bounded because $2/\delta \le \beta$, $3/\delta \le \gamma$, and $2/\delta \le \varrho$. The assertion is obtained upon exponentiating. □

Part (a) of the lemma translates immediately into the estimates that follow.

Proposition 3.2. *Let $\beta > 1/2$ and assume that the measure μ satisfies the assumption*

$$(\ln m_\xi)'' = \mathrm{O}(\xi^{-\beta}), \quad \xi \to \infty.$$

Put $\sigma = 1/2 \vee 1/\beta$. Then there exists a constant $C_\mu > 0$ such that $K_\mu(a)$ is Hilbert–Schmidt and the estimate

$$\|K_\mu(a)\|_{\mathcal{C}_2(\ell^2)} \le C_\mu \|a\|_{F\ell^2_\sigma}$$

holds whenever $a \in F\ell^2_\sigma$.

Proof. Put $\delta = 2\sigma = 1 \vee 2/\beta$ so that Lemma 3.1(a) is applicable. The operator $K_\mu(a)$ is Hilbert–Schmidt if and only if the sum $\sum_{(j,k)\in\mathbb{Z}^2_+} |a_{j-k}|^2(1-\varrho_{j,k})^2$ is finite (this quantity is the square of the Hilbert–Schmidt norm). We have that

$$\begin{aligned}
\sum_{(j,k)\in\mathbb{Z}^2_+} |a_{j-k}|^2(1-\varrho_{j,k})^2 &\le \sum_{(j,k)\notin\mathcal{I}_\delta} |a_{j-k}|^2 + \sum_{(j,k)\in\mathcal{I}_\delta} |a_{j-k}|^2(1-\varrho_{j,k})^2 \\
&\le \sum_{\substack{(d,s)\in\mathbb{Z}\times\mathbb{Z}_+ \\ |d|^\delta \ge s/2}} |a_d|^2 + \sum_{\substack{(d,s)\in\mathbb{Z}\times\mathbb{Z}_+ \\ |d|^\delta < s/2}} |a_d|^2 \frac{d^4}{s^{2\beta}} \\
&\le C\sum_{d\in\mathbb{Z}} |a_d|^2|d|^\delta + C\sum_{d\in\mathbb{Z}} |a_d|^2|d|^{4+\delta(1-2\beta)}.
\end{aligned}$$

Line one just uses $\varrho_{j,k} \in (0,1]$. In line two we make the substitution $d = j-k$, $s = j+k$ and employ Lemma 3.1(a), and the final line uses the fact $\beta > 1/2$. Furthermore, as $\delta\beta \ge 2$ we see that the second term in this last line does not exceed the first one, and that in turn is equal to the square of $\|a\|_{F\ell^2_\sigma}$ ($\delta = 2\sigma$). □

Next we establish two sufficient conditions for $K_\mu(a)$ to be trace class. It is not hard to show that one is not weaker than the other, i.e., neither of the two function classes pointed out below is contained in the other.

Proposition 3.3. *Let $\beta > 1$ and assume that the measure μ satisfies the assumption*

$$(\ln m_\xi)'' = \mathrm{O}(\xi^{-\beta}), \quad \xi \to \infty.$$

Put $\sigma = 1 \vee 2/\beta$. Then there exists $C_\mu > 0$ and, for each $\varepsilon > 0$, $C_{\mu,\varepsilon} > 0$ such that

(a) *$K_\mu(a)$ is trace class and the estimate*

$$\|K_\mu(a)\|_{\mathcal{C}_1(\ell^2)} \le C_\mu \|a\|_{F\ell^1_\sigma}$$

holds whenever $a \in F\ell^1_\sigma$;

(b) $K_\mu(a)$ *is trace class and the estimate*

$$\left\|K_\mu(a)\right\|_{\mathcal{C}_1(\ell^2)} \le C_{\mu,\varepsilon}\|a\|_{F\ell^2_{\sigma+\varepsilon}}$$

holds whenever $a \in F\ell^2_{\sigma+\varepsilon}$.

Proof. Here we put $\delta = \sigma = 1 \vee 2/\beta$ and notice that then Lemma 3.1(a) is again applicable.

(a) We first estimate the trace norm of $K_\mu(t^m)$, $m \in \mathbb{Z}$. Without loss of generality assume $m > 0$. Then $K_\mu(t^m)$ has entries on the mth diagonal given by $\{\varrho_{k+m,k} - 1\}_{k=0}^\infty$. This operator is trace class if and only if its trace norm

$$\sum_{k=0}^{\infty} |\varrho_{k+m,k} - 1| < \infty.$$

We split and overestimate this sum by a constant times

$$\sum_{(k+m,k)\notin\mathcal{I}_\delta} 1 + \sum_{(k+m,k)\in\mathcal{I}_\delta} \frac{m^2}{(2k+m)^\beta},$$

using Lemma 3.1(a) for the second part. Now $(k+m,k) \in \mathcal{I}_\delta$ means that $m^\delta < (2k+m)/2$, i.e., $2k > 2m^\delta - m$. Noting that $2k \le 2m^\delta - m$ implies $k < m^\delta$, and $2k > 2m^\delta - m$ implies $2k > m^\delta$, the previous terms are overestimated by

$$\sum_{0\le k<m^\delta} 1 + \sum_{k\ge m^\delta/2} \frac{m^2}{(2k)^\beta} \le m^\delta + Cm^{2+\delta(1-\beta)} \le (1+C)m^\delta.$$

Here we used $\beta > 1$ and $\delta\beta \ge 2$, and all estimates are uniform in m. Thus $\|K_\mu(t^m)\|_{\mathcal{C}_1(\ell^2)} = \mathrm{O}(|m|^\delta)$. From here the proof of (a) follows immediately.

(b) Introduce the diagonal operator $\Lambda = \mathrm{diag}((1+k)^{-1/2-\varepsilon})$, $\varepsilon > 0$, acting on ℓ^2. As Λ is Hilbert–Schmidt it suffices to prove that the operator with the matrix representation of $K_\mu(a)\Lambda^{-1}$ is Hilbert–Schmidt. The squared Hilbert–Schmidt norm of $K_\mu(a)\Lambda^{-1}$ equals

$$\sum_{(j,k)\in\mathbb{Z}_+^2} |a_{j-k}|^2(1+k)^{1+2\varepsilon}(1-\varrho_{j,k})^2.$$

As before we split the sum into two parts,

$$\sum_{(j,k)\notin\mathcal{I}_\delta} |a_{j-k}|^2(1+j+k)^{1+2\varepsilon} + \sum_{(j,k)\in\mathcal{I}_\delta} |a_{j-k}|^2(1-\varrho_{j,k})^2(1+j+k)^{1+2\varepsilon},$$

slightly overestimating it further. Now we make the substitution $d = j-k \in \mathbb{Z}$ and $s = j+k \in \mathbb{Z}_+$. We arrive at the upper estimate for the first term

$$\sum_{\substack{(d,s)\in\mathbb{Z}\times\mathbb{Z}_+\\ |d|^\delta\ge s/2}} |a_d|^2(1+s)^{1+2\varepsilon} \le C\sum_{d\in\mathbb{Z}} |a_d|^2\big(1+|d|\big)^{\delta(2+2\varepsilon)} \le C\|a\|^2_{F\ell^2_{\delta(1+\varepsilon)}}.$$

For the second term, employ $(1-\varrho_{j,k})^2 \le C(j-k)^4(1+j+k)^{-2\beta}$, by Lemma 3.1(a), to find that it is bounded by a constant times

$$\sum_{(j,k)\in\mathcal{I}_\delta} |a_{j-k}|^2\frac{(j-k)^4}{(1+j+k)^{2\beta-1-2\varepsilon}} \le \sum_{\substack{(d,s)\in\mathbb{Z}\times\mathbb{Z}_+\\ |d|^\delta<s/2}} |a_d|^2\frac{d^4}{(1+s)^{2\beta-1-2\varepsilon}}.$$

Without loss of generality we could have chosen $\varepsilon > 0$ small enough such that $\beta > 1 + \varepsilon$. Then we can estimate further by a constant times

$$\sum_{d\in\mathbb{Z}} |a_d|^2 |d|^{4+\delta(2+2\varepsilon-2\beta)} \le \sum_{d\in\mathbb{Z}} |a_d|^2 |d|^{\delta(2+2\varepsilon)} \le \|a\|^2_{F\ell^2_{\sigma(1+\varepsilon)}}.$$

This proves the assertion. □

Remark. *The condition $\beta > 1$ is (in a certain sense) necessary to ensure that $K_\mu(a)$ is trace class. More precisely, assume that the measure μ satisfies the condition*

$$(\ln m_\xi)'' = \frac{\alpha}{\xi^\beta} + \mathrm{O}(\xi^{-\varrho}), \quad \alpha > 0, 1/2 < \beta \le 1, \varrho > \beta. \tag{24}$$

Choose $\delta > 2/\beta > 1$. Using Lemma 3.1(b) it follows easily that

$$\varrho_{j,k} = 1 - \frac{\alpha(j-k)^2}{2(1+j+k)^\beta}\big(1+\mathrm{o}(1)\big)$$

for indices $(j,k) \in \mathcal{I}_\delta$. Moreover for each fixed m, the entries $(k, m+k)$ belongs to $\mathcal{I}_\delta$ for all sufficiently large $k \ge k_0(m)$. Thus the mth diagonal has entries

$$a_m(\varrho_{k,k+m} - 1) = a_m \frac{\alpha m^2}{2(1+m+2k)^\beta}\big(1+\mathrm{o}(1)\big), \quad k \ge k_0(m).$$

This growth (in k) is too large to allow $K_\mu(a)$ to be trace class unless $ma_m = 0$. That is, under (24), the operator $K_\mu(a)$ can only be trace class in the trivial case of constant symbol.

