CONVERGENCE OF THE SPECTRAL MEASURE
OF NON-NORMAL MATRICES

ALICE GUIONNET, PHILIP MATCHETT WOOD, AND OFER ZEITOUNI

(Communicated by Marius Junge)

Abstract. We discuss regularization by noise of the spectrum of large random non-normal matrices. Under suitable conditions, we show that the regularization of a sequence of matrices that converges in \( \ast \)-moments to a regular element \( a \) by the addition of a polynomially vanishing Gaussian Ginibre matrix forces the empirical measure of eigenvalues to converge to the Brown measure of \( a \).

1. Introduction

We discuss in this paper the changes in the spectrum of matrices when they are perturbed by noise. The behavior of the spectrum of matrices under (small) perturbation is not well understood when the matrices involved are non-normal; see [7] for background. A standard example of the issues involved (described e.g. in [6]) is the following. Consider the nilpotent matrix

\[
T_N = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & 0 & 1 \\
0 & \cdots & \cdots & \cdots & 0
\end{pmatrix}.
\]

The eigenvalues of \( T_N \) all vanish. However, adding to \( T_N \) the matrix \( P_N \) whose entries are all 0 except for the bottom left, which is taken as \( \epsilon_N \to 0 \), changes the eigenvalues drastically. As we will see below, as \( N \) increases, if \( \epsilon_N \) decays polynomially, the empirical measure of eigenvalues converges to the uniform measure on the unit circle in the complex plane. A natural question is, then, what happens when a small random matrix is added to \( T_N \)? One of the consequences of our results (see Corollaries 6 and 8) is that the spectrum of \( T_N + G_N \), where \( G_N \) is an additive Gaussian random perturbation that vanishes polynomially, behaves like the spectrum of \( T_N + P_N \), while this is false for matrices \( T_N \) that are themselves a low-rank perturbation of \( T_N \). The distinction involves the study of regular elements in \( W^\ast \) probability spaces.

Received by the editors October 11, 2011 and, in revised form, October 22, 2011; March 4, 2012; and March 15, 2012.

2010 Mathematics Subject Classification. Primary 60B20.

The work of the first author was partially supported by ANR project ANR-08-BLAN-0311-01.

The work of the third author was partially supported by NSF grant DMS-0804133 and by a grant from the Israel Science Foundation.

©2013 American Mathematical Society
To state our results we need to introduce some terminology. Consider a sequence \((A_N)_{N \geq 1}\), where \(A_N\) is an \(N \times N\) matrix of uniformly bounded operator norm, and assume that \(A_N\) converges in \(\ast\)-moments toward an element \(a\) in a \(W^*\) probability space \((\mathcal{A}, \| \cdot \|, \ast, \varphi)\); that is, for any non-commutative polynomial \(P\),

\[
\frac{1}{N} \text{tr} P(A_N, A_N^*) \xrightarrow{N \to \infty} \varphi(P(a, a^*)).
\]

We assume throughout that the tracial state \(\varphi\) is faithful; this does not represent a loss of generality. If \(A_N\) is a sequence of Hermitian matrices, the convergence above is enough in order to conclude that the empirical measure of eigenvalues of \(A_N\), that is, the measure

\[
L^A_N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(A_N)},
\]

where \(\lambda_i(A_N), i = 1, \ldots, N,\) are the eigenvalues of \(A_N\), converges weakly to a limiting measure \(\mu_a\), the spectral measure of \(a\), supported on a compact subset of \(\mathbb{R}\). (See [1, Corollary 5.2.16, Lemma 5.2.19] for this standard result and further background.) Significantly, in the Hermitian case, this convergence is stable under small bounded perturbations: with \(B_N = A_N + E_N\) and \(E_N\) Hermitian with \(\|E_N\| < \epsilon\), any subsequential limit of \(L^B_N\) will belong to \(B_L(\mu_a, \delta(\epsilon))\), with \(\delta(\epsilon) \xrightarrow{\epsilon \to 0} 0\) and \(B_L(\nu_a, r)\) the ball (in, say, the Lévy metric) centered at \(\nu_a\) and of radius \(r\).

Both these statements fail when \(A_N\) is not normal. Recall the matrix \(T_N\); see (1). Obviously, \(L^T_N = \delta_0\), while a simple computation reveals that \(T_N\) converges in \(\ast\)-moments to a unitary Haar element of \(\mathcal{A}\), that is,

\[
\frac{1}{N} \text{tr}(T_N^{\alpha_1}(T_N^*)^{\beta_1} \cdots T_N^{\alpha_k}(T_N^*)^{\beta_k}) \xrightarrow{N \to \infty} \begin{cases} 
1, & \text{if } \sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} \beta_i, \\
0, & \text{otherwise.}
\end{cases}
\]

Further, adding to \(T_N\) the matrix whose entries are all 0 except for the bottom left, which is taken as \(\epsilon\), changes the empirical measure of eigenvalues drastically: as already mentioned, when \(N\) increases, the empirical measure converges to the uniform measure on the unit circle in the complex plane.