4. Determinant asymptotics

Recall that given a function $a \in L^\infty(\mathbb{T})$ with Fourier coefficients a_n, the Toeplitz and the Hankel operator acting on $\ell^2 = \ell^2(\mathbb{Z}_+)$, are defined by their infinite matrix representations

$$T(a) = (a_{j-k}), \qquad H(a) = (a_{j+k+1}), \quad 0 \le j, k < \infty. \tag{25}$$

It is well known that the relations

$$T(ab) = T(a)T(b) + H(a)H(\tilde{b}), \tag{26}$$

$$H(ab) = T(a)H(b) + H(a)T(\tilde{b}), \tag{27}$$

hold, where $\tilde{b}(t) = b(t^{-1})$, $t \in \mathbb{T}$. For later use, introduce the flip operators and the projection operators acting on ℓ^2,

$$W_n : \{x_0, x_1, \ldots\} \mapsto \{x_{n-1}, x_{n-2}, \ldots, x_1, x_0, 0, 0, \ldots\},$$

$$P_n : \{x_0, x_1, \ldots\} \mapsto \{x_0, x_1, \ldots, x_{n-2}, x_{n-1}, 0, 0, \ldots\}, \quad Q_n = I - P_n,$$

where $n \in \{1, 2, 3, \ldots\}$, as well as the forward and backward shift operators,

$$V_n : \{x_k\}_{k=0}^\infty \mapsto \{y_k\}_{k=0}^\infty, \quad y_k = \begin{cases} 0 & \text{if } k < n, \\ x_{k-n} & \text{if } k \ge n, \end{cases}$$

$$V_{-n} : \{x_k\}_{k=0}^\infty \mapsto \{x_{n+k}\}_{k=0}^\infty.$$

It is easy to verify that $V_{\pm n} = T(t^{\pm n})$, $W_n = H(t^n)$, and

$$V_n V_{-n} = Q_n, \qquad V_{-n} V_n = I, \qquad W^n P_n = P_n W_n = W_n, \quad W_n^2 = P_n. \tag{28}$$

Consistent with previous notation, we denote by $T_n(a)$ and $M_{\mu,n}(a)$ the $n\times n$ upper-left submatrices of the matrix representation of $T(a)$ and $M_\mu(a)$, i.e.,

$$T_n(a) = P_n T(a) P_n, \qquad M_{\mu,n}(a) = P_n M_\mu(a) P_n.$$

Here we identify the upper-left $n\times n$ block in the matrix representation of the operators on the right-hand sides with the $\mathbb{C}^{n\times n}$ matrices on the left-hand sides.

In this section we are going to establish the main auxiliary result (Theorem 4.4), which reduces the asymptotics of the determinant $\det M_{\mu,n}(a)$ to the asymptotics of a trace (or already gives the determinant asymptotics up to the computation of a constant). This and the main results hold either for the Krein algebra $\mathrm{K} = L^\infty(\mathbb{T}) \cap F\ell^2_{1/2}$ (see [4], Ch. 10), or for several subalgebras of $C(\mathbb{T})$, which satisfy "suitable conditions." Therefore, it seems convenient to formulate Theorem 4.4 below in a quite general context and to make use of the following definition.

Definition 4.1. *Given a unital Banach algebra B which is continuously embedded in $L^\infty(\mathbb{T})$, denote by $\Phi(B)$ the set of all $a\in B$ such that the Toeplitz operator $T(a)$ is invertible on ℓ^2. We say such a Banach algebra B suitable if*:

(a) *B is continuously embedded in $\mathrm{K} = L^\infty(\mathbb{T}) \cap F\ell^2_{1/2}$.*
(b) *If $a\in\Phi(B)$, then $a^{-1}\in\Phi(B)$.*

The next proposition demonstrates the suitability of several Banach algebras which appear in the main results.

Proposition 4.2. *With $W = F\ell^1_0$ denoting the Wiener algebra, the following are suitable Banach algebras*:

(i) *$W\cap F\ell^2_\sigma = F\ell^2_\sigma$ for $\sigma > 1/2$;*
(ii) *$F\ell^1_\sigma$ for $\sigma \geq 1/2$;*
(iii) *$W\cap F\ell^2_{1/2}$ and $\mathrm{K} = L^\infty(\mathbb{T})\cap F\ell^2_{1/2}$;*
(iv) *$W\cap F\ell^2(\nu)$ provided that $\nu_{-n} = \nu_n \geq n^{1/2}$, $\{\nu_n\}_{n=1}^\infty$ is increasing, and $\sup_{n\geq 1}\frac{\nu_{2n}}{\nu_n} < \infty$.*

Proof. First of all, the above are indeed Banach algebras. This is elementary for $F\ell^1_\sigma$. A proof for $W\cap F\ell^2_\sigma$, $\sigma\geq 0$, can be found in [4], Thm. 6.54, while the more general space $W\cap F\ell^2(\nu_\sigma)$ is treated in [16]. For K see, e.g., [4], Thm. 10.9. As for (i), note that $F\ell^2_\sigma$ is continuously embedded in W whenever $\sigma > 1/2$. Further, property (a) of suitability is immediate for these spaces.

Recall that a unital Banach algebra B is called inverse closed in Banach algebra $B_0 \supset B$ if $a\in B$ and $a^{-1}\in B_0$ implies that $a^{-1}\in B$. For all the Banach algebras B above, except for K, using simple Gelfand theory and the density of the Laurent polynomials it is easily seen that the maximal ideal space can be naturally identified with $\mathbb{T}$. (In the case of (iv), this is also proved in [16].) By a standard argument, this implies that these Banach algebras are inverse closed in $C(\mathbb{T})$, thus also in $L^\infty(\mathbb{T})$. For a proof of the inverse closedness of K in $L^\infty(\mathbb{T})$ see again [4], Thm. 10.9.

As for property (b), take $a\in\Phi(B)$, i.e., $a\in B$ such that $T(a)$ is invertible on ℓ^2. From the theory of Toeplitz operators it is well known that then a is invertible in $L^\infty(\mathbb{T})$. By the inverse closedness we thus have $a^{-1}\in B$. Now we observe that $b\in\mathrm{K}$ implies that both $H(b)$ and $H(\tilde{b})$ are Hilbert–Schmidt. Using the formulas

$$I = T(a)T\bigl(a^{-1}\bigr) + H(a)H\bigl(\tilde{a}^{-1}\bigr), \qquad I = T\bigl(a^{-1}\bigr)T(a) + H\bigl(a^{-1}\bigr)H(\tilde{a}), \tag{29}$$

and the implied compactness of the Hankel operators, it follows that $T(a^{-1})$ is a Fredholm regularizer for $T(a)$. (For information about Fredholm operators, see, e.g., [11].) Hence $T(a^{-1})$ is also Fredholm with index zero and thus invertible (by Coburn's lemma [4], Sec. 2.6). But this means that $a^{-1}\in\Phi(B)$. □

The next proposition shows (besides a technical result (ii)) that the constant $G[a]$ is well defined for all $a\in\Phi(B)$. This constant appears in our limit theorem as it did appear in the classical Szegö–Widom limit theorem. We follow closely the arguments of [4], Ch. 10.

Proposition 4.3. *Let B be a suitable Banach algebra, and $a \in \Phi(B)$.*

(i) *With $[*]_{00}$ the $(0,0)$-entry of the matrix representation on ℓ^2, the constant*

$$G[a] := \left[T^{-1}\left(a^{-1}\right)\right]_{00} \tag{30}$$

is nonzero.

(ii) *With $A_n = P_n T^{-1}(a^{-1}) P_n$, we have* $\det A_n = G[a]^n$, *and*

$$A_n^{-1} \to T\left(a^{-1}\right), \qquad \left(A_n^*\right)^{-1} \to T\left(a^{-1}\right)^*$$

strongly on ℓ^2 as $n \to \infty$. (A^* is the adjoint of A.) Moreover, the mappings

$$\Lambda_n : a \in \Phi(B) \mapsto A_n^{-1} \in \mathcal{L}\left(\ell^2\right),$$

are equi-continuous, $\mathcal{L}(\ell^2)$ being the space of bounded linear operators on ℓ^2.

(iii) *If $b \in B$, then $\mathrm{e}^b \in \Phi(B)$ and $G[\mathrm{e}^b] = \mathrm{e}^{b_0}$, where b_0 is the 0th Fourier coefficient.*

Proof. (i)–(ii) If $a \in \Phi(B)$, then $a^{-1} \in \Phi(B)$ and $T(a^{-1})$ is invertible. Hence the definitions of $G[a]$ and A_n make sense. Notice that for $n = 1$, we have $\det A_1 = A_1 = [T^{-1}(a^{-1})]_{00} = G[a]$. Consequently, (i) will follow from the invertibility of A_n in the case $n = 1$.

To show the invertibility of A_n we use a block operator inversion formula, sometimes also referred to as Kozak's formula. If P is a projection, $Q = I - P$ is the complementary projection, and A is an invertible operator, then $PAP|_{\mathrm{Im}(P)}$ is invertible if and only if so is $QA^{-1}Q|_{\mathrm{Im}(Q)}$. In fact, the formula

$$(PAP)|_{\mathrm{Im}(P)}^{-1} = PA^{-1}P|_{\mathrm{Im}(P)} - PA^{-1}Q\left(QA^{-1}Q\right)|_{\mathrm{Im}(Q)}^{-1} QA^{-1}P|_{\mathrm{Im}(P)} \tag{31}$$

holds, which can be easily verified (see also [4], Prop. 7.15).

Applying the above to $A_n = P_n T^{-1}(a^{-1}) P_n$ we see that A_n is invertible if and only if $Q_n T(a^{-1}) Q_n$ is invertible, and in this case we have

$$A_n^{-1} = P_n T\left(a^{-1}\right) P_n - P_n T\left(a^{-1}\right) Q_n \left(Q_n T\left(a^{-1}\right) Q_n\right)^{-1} Q_n T\left(a^{-1}\right) P_n. \tag{32}$$

Notice that $Q_n T(a^{-1}) Q_n$ is nothing but the "shifted" Toeplitz operator. Using the shift operators $V_{\pm n}$ satisfying the relations (28), we obtain $(Q_n T(a^{-1}) Q_n)^{-1} = V_n T^{-1}(a^{-1}) V_{-n}$ and hence

$$A_n^{-1} = P_n T\left(a^{-1}\right) P_n - P_n T\left(a^{-1}\right) V_n T^{-1}\left(a^{-1}\right) V_{-n} T\left(a^{-1}\right) P_n. \tag{33}$$

We have thus shown that A_n is invertible and in particular (i). Moreover, from this representation it follows immediately that the mappings Λ_n are equi-continuous. If suffices to remark that the operators P_n and $V_{\pm n}$ have norm one, and that the various mappings $b \in \Phi(B) \mapsto b^{-1} \in \Phi(B)$, $b \in B \mapsto T(b) \in \mathcal{L}(\ell^2)$, $B \in G\mathcal{L}(\ell^2) \mapsto B^{-1} \in G\mathcal{L}(\ell^2)$ are continuous. (Here $G\mathcal{L}(\ell^2)$ stands for the group of all invertible bounded linear operator on ℓ^2.) Using that $P_n = P_n^* \to I$ strongly, and $V_n^* = V_{-n} \to 0$ strongly on ℓ^2, it follows that A_n^{-1} and their adjoints converge strongly.

In order to prove $\det P_n T^{-1}(a^{-1}) P_n = G[a]^n$ is suffices to prove that

$$\frac{\det A_n}{\det A_{n-1}} = G[a] \tag{34}$$

for $n \geq 1$. For $n = 1$ with $\det A_0 := 1$, this is just the definition of $G[a]$. By noting that $A_{n-1} = P_{n-1} A_n P_{n-1}$ it follows from Cramer's rule that

$$\frac{\det A_{n-1}}{\det A_n} = \left[A_n^{-1}\right]_{n-1,n-1}$$

for $n \geq 2$ while the statement is obvious for $n = 1$. Reformulating the above expression (33) for A_n^{-1} one step further, we have

$$A_n^{-1} = W_n T(\tilde{a}^{-1}) W_n - W_n H(\tilde{a}^{-1}) T^{-1}(a^{-1}) H(a^{-1}) W_n = W_n T^{-1}(\tilde{a}) W_n. \tag{35}$$

Here we use the general formulas

$$P_n T(b) P_n = W_n T(\tilde{b}) W_n, \qquad P_n T(b) V_n = W_n H(\tilde{b}), \qquad V_{-n} T(b) P_n = H(b) W_n.$$

as well as an identity relating the inverses of $T(a^{-1})$ and $T(\tilde{a})$ to each other (which either can be derived from (31) or by using (26), (27)). Due to the definition of the W_n, we see that the lower-right entry of A_n^{-1} does not depend on n for $n \geq 1$, i.e.,

$$[A_n^{-1}]_{n-1,n-1} = [T^{-1}(\tilde{a})]_{00} = 1/G[a],$$

the last equality following from (34) for $n = 1$. This completes the proof of (34) for all n.

(iii) Using (29) it can be seen that $T(e^{-\lambda b})$ is a Fredholm regularizer of $T(e^{\lambda b})$, $\lambda \in [0, 1]$. Due to the stability of the Fredholm index under perturbation, all these operators have Fredholm index zero; hence they are invertible (Coburn's lemma [4], Sec. 2.6). This proves $e^b \in \Phi(B)$. A proof of $G[e^b] = e^{b_0}$ can now be given via an approximation argument and by using Wiener–Hopf factorization (see [4], Prop. 10.4). □

Before stating the main result of this section, we introduce two conditions on a Banach algebra $B \subseteq L^\infty(\mathbb{T})$.

(TC) For all $a \in B$ the operator $K_\mu(a)$ is trace class and $\|K_\mu(a)\|_{\mathcal{C}_1(\ell^2)} \leq C\|a\|_B$.

(HS) For all $a \in B$ the operator $K_\mu(a)$ is Hilbert–Schmidt and $\|K_\mu(a)\|_{\mathcal{C}_2(\ell^2)} \leq C\|a\|_B$.

Propositions 3.2 and 3.3 identify Banach algebras B which satisfy (TC) or (HS), depending naturally on the underlying measure μ (the constant $C = C(\mu)$).

Theorem 4.4. *Let $B \subset L^\infty(\mathbb{T})$ be a suitable Banach algebra.*

(a) *Suppose B satisfies* (TC). *Then for $a \in \Phi(B)$ we have*

$$\lim_{n\to\infty} \frac{\det M_{\mu,n}(a)}{G[a]^n} = E[a], \tag{36}$$

where

$$E[a] = \det\big(T(a^{-1}) M_\mu(a)\big).$$

The constant $E[a]$ is a well-defined operator determinant, and the convergence (36) is uniform in $a \in \Phi(B)$ on compact subsets of $\Phi(B)$.

(b) *Suppose B satisfies* (HS). *Then for $a \in \Phi(B)$ we have*

$$\lim_{n\to\infty} \frac{\det M_{\mu,n}(a)}{G[a]^n \cdot \exp(\operatorname{trace} P_n T(a^{-1}) K_\mu(a) P_n)} = H[a] \tag{37}$$

with

$$H[a] = \det\big(T(a^{-1}) M_\mu(a) e^{-T(a^{-1}) K_\mu(a)}\big).$$

Again, the constant $H[a]$ is a well-defined operator determinant, and the convergence (37) is uniform in $a \in \Phi(B)$ on compact subsets of $\Phi(B)$.

Proof. The first steps in the proof of (a) and (b) are the same. As in the previous proposition define $A_n = P_n T^{-1}(a^{-1}) P_n$. Recall (29) to conclude that

$$T(a) = T^{-1}(a^{-1}) + L(a), \quad L(a) := -T^{-1}(a^{-1}) H(a^{-1}) H(\tilde{a})$$

with $L(a)$ being trace class. The latter follows from the fact that $H(b)$ and $H(\tilde{b})$ are Hilbert–Schmidt for $b \in B \subseteq \mathrm{K}$, while appropriate norm estimates also hold. Moreover, property (b) of the suitability of B implies that the mapping

$$a \in \Phi(B) \mapsto L(a) \in \mathcal{C}_1(\ell^2)$$

is continuous. Now we can write

$$\begin{aligned} M_{\mu,n}(a) &= P_n\big(T^{-1}(a^{-1}) + L(a) + K_\mu(a)\big)P_n \\ &= A_n + P_n\big(L(a) + K_\mu(a)\big)P_n. \end{aligned}$$

Using Proposition 4.3(ii) we obtain

$$\frac{\det M_{\mu,n}(a)}{G[a]^n} = \det\big(P_n + A_n^{-1} P_n\big(L(a) + K_\mu(a)\big)P_n\big). \tag{38}$$

(a) Assume condition (TC). Then $K_\mu(a)$ is trace class, and the mapping $a \in \Phi(a) \mapsto K_\mu(a) \in \mathcal{C}_1(\ell^2)$ is continuous. Consequently, again by Proposition 4.3(ii),

$$\det\big(P_n + A_n^{-1} P_n\big(L(a) + K_\mu(a)\big)P_n\big)$$

converges to the well defined operator determinant

$$\det\big(I + T(a^{-1})\big(L(a) + K_\mu(a)\big)\big),$$

which equals

$$\det\big(T(a^{-1})\big(T(a) + K_\mu(a)\big)\big) = \det\big(T(a^{-1}) M_\mu(a)\big).$$

As to the uniform convergence on compact subset of $\Phi(B)$, it is enough to show that the family of maps

$$a \in \Phi(B) \mapsto \det\big(P_n + A_n^{-1} P_n\big(L(a) + K_\mu(a)\big)P_n\big) \in \mathbb{C}$$

are equi-continuous. To see this we use the equi-continuity of $a \in \Phi(B) \mapsto A_n^{-1} \in \mathcal{L}(\ell^2)$ and the continuity of $a \in \Phi(B) \mapsto L(a) + K_\mu(a) \in \mathcal{C}_1(\ell^2)$ along with fact that $\sup \|A_n^{-1}\| < \infty$ for each $a \in \Phi(B)$. This implies that the maps

$$a \in \Phi(B) \mapsto A_n^{-1} P_n\big(L + K_\mu(a)\big)P_n$$

are equi-continuous and bounded. Finally, in order to pass to the determinant we use the general estimate

$$\big|\det(I + A) - \det(I + C)\big| \le \|A - C\|_1 \exp\big(\max\big\{\|A\|_1, \|C\|_1\big\}\big),$$

which holds for trace class operators A, C.

(b) Now assume condition (HS). In view of (38) introduce

$$C_n = A_n^{-1} P_n\big(L(a) + K_\mu(a)\big)P_n.$$

Then

$$C_n = A_n^{-1} P_n L(a) P_n + P_n T(a^{-1}) K_\mu(a) P_n + D_n$$

with

$$D_n = \left(A_n^{-1} P_n - P_n T\left(a^{-1}\right)\right) K_\mu(a) P_n.$$

From (32) and $P_n = I - Q_n$ we obtain

$$\begin{aligned} A_n^{-1} P_n - P_n T\left(a^{-1}\right) &= -P_n T\left(a^{-1}\right) Q_n - P_n T\left(a^{-1}\right) Q_n \left(Q_n T\left(a^{-1}\right) Q_n\right)^{-1} Q_n T\left(a^{-1}\right)(I - Q_n) \\ &= -P_n T\left(a^{-1}\right) Q_n \left(Q_n T\left(a^{-1}\right) Q_n\right)^{-1} Q_n T\left(a^{-1}\right). \end{aligned}$$

Using the same arguments as in the derivation of (33) and (35), this equals

$$-W_n H\left(\tilde{a}^{-1}\right) T^{-1}\left(a^{-1}\right) V_{-n} T\left(a^{-1}\right),$$

whence

$$D_n = -W_n H\left(\tilde{a}^{-1}\right) T^{-1}\left(a^{-1}\right) V_{-n} T\left(a^{-1}\right) K_\mu(a) P_n.$$

Since $H(\tilde{a}^{-1})$ and $K_\mu(a)$ are each Hilbert–Schmidt, and $V_{-n} \to 0$ strongly, it follows that $D_n \to 0$ in the trace norm. Moreover, from the explicit representation it is seen that the family of mappings $a \in \Phi(B) \mapsto D_n \in \mathcal{C}_1(\ell^2)$ is equi-continuous.

Further, by Proposition 4.3(ii), $A_n^{-1} P_n L(a) P_n \to T(a^{-1}) L(a)$ converges in the trace norm, and the family of maps $a \in \Phi(B) \mapsto A_n^{-1} P_n L(a) P_n \in \mathcal{C}_1(\ell^2)$ is equi-continuous.

In contrast, $P_n T(a^{-1}) K_\mu(a) P_n$ converges only in the Hilbert–Schmidt norm to $T(a^{-1}) K_\mu(a)$, while the mappings $a \in \Phi(B) \mapsto P_n T(a^{-1}) K_\mu(a) P_n \in \mathcal{C}_2(\ell^2)$ are equi-continuous.

We can now conclude that on each compact subset of $\Phi(B)$, the afore-mentioned maps are actually uniformly equi-continuous and uniformly bounded. Hence we have uniform convergence of the corresponding sequences of operators in the trace class or Hilbert–Schmidt norm.

With $C = T(a^{-1}) L(a) + T(a^{-1}) K_\mu(a) = T(a^{-1}) M_\mu(a)$, noting that $L(a) = T(a) - T(a^{-1})^{-1}$, it follows that, as $n \to \infty$,

$$(I + C_n)\mathrm{e}^{-P_n T(a^{-1}) K_\mu(a) P_n} - I \to (I + C)\mathrm{e}^{-T(a^{-1}) K_\mu(a)} - I,$$

uniformly on compact subset of $\Phi(B)$ in trace norm. Consequently,

$$\lim_{n\to\infty} \det\left((I + C_n)\mathrm{e}^{-P_n T(a^{-1}) K_\mu(a) P_n}\right) = \det\left((I + C)\mathrm{e}^{-T(a^{-1}) K_\mu(a)}\right),$$

also uniformly. □

Let us summarize what we have achieved thus far:

Assuming the moment condition (C1), i.e., "$\beta > 1$," we have both the trace class condition (TC) and the Hilbert–Schmidt condition (HS) available (see Proposition 3.2 and 3.3). The easiest way is to assume (TC) and use Theorem 4.4(a) to conclude a limit theorem. However, the trace class conditions are much stronger than the Hilbert–Schmidt conditions, and it is worthwhile to see what can be done assuming only the latter. Then we can apply Theorem 4.4(b), and are left with the computation of traces (which will be done in Proposition 5.1 below). While we get a better result assuming only (HS), the constant expression will be more complicated.

Assuming the moment condition (C2), i.e., "$1/2 < \beta \le 1$," $K_\mu(a)$ will in general not be trace class (see the remark at the end of Section 3). Therefore we are left with Theorem 4.4(b) and the computation of the traces, which in this case is more difficult and will occupy most of the next section.

5. Asymptotics of the trace

As just pointed out, in order to make use of part (b) of Theorem 4.4, we need to evaluate the trace term. We distinguish between the two cases indicated above.

The case of $\beta > 1$ is completely settled by the following proposition, which shows that the trace converges to a constant.

Proposition 5.1. *Assume the moment condition* (C1), *and put* $\sigma = 1/2 \vee 1/\beta$. *Then, for* $a, b \in F\ell^2_\sigma$, *we have*

$$\operatorname{trace}\big(P_n T(b) K_\mu(a) P_n\big) = \tau_\mu(a,b) + \mathrm{o}(1), \quad n \to \infty, \tag{39}$$

where

$$\tau_\mu(a,b) := \sum_{j,k=0}^{\infty} b_{k-j} a_{j-k} (\varrho_{j,k} - 1). \tag{40}$$

The series (40) *converges absolutely. Moreover, the convergence* (39) *is uniform in* (a,b) *on compact subsets of* $F\ell^2_\sigma \times F\ell^2_\sigma$.

Proof. By Proposition 3.2 the operator $K_\mu(a)$ is a Hilbert–Schmidt and hence bounded and linear. Consequently the trace equals

$$\operatorname{trace}\big(P_n T(b) K_\mu(a) P_n\big) = \sum_{j=0}^{\infty} \sum_{k=0}^{n-1} b_{k-j} a_{j-k} (\varrho_{j,k} - 1).$$

We claim that the estimate

$$\sum_{j,k=0}^{\infty} \big| b_{k-j} a_{j-k} (\varrho_{j,k} - 1) \big| \le C \|a\|_{F\ell^2_\sigma} \|b\|_{F\ell^2_\sigma} \tag{41}$$

holds. Indeed, put $\delta = 2\sigma = 1 \vee 2/\beta$, recall $0 < \varrho_{j,k} \le 1$, and split the sum into

$$\sum_{(j,k)\notin \mathcal{I}_\delta} |b_{k-j} a_{j-k}| + \sum_{(j,k)\in \mathcal{I}_\delta} \big| b_{j-k} a_{j-k} (\varrho_{j,k} - 1) \big|,$$

where $\mathcal{I}_\delta$ is defined in (21). Using Lemma 3.1(a) and substituting $m = j - k$ and $\ell = j + k$ we can overestimate this by

$$\sum_{\substack{(m,\ell)\in\mathbb{Z}\times\mathbb{Z}_+ \\ 2|m|^\delta \ge \ell}} |b_{-m} a_m| + \sum_{\substack{(m,\ell)\in\mathbb{Z}\times\mathbb{Z}_+ \\ 2|m|^\delta < \ell}} |b_{-m} a_m| \frac{m^2}{\ell^\beta} \le C \sum_{m=-\infty}^{\infty} |b_{-m} a_m| |m|^\delta + C \sum_{m=-\infty}^{\infty} |b_{-m} a_m| |m|^{2+\delta(1-\beta)}.$$

Since $\delta\beta \ge 2$, we obtain (41) via Cauchy–Schwarz.

The convergence (39) of the trace now follows from (41) by dominated convergence. The absolute convergence of (40) is also a consequence of (41). Finally, again by (41), the mappings

$$\Lambda_n : (a,b) \in F\ell^2_\sigma \times F\ell^2_\sigma \mapsto \operatorname{trace}\big(P_n T(b) K_\mu(a) P_n\big), \quad n \ge 1$$

are equi-continuous. Convergence and equi-continuity imply the uniform convergence on compact subsets. □

We remark that the function $\tau_\mu(a,b)$ is bilinear and continuous in $a, b \in F\ell^2_\sigma$. Formally $\tau_\mu(a,b)$ equals the trace of $T(b)K_\mu(a)$, though note the assumptions made in the proposition are not sufficient to insure $T(b)K_\mu(a)$ is trace class. Indeed, there exists $a \in F\ell^2_\sigma$ such that $K_\mu(a)$ is not trace class (and one can choose $b = 1$). Of course, if $K_\mu(a)$ is trace class, we have equality (and the proposition is a triviality).

Now we turn to the case $1/2 < \beta \le 1$, for which the trace does not converge to a constant. It provides the second order asymptotics of the det $M_{\mu,n}(a)$. In terms of the random matrix interpretation, the asymptotics of the trace gives the shape of the variance for the corresponding linear statistics. We begin with the following estimate.

Lemma 5.2. *Assume the moment condition* (C2), *and put* $\delta = 2\sigma = 2/\beta \vee 3/\gamma$. *Then for* $a, b \in F\ell^2_\sigma$ *it holds*

$$\operatorname{trace}\bigl(P_n T(b) K_\mu(a) P_n\bigr) = -\frac{1}{2}\sum_{m=-\infty}^{\infty} m^2 b_{-m} a_m p^{(\delta)}_{n,m} + E_1(a, b; \delta) + \mathrm{o}(1), \quad n \to \infty. \tag{42}$$

Here E_1 *is some constant and*

$$p^{(\delta)}_{n,m} = \sum_{2|m|^\delta < \ell \le 2n}{}' h_\mu(\ell), \tag{43}$$

where the prime indicates that the summation is taken over all $\ell \in \mathbb{Z}_+$ *with the same parity as* m. *The convergence* (42) *is uniform in* (a, b) *on compact subsets of* $F\ell^2_\sigma \times F\ell^2_\sigma$.

Proof. As in the previous lemma, the operator $K_\mu(a)$ is Hilbert–Schmidt and the trace evaluates to

$$\operatorname{trace}\bigl(P_n T(b) K_\mu(a) P_n\bigr) = \sum_{j=0}^{\infty}\sum_{k=0}^{n-1} b_{k-j} a_{j-k} (\varrho_{j,k} - 1).$$

We can split the double series into

$$\sum_{\substack{(j,k)\notin \mathcal{I}_\delta \\ k<n}} b_{k-j} a_{j-k} (\varrho_{j,k} - 1) \quad \text{and} \quad \sum_{\substack{(j,k)\in \mathcal{I}_\delta \\ k<n}} b_{k-j} a_{j-k} (\varrho_{j,k} - 1), \tag{44}$$

where the first term is dominated by

$$\sum_{(j,k)\notin \mathcal{I}_\delta} |b_{k-j} a_{j-k}| \le C \|a\|_{F\ell^2_\sigma} \|b\|_{F\ell^2_\sigma}.$$

Consequently, the first term in (44) converges as $n \to \infty$ to the constant

$$\sum_{(j,k)\notin \mathcal{I}_\delta} b_{k-j} a_{j-k} (\varrho_{j,k} - 1), \tag{45}$$

and using equi-continuity we see that the convergence is uniform on compact subsets.

For the second term in (44) we bring in the estimate of Lemma 3.1(b),

$$\varrho_{j,k} = 1 - \frac{m^2}{2} h_\mu(\ell) + \mathrm{O}\biggl(\frac{m^4}{\ell^{2\beta}} \vee \frac{|m|^3}{\ell^\gamma} \vee \frac{m^2}{\ell^\varrho}\biggr), \quad (j, k) \in \mathcal{I}_\delta,$$

together with the substitution $\ell = j + k$, $m = j - k$. As to the applicability of this lemma, note that $\delta\varrho > \delta \ge 2/\beta \ge 2$. Hence the second term in (44) equals

$$-\sum_{\substack{(j,k)\in \mathcal{I}_\delta \\ k<n}} b_{-m} a_m \frac{m^2}{2} h_\mu(\ell) + \sum_{\substack{(j,k)\in \mathcal{I}_\delta \\ k<n}} b_{-m} a_m \mathrm{O}\biggl(\frac{m^4}{\ell^{2\beta}} \vee \frac{|m|^3}{\ell^\gamma} \vee \frac{m^2}{\ell^\varrho}\biggr). \tag{46}$$

The error term here can be overestimated by a constant multiple of

$$\sum_{m\in\mathbb{Z}} |b_{-m} a_m| \cdot \bigl(|m|^{4+\delta(1-2\beta)} \vee |m|^{3+\delta(1-\gamma)} \vee |m|^{2+\delta(1-\varrho)}\bigr) \le \|a\|_{F\ell^2_\sigma} \|b\|_{F\ell^2_\sigma}.$$

Here, we first converted the sum over (j, k) to that over $(m, \ell) \in \mathbb{Z} \times \mathbb{Z}_+$ restricted to $2|m|^\delta < \ell$ and then summed over the ℓ variable. After this one notes that our conditions imply that the exponents $4 + \delta(1 - 2\beta)$, $3 + \delta(1 - \gamma)$, and

$2+\delta(1-\varrho)$ are all less than $\delta=2\sigma$. In other words, the error in (46) is dominated by a corresponding absolutely convergent series. As such it converges to the constant

$$\sum_{(j,k)\in\mathcal{I}_\delta} b_{-m}a_m\left(\varrho_{j,k}-1+\frac{m^2}{2}h_\mu(\ell)\right) \tag{47}$$

as $n\to\infty$. In fact, the convergence is uniform on compact subsets of $F\ell^2_\sigma\times F\ell^2_\sigma$, which can be most easily seen by equi-continuity. In view of what follows, the constant $E_1(a,b;\delta)$ is now identified as the sum of (45) and (47).

Turning to the first term in (46), the summation expressed in terms of $(m,\ell)\in\mathbb{Z}\times\mathbb{Z}_+$ is over all indices such that $\ell<2n+m$, $2|m|^\delta<\ell$, and such that the parity of ℓ and m is the same. That is, what we have for the leading order is

$$\sum_{m=-\infty}^{\infty} b_{-m}a_m\frac{m^2}{2}\left(\sideset{}{'}\sum_{2|m|^\delta<\ell<2n+m} h_\mu(\ell)\right) \tag{48}$$

while

$$\sum_{m=-\infty}^{\infty} b_{-m}a_m\frac{m^2}{2}\left(\sideset{}{'}\sum_{2|m|^\delta<\ell\le 2n} h_\mu(\ell)\right) \tag{49}$$

is what is claimed in (42).

We next show that

$$s_{n,m}:=\sideset{}{'}\sum_{2|m|^\delta<\ell<2n+m} m^2h_\mu(\ell)-\sideset{}{'}\sum_{2|m|^\delta<\ell\le 2n} m^2h_\mu(\ell)=\mathrm{O}\left(\frac{|m|^\delta}{n^\varepsilon}\vee\frac{|m|^\delta}{n^\beta}\right), \tag{50}$$

as $n\to\infty$, uniformly in m, where $\varepsilon=\beta+1-3/\delta>0$. This will imply that the difference between (48) and (49) converges (uniformly) to zero as $n\to\infty$.

To see (50) we distinguish four cases:

1. $m>0$ and $2|m|^\delta<2n$. Then $s_{n,m}=\mathrm{O}(m^3/n^\beta)$. Since $m<n^{1/\delta}$ we have

$$\frac{m^3}{n^\beta}\le\frac{m^\delta n^{(3-\delta)/\delta}}{n^\beta}=\frac{m^\delta}{n^\varepsilon}$$

in case $\delta<3$, while the bound is m^δ/n^β in the case $\delta\ge 3$.

2. $m>0$ and $2n\le 2|m|^\delta$. Then $s_{n,m}=\mathrm{O}(m^3/m^{\beta\delta})$, and since $m\ge n^{1/\delta}$, we have

$$\frac{m^3}{m^{\beta\delta}}=\frac{m^\delta}{m^{\beta\delta+\delta-3}}\le\frac{m^\delta}{n^{\beta+1-3/\delta}}=\frac{m^\delta}{n^\varepsilon}.$$

3. $m<0$ and $2|m|^\delta<2n+m$. Then $s_{n,m}=\mathrm{O}(|m|^3/(2n-|m|)^\beta)$, $|m|<(n-|m|/2)^{1/\delta}\le n^{1/\delta}$, and we have

$$\frac{|m|^3}{(n-|m|/2)^\beta}\le\frac{|m|^\delta(n-|m|/2)^{(3-\delta)/\delta}}{(n-|m|/2)^\beta}=\frac{|m|^\delta}{(n-|m|/2)^\varepsilon}\le\frac{|m|^\delta}{(n-n^{1/\delta}/2)^\varepsilon}$$

in case $\delta<3$, or $|m|^\delta/(n-n^{1/\delta}/2)^\beta$ in the case $\delta\ge 3$.

4. $m<0$ and $2n+m\le 2|m|^\delta$. Then $s_{n,m}=\mathrm{O}(|m|^3/|m|^{\beta\delta})$, $n\le|m|^\delta+|m|/2\le 2|m|^\delta$, and

$$\frac{|m|^3}{|m|^{\beta\delta}}=\frac{|m|^\delta}{|m|^{\beta\delta+\delta-3}}\le C\frac{|m|^\delta}{n^{\beta+1-3/\delta}}=C\frac{|m|^\delta}{n^\varepsilon}.$$

From here it follows that difference of (48) and (49) is bounded by a constant multiple of $n^{-\varepsilon\wedge\beta}\|a\|_{F\ell^2_\sigma}\|b\|_{F\ell^2_\sigma}$, and the indicated convergence is uniform in (a,b) even on bounded subsets of $F\ell^2_\sigma \times F\ell^2_\sigma$. The proof is finished. □

Next we estimate the leading term from the previous lemma.

Lemma 5.3. *Assume the moment assumption* (C2), *and define* $p^{(\delta)}_{n,m}$ *for* $\delta > 1$ *by* (43).

(i) *If* $c \in W = F\ell^1$, *then*

$$\sum_{m=-\infty}^{\infty} c_m p^{(\delta)}_{n,m} = \iota_\mu(2n) \sum_{m=-\infty}^{\infty} c_m + \mathrm{o}\bigl(\iota_\mu(2n)\bigr), \quad n \to \infty. \tag{51}$$

(ii) *If* $c \in F\ell^1(\hat{\nu})$ *with* $\hat{\nu}_m = 1 + \iota_\mu(2|m|^\delta)$, *then, with some constant* E_2,

$$\sum_{m=-\infty}^{\infty} c_m p^{(\delta)}_{n,m} = \iota_\mu(2n) \sum_{m=-\infty}^{\infty} c_m + E_2(c;\delta) + \mathrm{o}(1), \quad n \to \infty. \tag{52}$$

The convergence holds uniformly in c *on compact subsets of* W *and* $F\ell^2(\hat{\nu})$, *respectively.*

Proof. First set

$$s^{\pm}_\mu(x) = \sum_{\substack{1\le\ell\le x \\ (-1)^\ell=\pm 1}} h_\mu(\ell).$$

Standard estimates using the assumptions on h_μ and the fact that the functions $s^{\pm}_\mu(x)$ are increasing gives $s^{\pm}_\mu(x) = \iota_\mu(x) + C_\pm + \mathrm{o}(1)$ as $x \to \infty$ for constants $C_\pm$. Granted this, for either point (i) or (ii), we split the sum over even and odd indices. In particular,

$$\begin{aligned}\sum_{m \text{ even}} c_m p^{(\delta)}_{n,m} &= \sum_{m \text{ even}} c_m \max\bigl\{0, s^+_\mu(2n) - s^+_\mu\bigl(2|m|^\delta\bigr)\bigr\} \\ &= s^+_\mu(2n) \sum_{m \text{ even}} c_m - \sum_{m \text{ even}} c_m \min\bigl\{s^+_\mu(2n), s^+_\mu\bigl(2|m|^\delta\bigr)\bigr\}.\end{aligned}$$

The first term on the right-hand side gives one half of the leading asymptotics. Next we show that for part (i), the second term is $\mathrm{o}(s^+_\mu(2n))$, while for part (ii) the second term is a constant plus $\mathrm{o}(1)$.

Indeed, for part (i), we write the second term as

$$s^+_\mu(2n) \sum_{m \text{ even}} c_m \min\left\{1, \frac{s^+_\mu(2|m|^\delta)}{s^+_\mu(2n)}\right\}.$$

This renormalized series is dominated by the series $\sum |c_m|$. Moreover, for each fixed m, the minimum converges to zero as $n \to \infty$. Dominated convergence then implies that the series is $\mathrm{o}(1)$ as $n \to \infty$. Similar considerations can be carried out for the odd term, concluding the proof of part (i).

As for part (ii), take again the even terms:

$$\sum_{m \text{ even}} c_m \min\bigl\{s^+_\mu(2n), s^+_\mu\bigl(2|m|^\delta\bigr)\bigr\}.$$

This sum is now dominated by (a constant times)

$$\sum_{m=-\infty}^{\infty} |c_m|\bigl(1 + \iota_\mu\bigl(2|m|^\delta\bigr)\bigr) < \infty, \tag{53}$$

while for each fixed m, the minimum converges to $s_\mu^+(2|m|^\delta)$ as $n \to \infty$. So dominated convergence yields that the above equals

$$\sum_{m \text{ even}} c_m s_\mu^+\big(2|m|^\delta\big) + \mathrm{o}(1).$$

The terms involving the summation over odd m give a similar contribution, and collecting everything we arrive at, in case (ii):

$$\sum c_m p_{n,m}^{(\delta)} = \sum_{m \text{ even}} c_m\big(s_\mu^+(2n) - s_\mu^+\big(2|m|^\delta\big)\big) + \sum_{m \text{ odd}} c_m\big(s_\mu^-(2n) - s_\mu^-\big(2|m|^\delta\big)\big) + \mathrm{o}(1).$$

From here the constant

$$E_2(c;\delta) = C_+ \sum_{m \text{ even}} c_m + C_- \sum_{m \text{ odd}} c_m - \sum_{m=-\infty}^{\infty} c_m \sum_{1\le \ell \le 2|m|^\delta}{}' h_\mu(\ell)$$

is identified. The uniform convergence on compacts is seen by using the equi-continuity of the corresponding mappings. □

We now combine the previous two lemmas into the following theorem. Notice that part (i) will be used to prove Theorem 1.1, while part (ii) is used to show Theorem 1.3(a).

Theorem 5.4. *Assume the moment condition* (C2), *and put* $\sigma = 1/\beta \vee 3/(2\gamma)$.

(i) *If* $a, b \in F\ell^2_\sigma$, *then*

$$\operatorname{trace}\big(P_n T(b) K_\mu(a) P_n\big) = \Omega(a,b) \cdot \iota_\mu(2n) + \mathrm{o}\big(\iota_\mu(2n)\big), \quad n \to \infty, \tag{54}$$

where

$$\Omega(a,b) = -\frac{1}{2}\sum_{m=-\infty}^{\infty} m^2 a_m b_{-m} = -\frac{1}{4\pi}\int_0^{2\pi} a'\big(\mathrm{e}^{\mathrm{i}t}\big) b'\big(\mathrm{e}^{\mathrm{i}t}\big)\,\mathrm{d}t,$$

and the convergence (54) *is uniform in* (a,b) *on compact subsets of* $F\ell^2_\sigma \times F\ell^2_\sigma$.

(ii) *Let* $B = F\ell^2_\sigma \cap F\ell^2(\nu)$ *with* $\nu_m = \sqrt{1 + m^2 \iota_\mu(2|m|^{2\sigma})}$. *Then, for* $a, b \in B$,

$$\operatorname{trace}\big(P_n T(b) K_\mu(a) P_n\big) = \Omega(a,b) \cdot \iota_\mu(2n) + C_\mu(a,b) + \mathrm{o}(1), \quad n \to \infty, \tag{55}$$

with a certain constant $C_\mu(a,b)$. *The convergence* (54) *is uniform in* (a,b) *on compact subsets of* $B \times B$.

Proof. (i) We employ Lemma 5.2 and Lemma 5.3(i) with $c_m = m^2 b_{-m} a_m$ and $\delta = 2\sigma$. Since $\sigma \ge 1/\beta \ge 1$, we obtain from Cauchy–Schwarz that $c \in F\ell^1_{2\sigma-2} \subseteq W$. Hence

$$\operatorname{trace}\big(P_n T(b) K_\mu(a) P_n\big) = -\frac{\iota_\mu(2n)}{2}\sum_{m=-\infty}^{\infty} c_m + \mathrm{o}\big(\iota_\mu(2n)\big), \quad n \to \infty,$$

with the convergence being uniform in a, b on compact subsets of $F\ell^2_\sigma$. The computation of the constant $\Omega(a,b)$ is straightforward.

(ii) Lemma 5.2 is applied without any change. This produces the constant factor E_1 which could be neglected in case (i). Lemma 5.3(ii) is now applicable because $a, b \in F\ell^2(\nu)$ along with Cauchy–Schwarz implies that $c \in F\ell^1(\hat{\nu})$.

We thus obtain the asymptotics (52). Combined with Lemma 5.2 we arrive at (55) with the overall constant evaluated from E_1 and E_2,

$$C_\mu(a,b) = \sum_{j,k=0}^{\infty} b_{k-j}a_{j-k}\left(\varrho_{j,k} - 1 + \frac{(j-k)^2}{2}h_\mu(j+k)\right) - \frac{C_+}{2}\sum_{m \text{ even}} m^2 a_m b_{-m} - \frac{C_-}{2}\sum_{m \text{ odd}} m^2 a_m b_{-m}. \quad (56)$$

The constant $C_\pm$ were defined at the beginning of the proof of Lemma 5.3. The absolute convergence of the above series is guaranteed by estimates on a_m and b_m that follow from the choice of B. □

6. Limit theorems: the case $\beta > 1$ (C1)

As pointed out at the end of Section 4, we can proceed in two ways, by using either Theorem 4.4 (a) or (b) depending whether we have the trace class (TC) or Hilbert–Schmidt (HS) condition available. In turn, Propositions 3.2 and 3.3 indicate which condition is in effect given the underlying assumptions. We start with the proof of Theorem 1.3(b).

Put $B = F\ell^1_\sigma$, or $B = F\ell^2_{\sigma+\varepsilon}$, $\varepsilon > 0$ with $\sigma = 1 \vee 2/\beta$. Then Proposition 3.3 implies that B satisfies the trace class condition (TC), and Proposition 4.2 shows that the Banach algebra B is suitable. Now apply Theorem 4.4(a) in order to get (15) in Theorem 1.3(b). In particular, we obtain the correct identification of the constant $E[a]$ as a well-defined operator determinant.

As for the constant $G[a]$, which is given by (30) in Proposition 4.3(iii), notice first that standard Toeplitz theory implies that $a \in C(\mathbb{T})$ does not vanish on $\mathbb{T}$ and has winding zero. (Recall that $a \in B \subset C(\mathbb{T})$ and the $T(a)$ is assumed to be invertible.) Hence there exists a continuous logarithm $\log a$. Because the Banach algebra B under consideration contain all smooth functions and contains the set of Laurent polynomials as a dense subset, an approximation argument implies that $\log a \in B$. Now Proposition 4.3(iii) implies that $G[a]$ is also given by (14).

Next, put $B = L^\infty(\mathbb{T}) \cap F\ell^2_{1/2}$ ($\beta \geq 2$) or $B = F\ell^2_{1/\beta}$ ($1 < \beta < 2$). Again suitability of B is guaranteed by Proposition 4.2, and Proposition 3.2 implies (HS). Now we can use Theorem 4.4(b), and we are left with the asymptotics of the trace, which is settled by Proposition 5.1. We obtain the same convergence (15) in Theorem 1.3(b) under the stated (more general) conditions, but the constant $E[a]$ must be identified as

$$E[a] = \mathrm{e}^{\tau_\mu(a,a^{-1})}\det\bigl(T\bigl(a^{-1}\bigr)M_\mu(a)\mathrm{e}^{-T(a^{-1})K_\mu(a)}\bigr).$$

Clearly, if a satisfies the stronger conditions, then both expressions for $E[a]$ coincide (see also the remark after Proposition 5.1). This concludes the proof of Theorem 1.3(b).

For the main application (Theorem 1.2), the behavior of the (centered) linear statistic $X_{f,n} - nf_0 = X_{f-f_0,n}$ is accessed through considering symbols $a_\lambda = \mathrm{e}^{\mathrm{i}\lambda(f-f_0)}$. Notice that Proposition 4.3(iii) implies $a_\lambda \in \Phi(B)$ and $G[a_\lambda] = 1$. Applying what we have just proved (Theorem 1.3(b)) and (2) we immediately obtain

$$\lim_{n\to\infty} \mathbb{E}_{\mathfrak{m},n}\bigl[\mathrm{e}^{\mathrm{i}\lambda(X_{f,n}-nf_0)}\bigr] = E(f,\lambda) \quad (57)$$

with

$$E(f,\lambda) := \mathrm{e}^{\tau_\mu(a_\lambda^{-1},a_\lambda)}\det\bigl(T\bigl(a_\lambda^{-1}\bigr)M_\mu(a_\lambda)\mathrm{e}^{-T(a_\lambda^{-1})K_\mu(a_\lambda)}\bigr) \quad (58)$$

under the conditions stated in Theorem 1.2(a). The convergence (57) is locally uniform in λ. Hence $E(f,\lambda)$ is analytic in λ and $E(f,0) = 1$. This implies that $E(f,\lambda)$ is a proper moment generating function, and hence $X_{f,n} - nf_0$ converges in distribution to some random variable $\mathcal{Z}$.

In order to identify the mean and the variance of $\mathcal{Z}$ we rewrite (57) as

$$\lim_{n\to\infty} \det M_{\mu,n}\bigl(\mathrm{e}^{\lambda b}\bigr) = \mathrm{e}^{\mathrm{i}\lambda c_1 - \lambda^2 c_2/2+\cdots}$$

abbreviating $b = i(f - f_0)$, where c_1 stands for the means and c_2 for the variance, and the series expansion hold in a neighborhood of zero. Due to locally uniform convergence we can take the first and second derivative of the logarithm and put $\lambda = 0$ to get c_1 and c_2. Using the formula $(\log\det F(\lambda))' = \operatorname{trace} F'(\lambda)F^{-1}(\lambda)$ we obtain

$$c_1 = \lim_{n\to\infty} \operatorname{trace} M_{\mu,n}(f - f_0) = 0,$$

$$c_2 = -\lim_{n\to\infty} \operatorname{trace}\bigl(M_{\mu,n}(b^2) - \bigl(M_{\mu,n}(b)\bigr)^2\bigr).$$

We decompose the $M_{\mu,n}$ matrices into the Toeplitz matrices T_n and the error terms $K_{\mu,n}$ and use a general formula for Toeplitz matrices,

$$T_n(\phi\psi) = T_n(\phi)T_n(\psi) + P_n H(\phi)H(\tilde{\psi})P_n + W_n H(\tilde{\phi})H(\psi)W_n,$$

see Section 4 for the notation. Noting that the trace of $K_{\mu,n}(b^2)$ equals zero it follows that

$$\begin{aligned} c_2 &= -\lim_{n\to\infty} \operatorname{trace}\bigl(P_n H(b)H(\tilde{b})P_n + W_n H(\tilde{b})H(b)W_n - 2T_n(b)K_{\mu,n}(b) - K_{\mu,n}(b)^2\bigr) \\ &= -2\sum_{k=1}^{\infty} k b_k b_{-k} + 2\sum_{j,k=0}^{\infty} (\varrho_{j,k} - 1) b_{j-k} b_{k-j} + \sum_{j,k=0}^{\infty} (\varrho_{j,k} - 1)^2 b_{j-k} b_{k-j} \\ &= 2\sum_{k=1}^{\infty} k f_k f_{-k} + \sum_{j,k=0}^{\infty} \bigl(1 - \varrho_{j,k}^2\bigr) f_{j-k} f_{k-j}. \end{aligned} \tag{59}$$

Notice that for f real, $f_{-k} = \overline{f_k}$, so that in any case $c_2 > 0$ unless f is constant. The absolute convergence of the above series follows from the same estimate used in the proof of Proposition 5.1 for (41). Recall that it is assumed $f \in F\ell^2_\sigma$, $\sigma = 1/2 \vee 1/\beta$. This concludes the proof of the first part of Theorem 1.2.

For second part, notice that under the stronger conditions, the constant simplifies to

$$E(f,\lambda) = \det\bigl(T\bigl(\mathrm{e}^{-\mathrm{i}\lambda(f-f_0)}\bigr)M_\mu\bigl(\mathrm{e}^{\mathrm{i}\lambda(f-f_0)}\bigr)\bigr) = \det\bigl(T\bigl(\mathrm{e}^{-\mathrm{i}\lambda f}\bigr)M_\mu\bigl(\mathrm{e}^{\mathrm{i}\lambda f}\bigr)\bigr). \tag{60}$$

From this expression, via differentiation, the formulas for the zero mean and the variance (59) can be obtained as well. What exactly $\mathcal{Z}$ is though is hard to understand from (58) or (60). The following is the best we have; it completes the proof of Theorem 1.2.

Proposition 6.1. *Let $\beta > 1$, $\sigma = 1 \vee 2/\beta$ and assume either $b \in F\ell^1_\sigma$ or $b \in F\ell^2_{\sigma+\varepsilon}$, $\varepsilon > 0$. Then there exists $\delta > 0$ such that for $\lambda \in \mathbb{C}$ with $|\lambda| < \delta$ it holds that*

$$\det\bigl(T\bigl(\mathrm{e}^{-\lambda b}\bigr)M_\mu\bigl(\mathrm{e}^{\lambda b}\bigr)\bigr) = \exp\left(\frac{\lambda^2}{2}\operatorname{trace}\bigl(H(b)H(\tilde{b})\bigr) + \sum_{n=2}^{\infty} \frac{\lambda^n}{n!}\operatorname{trace}(B_n)\right), \tag{61}$$

where the (trace class) operators B_n are defined by the recursion

$$B_{n+1} = M_\mu\bigl(b^{n+1}\bigr) - \sum_{k=1}^{n} \binom{n}{k} B_{n+1-k} M_\mu\bigl(b^k\bigr), \quad n \ge 0.$$

Ahead of the proof, we write out the first couple of B_n's. With $M_k = M_\mu(b^k)$ we obtain $B_1 = M_1$,

$$\begin{aligned} B_2 &= M_2 - M_1^2, \\ B_3 &= M_3 - 2M_2M_1 - M_1M_2 + 2M_1^3, \\ B_4 &= M_4 - 3M_3M_1 - M_1M_3 - 2M_2^2 + 6M_2M_1^2 + 3M_1M_2M_1 + 3M_1^2M_2 - 6M_1^4. \end{aligned}$$

When μ is the unit mass at 1, then $M_\mu(b) = T(b)$ and one has that $\log\det(T(\mathrm{e}^{-\lambda b})T(\mathrm{e}^{\lambda b}))$ equals $\lambda^2 \operatorname{trace}(H(b)H(\tilde{b}))$ (according to the Szegö–Widom limit theorem). That is, we have the above expressions with M_k replaced by $T_k =$

$T(b^k)$ while at the same time trace $B_2 = \operatorname{trace}(H(b)H(\tilde{b}))$ and trace $B_m = 0$ for all $m \geq 3$. (This means that the cumulants of $\mathcal{Z}$ of order three and higher are vanishing.) Back in the general case, we can subtract from the B_k given by the above formulas the corresponding expressions for the special case $M_k = T_k$ and then take traces. Substituting $M_k = T_k + K_k$ with $K_k = K(b^k)$, yields

$$\begin{aligned}
\operatorname{trace}(B_2) &= \operatorname{trace}\big(H(b)H(\tilde{b})\big) - \operatorname{trace}\big(2T_1K_1 + K_1^2\big),\\
\operatorname{trace}(B_3) &= -3\operatorname{trace}(K_2T_1 + K_1T_2 + K_2K_1) + 2\operatorname{trace}\big(3K_1^2T_1 + 3K_1T_1^2 + K_1^3\big),\\
\operatorname{trace}(B_4) &= -4\operatorname{trace}(T_3K_1 + K_3T_1 + K_3K_1) - 2\operatorname{trace}\big(2T_2K_2 + K_2^2\big)\\
&\quad + 12\operatorname{trace}\big(T_2T_1K_1 + T_2K_1T_1 + T_2K_1^2 + K_2T_1^2 + K_2T_1K_1 + K_2K_1T_1 + K_2K_1^2\big)\\
&\quad - 6\operatorname{trace}\big(4T_1^3K_1 + 4T_1^4K_1^2 + 2T_1K_1T_1K_1 + 4T_1K_1^3 + K_1^4\big).
\end{aligned}$$

All products under the traces are trace class operators and thus each of the above objects can be computed explicitly in terms of infinite sums. Still, the expressions become increasingly intractable, and we do not see how further simplifications are possible.

Proof of Proposition 6.1. Set $a_\lambda = e^{\lambda b}$ and split the determinant $E[a_\lambda] = \det T(a_\lambda^{-1})M_\mu(a_\lambda)$ into two parts $E[a_\lambda] = E_1(\lambda)E_2(\lambda)$ where

$$E_1(\lambda) = \det T\big(a_\lambda^{-1}\big)e^{\lambda T(b)}, \qquad E_2(\lambda) = \det e^{-\lambda T(b)} M_\mu(a_\lambda).$$

First of all, both expressions are well defined because the expressions under the determinant are of the form identity plus trace class. Indeed, this has been shown for $T(a_\lambda^{-1})e^{\lambda T(b)}$ in [7], Prop. 7.1. Now observe that $M_\mu(a_\lambda)$ is a trace class perturbation of $T(a_\lambda)$.

It has been shown in [8], Sec. 3 (see also the proof of Thm. 2.5 in [3]) that

$$E_1(\lambda) = \exp\left(\frac{\lambda^2}{2}\operatorname{trace}\big(H(b)H(\tilde{b})\big)\right).$$

It is straightforward to verify that $E_2(\lambda)$ depends analytically on λ (see again [7,8]). Assume now that $|\lambda|$ is sufficiently small such that $M_\mu(a_\lambda)$, being close to the identity operator, is invertible and hence the determinants $E_2(\lambda)$ are nonzero. Notice that $E_2(0) = 1$, whence there is no problem of defining a logarithm in a small neighborhood of zero,

$$f(\lambda) := \log\det e^{-\lambda T(b)} M_\mu(a_\lambda).$$

Recall that for invertible analytic operator-valued functions $F(\lambda)$ of the form identity plus trace class we have the well-known formula $(\log\det F(\lambda))' = \operatorname{trace} F'(\lambda)F^{-1}(\lambda)$. As a consequence, for invertible $A(\lambda)$ and $B(\lambda)$, whose product is identity plus trace class, we have

$$\big(\log\det A(\lambda)B(\lambda)\big)' = \operatorname{trace}\big(A^{-1}(\lambda)A'(\lambda) + B'(\lambda)B^{-1}(\lambda)\big). \tag{62}$$

From this we obtain

$$f'(\lambda) = \operatorname{trace}\big(M_\mu(a_\lambda)'M_\mu^{-1}(a_\lambda) - T(b)\big).$$

For small $|\lambda|$ introduce the well-defined analytic function $B(\lambda)$ defined by $B(0) = 0$ and

$$B'(\lambda) = M_\mu(a_\lambda)'M_\mu^{-1}(a_\lambda).$$

Writing out this relation in terms of power series (with $B(\lambda) = \sum_{k=1}^\infty \lambda^k B_k/k!$) it follows that

$$\left(\sum_{k=0}^\infty \frac{\lambda^k B_{k+1}}{k!}\right)\left(\sum_{k=0}^\infty \frac{\lambda^k M_\mu(b^k)}{k!}\right) = \sum_{n=0}^\infty \frac{\lambda^n M_\mu(b^{n+1})}{n!}.$$

Inspection of the nth coefficient ($n \geq 0$) produces

$$M_\mu\big(b^{n+1}\big) = B_{n+1} + \sum_{k=1}^{n} \binom{n}{k} B_{n+1-k} M_\mu\big(b^k\big)$$

which implies the recursion. Noting that $f(0)=0$, $B(0)=0$, and $f'(\lambda) = \operatorname{trace}(B'(\lambda) - T(b))$ yields

$$E_2(\lambda) = \det \mathrm{e}^{-\lambda M_\mu(b)} M_\mu(a_\lambda) = \exp\big(\operatorname{trace}\big(B(\lambda) - \lambda T(b)\big)\big).$$

Since we have $B_1 = M_\mu(b)$ from the recursion and trace $K_\mu(b) = 0$ ($\varrho_{kk} = 1$) the proof is finished. □

7. Limit theorems: the case 1/2 < β ≤ 1 (C2)

We first prove Theorem 1.3(a). Put $B = F\ell^2(\nu)$ with the conditions on ν stated there. It follows immediately that $B \subseteq F\ell^2_\sigma$ with $\sigma \geq 1/\beta \geq 1$. Hence by Proposition 3.2 the Hilbert–Schmidt condition (HS) holds. Moreover, Proposition 4.2 implies that B is a suitable Banach algebra. Hence we can use Theorem 4.4(b) and obtain (37) with the constant $H[a]$. We are left with determining the asymptotics of the trace of $P_n T(a^{-1}) K_\mu(a) P_n$, for which we can use Theorem 5.4(ii). Therein our Banach algebra is continuously embedded into the Banach space $F\ell^2_\sigma \cap F\ell^2(\nu)$ (with possibly different ν). With $b = a^{-1}$ the asymptotics equals $\Omega(a, a^{-1}) \cdot \iota_\mu(2n) + C_\mu(a, a^{-1}) + \mathrm{o}(1)$ with

$$\Omega[a] := \Omega\big(a, a^{-1}\big) = -\frac{1}{4\pi} \int_0^{2\pi} a'\big(\mathrm{e}^{\mathrm{i}t}\big)\big(a^{-1}\big(\mathrm{e}^{\mathrm{i}t}\big)\big)' \mathrm{d}t = \frac{1}{4\pi} \int_0^{2\pi} \left(\frac{a'(\mathrm{e}^{\mathrm{i}t})}{a(\mathrm{e}^{\mathrm{i}t})}\right)^2 \mathrm{d}t.$$

This gives the correct constant in (14). As for the constant $F[a]$ in (13) we remark that

$$F[a] = \mathrm{e}^{C_\mu(a,a^{-1})} \det\big(T\big(a^{-1}\big) M_\mu(a) \mathrm{e}^{-T(a^{-1})K_\mu(a)}\big), \tag{63}$$

where $C_\mu(a, a^{-1})$ is given by (56), but we make no attempt to simplify the expression. Notice that both Theorem 5.4(ii) and Proposition 4.2(iv) require the rather complicated Banach algebra $B = F\ell^2(\nu)$.

We now turn to the proof of Theorem 1.1, assuming that $B = F\ell^2_\sigma$ with $\sigma = 1/\beta \vee 3/(2\gamma)$. There is no change in the applicability of Theorem 4.4(b), though the function to which it is applied is appropriately rescaled in n. This is where the statements about uniform convergence are needed.

We replace $X_{f,n}$ with $f \in B$ by

$$X^{\mathrm{scal}}_{f,n} := \frac{X_{f,n} - nf_0}{\sqrt{\iota_\mu(2n)}} = X_{g_n,n}, \qquad g_n\big(\mathrm{e}^{\mathrm{i}x}\big) := \frac{f(\mathrm{e}^{\mathrm{i}x}) - f_0}{\sqrt{\iota_\mu(2n)}}, \tag{64}$$

and have that

$$\mathbb{E}_{\mathfrak{m},n}\big[\mathrm{e}^{\mathrm{i}\lambda X^{\mathrm{scal}}_{f,n}}\big] = \det M_{\mu,n}(a_{\lambda,n})$$

with $a_{\lambda,n} = \mathrm{e}^{\mathrm{i}\lambda g_n}$. Since $\iota_\mu(2n) \to \infty$, the elements g_n ($n \in \mathbb{N}$) lie in a compact subset of B, and so $a_{\lambda,n}$ lie in a compact subset of $\Phi(B)$ (recall Proposition 4.3(iii)).

By Theorem 4.4(b)

$$\lim_{n\to\infty} \frac{\det M_{\mu,n}(a_{\lambda,n})}{G[a_{\lambda,n}]^n \cdot \exp(\operatorname{trace} P_n T(a^{-1}_{\lambda,n}) K_\mu(a_{\lambda,n}) P_n)} = \lim_{n\to\infty} H[a_{\lambda,n}],$$

due to uniform convergence on compact subsets. The regularized determinant $H[a_{\lambda,n}]$ converges to $H[1] = 1$ since $T(a^{-1}_{\lambda,n}) K_\mu(a_{\lambda,n}) \to T(1) K_\mu(1) = 0$ in the Hilbert–Schmidt norm. Here we have to use Proposition 3.3 and the estimate implied by (HS).

Again by Proposition 4.3(iii), $G[a_{\lambda,n}] = 1$. To evaluate the trace we will use Theorem 5.4(i). Define

$$h = \mathrm{i}\lambda(f - f_0) \quad \text{and} \quad s_n = \sqrt{\iota_\mu(2n)}$$

and introduce the functions $p_n, q_n \in B$ via series expansion

$$a_{\lambda,n} = \mathrm{e}^{h/s_n} = 1 + h/s_n + p_n/s_n^2, \qquad a_{\lambda,n}^{-1} = \mathrm{e}^{-h/s_n} = 1 - h/s_n + q_n/s_n^2.$$

Notice immediately that $p_n \to h^2/2$ and $q_n \to h^2/2$ in the norm of B. Denoting $t_n(b,a) = \mathrm{trace}(P_n T(b) K_\mu(a) P_n)$ we have that

$$t_n\big(a_{\lambda,n}^{-1}, a_{\lambda,n}\big) = -\frac{t_n(h,h)}{s_n^2} + \frac{-t_n(h,p_n) + t_n(q_n,h) + s_n^{-1} t_n(p_n,q_n)}{s_n^3}$$

because in general $t_n(b,1) = t_n(1,a) = 0$. Theorem 5.4(i) says that for $a, b \in B$ we have $t_n(b,a) = \Omega(a,b) s_n^2 + \mathrm{o}(s_n^2)$ and that the convergence is uniform on compact sets. Hence, applying this to all of the above expressions involving t_n and using that p_n and q_n are from compact subsets of B, it follows that

$$\lim_{n\to\infty} t_n\big(a_{\lambda,n}^{-1}, a_{\lambda,n}\big) = -\Omega(h,h) = -\frac{\lambda^2}{2} \sum_{k=-\infty}^{\infty} k^2 f_k f_{-k} = -\frac{\lambda^2}{4\pi} \int_0^{2\pi} \big(f'\big(\mathrm{e}^{\mathrm{i}x}\big)\big)^2 \,\mathrm{d}x.$$

This implies, uniformly on bounded sets of λ and compact sets of $f \in B$,

$$\lim_{n\to\infty} \mathbb{E}_{\mathfrak{m},n}\big[\mathrm{e}^{\mathrm{i}\lambda X_f^{\mathrm{scal}}}\big] = \exp\bigg(-\frac{\lambda^2}{2} \sum_{k\in\mathbb{Z}} k^2 f_k f_{-k}\bigg) = \exp\bigg(-\frac{\lambda^2}{4\pi} \int_0^{2\pi} \big(f'\big(\mathrm{e}^{\mathrm{i}x}\big)\big)^2 \,\mathrm{d}x\bigg) \tag{65}$$

completing the proof of Theorem 1.1.

Appendix: On the Toeplitz ∘ Hankel formula

We wish to compute the integral

$$\mathcal{I}_{\mathfrak{m},n}(\varphi) = \frac{1}{Z_{\mathfrak{m},n}} \int_{\mathbb{C}^n} \prod_{k=1}^{n} \varphi(\arg z_k) \prod_{k<\ell} |z_k - z_\ell|^2 \prod_{k=1}^{n} \mathrm{d}\mathfrak{m}(z_k),$$

where $\mathfrak{m}$ is radial and $Z_{\mathfrak{m},n}$ is chosen so that $\mathcal{I}_{\mathfrak{m},n}(1) = 1$.

To begin, write

$$\prod |z_k - z_\ell|^2 = \det\big[\big[z_k^{\ell-1}\big] \cdot \big[\bar{z}_k^{\ell-1}\big]^T\big],$$

where $[z_k^{\ell-1}]$ denotes the $n \times n$ matrix with $z_k^{\ell-1}$ in row k and column ℓ. That is to say,

$$\prod |z_k - z_\ell|^2 = \det \begin{bmatrix} n & \sum_{k=1}^n \bar{z}_k & \sum_{k=1}^n \bar{z}_k^2 & \cdots \\ \sum_{k=1}^n z_k & \sum_{k=1}^n z_k \bar{z}_k & \sum_{k=1}^n z_k \bar{z}_k^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Now expand the first column on the right-hand side via the linearity of the determinant, writing it as sum of n determinants with first column $[1, z_k, z_k^2, \ldots, z_k^{n-1}]$. By the product structure of $\prod \varphi(\arg z_k)\,\mathrm{d}\mathfrak{m}(z_k)$ each of the resulting n integrals are the same. Thus, we can replace the $\prod |z_k - z_\ell|^2$ in the measure with

$$\det \begin{bmatrix} 1 & \sum_{k=1}^n \bar{z}_k & \sum_{k=1}^n \bar{z}_k^2 & \cdots \\ z_1 & \sum_{k=1}^n z_k \bar{z}_k & \sum_{k=1}^n z_k \bar{z}_k^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \det \begin{bmatrix} 1 & \sum_{k=2}^n \bar{z}_k & \sum_{k=2}^n \bar{z}_k^2 & \cdots \\ z_1 & \sum_{k=2}^n z_k \bar{z}_k & \sum_{k=2}^n z_k \bar{z}_k^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

at the cost of introducing a constant factor which may be absorbed into $Z_{\mathfrak{m},n}$. This procedure may be repeated, and after the nth iteration we conclude that

$$\mathcal{I}_{\mathfrak{m},n}(\varphi) = \frac{1}{Z_{\mathfrak{m},n}} \int_{\mathbb{C}^n} \prod_{k=1}^{n} \varphi(\arg z_k) \det\big[z_k^{\ell-1} \bar{z}_k^{k-1}\big]_{1\le k,\ell\le n} \prod_{k=1}^{n} \mathrm{d}\mathfrak{m}(z_k)$$
$$= \frac{1}{\tilde{Z}_{\mathfrak{m},n}} \det\bigg[\frac{1}{2\pi} \int_{\mathbb{C}} \varphi(\arg z) z^{\ell} \bar{z}^{k} \,\mathrm{d}\mathfrak{m}(z)\bigg]_{0\le k,\ell\le n-1},$$

after using the linearity of the determinant once more. And, as

$$\frac{1}{2\pi} \int_{\mathbb{C}} \varphi(\arg z) z^{\ell} \bar{z}^{k} \,\mathrm{d}\mathfrak{m}(z) = \varphi_{k-\ell} \int_0^{\infty} r^{k+\ell} \,\mathrm{d}\mu(r) = \varphi_{k-\ell} m_{k+\ell},$$

setting $\varphi \equiv 1$ we find that $\tilde{Z}_{\mathfrak{m},n} = \prod_{k=0}^{n-1} m_{2k}$, and so formula (2).

Acknowledgments

The work of the first named author was supported in part by NSF Grant DMS-09-01434; that of the second by NSF Grant DMS-06-45756. We would also like to thank the anonymous referee for several thoughtful comments on an earlier draft.

References

[1] Y. Ameur, H. Hedenmalm and N. Makarov. Random normal matrices and ward identities. Preprint, 2011. Available at arXiv:1109.5941.
[2] Z. D. Bai. Circular law. *Ann. Probab.* **25** (1997) 494–529. MR1428519
[3] E. Basor and T. Ehrhardt. Asymptotic formulas for determinants of a sum of finite Toeplitz and Hankel matrices. *Math. Nachr.* **228** (2001) 5–45. MR1845906
[4] A. Böttcher and B. Silbermann. *Analysis of Toeplitz Operators*, 2nd edition. Springer, Berlin, 2006. MR2223704
[5] S.-J. Chen and J. D. Vaaler. The distribution of values of Mahler's measure. *J. Reine Angew. Math.* **540** (2001) 1–47. MR1868596
[6] P. Diaconis and S. Evans. Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.* **353** (2001) 2615–2633. MR1828463
[7] T. Ehrhardt. A new algebraic approach to the Szegö–Widom limit theorem. *Acta Math. Hungar.* **99** (2003) 233–261. MR1973097
[8] T. Ehrhardt. A generalization of Pincus' formula and Toeplitz operator determinants. *Arch. Math. (Basel)* **80** (2003) 302–309. MR1981184
[9] P. J. Forrester. Fluctuation formula for complex random matrices. *J. Phys. A: Math. and General* **32** (1999) 159–163. MR1687948
[10] J. Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** (1965) 440–449. MR0173726
[11] I. Gohberg and M. G. Krein. *Introduction to the Theory of Linear Nonselfadjoint Operators on Hilbert Space*. *Transl. Math. Monographs* **18**. Amer. Math. Soc., Providence, RI, 1969. MR0246142
[12] J. M. Hammersley. The zeros of a random polynomial. In *Proc. of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II* 89–111. Univ. California Press, Berkeley and Los Angeles, 1956. MR0084888
[13] J. B. Hough, M. Krishnapur, Y. Peres and B. Virág. Determinantal processes and independence. *Probab. Surv.* **3** (2006) 206–229. MR2216966
[14] C. P. Hughes, J. P. Keating and N. O'Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.* **220** (2001) 429–451. MR1844632
[15] K. Johansson. On random matrices from the compact classical groups. *Ann. Math.* **145** (1997) 519–545. MR1454702
[16] A. Y. Karlovich. Some algebras of functions with Fourier coefficients in weighted Orlicz sequence spaces. In *Operator Theoretical Methods and Applications to Math. Physics* 287–296. *Operator Theory: Advances and Applications* **147**. Birkhäuser, Basel, 2004. MR2053695
[17] M. Krishnapur. From random matrices to random analytic functions. *Ann. Probab.* **37** (2009) 314–346. MR2489167
[18] O. Macchi. The coincidence approach to stochastic point processes. *Adv. Appl. Probab.* **7** (1975) 83–122. MR0380979
[19] Y. Peres and B. Virág. Zeros of the i.i.d. Gaussian power series: A conformally invariant determinantal process. *Acta Math.* **194** (2005) 1–35. MR2231337
[20] B. Rider. Deviations from the circular law. *Probab. Theory Related Fields* **130** (2004) 337–367. MR2095933
[21] B. Rider and B. Virág. The noise in the circular law and the Gaussian free field. *Int. Math. Res. Not.* **2007** (2007) Art. ID rnm006-32. MR2361453
[22] A. Soshnikov. Determinantal random fields. *Russian Math. Surveys* **55** (2000) 923–975. MR1799012
[23] A. Soshnikov. Gaussian limits for determinantal random point fields. *Ann. Probab.* **30** (2002) 171–181. MR1894104
[24] H. Widom. Asymptotic behavior of block Toeplitz matrices and determinants. II. *Adv. in Math.* **21** (1976) 1–29. MR0409512