We explore this phenomenon in the context of small random perturbations of matrices. We recall some notions. For \(a \in \mathcal{A}\), the Brown measure \(\nu_a\) on \(\mathbb{C}\) is the measure satisfying

\[
\log \det(z - a) = \int \log |z - z'| d\nu_a(z'), \quad z \in \mathbb{C},
\]

where \(\det\) is the Fuglede-Kadison determinant; we refer to [2][4] for definitions. We have in particular that

\[
\log \det(z - a) = \int \log x d\nu^z_a(x), \quad x \in \mathbb{C},
\]

where \(\nu^z_a\) denotes the spectral measure of the operator \(|z - a|\). In the sense of distributions, we have

\[
\nu_a = \frac{1}{2\pi} \Delta \log \det(z - a).
\]
That is, for a smooth compactly supported function $\psi$ on $\mathbb{C}$,
\[
\int \psi(z) d\nu_a(z) = \frac{1}{2\pi} \int \Delta \psi(z) \left( \int \log |z - z'| \, d\nu_a(z') \right) \, dz
= \frac{1}{2\pi} \int \Delta \psi(z) \left( \int \log x \, d\nu_a^z(x) \right) \, dz.
\]

A crucial assumption in our analysis is the following.

**Definition 1 (Regular elements).** An element $a \in A$ is regular if
\[
\lim_{\epsilon \to 0} \int_{\mathbb{C}} \Delta \psi(z) \left( \int_0^\epsilon \log x \, d\nu_a^z(x) \right) \, dz = 0,
\]
for all smooth functions $\psi$ on $\mathbb{C}$ with compact support.

Note that regularity is a property of $a$, not merely of its Brown measure $\nu_a$. It is easy to check that the unitary Haar element in $A$ is regular (see Section 4), while elements with dense, purely atomic spectrum are not.

We next introduce the class of Gaussian perturbations we consider.

**Definition 2 (Polynomially vanishing Gaussian matrices).** A sequence $(G_N)_{N \geq 1}$ with $G_N$ an $N$-by-$N$ random Gaussian matrix is called polynomially vanishing if the entries $(G_N(i,j))$ are independent centered complex Gaussian variables and there exist $\kappa > 0$, $\kappa' \geq 1 + \kappa$ so that
\[
N^{-\kappa'} \leq |E[G_{ij}]|^2 \leq N^{1-\kappa}.
\]

**Remark 3.** As will be clear below (see the beginning of the proof of Lemma 10), the Gaussian assumption only intervenes in obtaining a uniform lower bound on singular values of certain random matrices. As pointed out to us by R. Vershynin, this uniform estimate extends to other situations, most notably to the polynomial scaling of matrices whose entries are iid and possess a bounded density. We do not discuss such extensions here.

Our first result is a stability, with respect to polynomially vanishing Gaussian perturbations, of the convergence of spectral measures for non-normal matrices. Throughout, we denote by $\|M\|_{op}$ the operator norm of a matrix $M$.

**Theorem 4.** Assume that the uniformly bounded (in the operator norm) sequence of $N$-by-$N$ matrices $A_N$ converges in $\ast$-moments to a regular element $a$. Assume further that $L^A_N$ converges weakly to the Brown measure $\nu_a$. Let $G_N$ be a sequence of polynomially vanishing Gaussian matrices, and set $B_N = A_N + G_N$. Then, $L^B_N \to \nu_a$ weakly, in probability.

Theorem 4 puts rather stringent assumptions on the sequence $A_N$. In particular, its assumptions are not satisfied by the sequence of nilpotent matrices $T_N$ in [2]. Our second result corrects this deficiency, by showing that small Gaussian perturbations “regularize” matrices that are close to matrices satisfying the assumptions of Theorem 4.

**Theorem 5.** Let $A_N$, $E_N$ be a sequence of bounded (for the operator norm) $N$-by-$N$ matrices so that $A_N$ converges in $\ast$-moments to a regular element $a$. Assume that $\|E_N\|_{op}$ converges to zero polynomially fast in $N$ and that $L^A_N + E \to \nu_a$ weakly. Let $G_N$ be a sequence of polynomially vanishing Gaussian matrices, and set $B_N = A_N + G_N$. Then, $L^B_N \to \nu_a$ weakly, in probability.
Figure 1. The eigenvalues of $T_N + N^{-3-1/2} G_N$, where $G_N$ is iid complex Gaussian with mean 0 and variance 1 entries and $T_N$ as in (1).

Theorem 5 should be compared to earlier results of Sniady [6], who used stochastic calculus to show that a perturbation by an asymptotically vanishing Ginibre Gaussian matrix regularizes arbitrary matrices. Compared with his results, we allow for more general Gaussian perturbations (both structurally and in terms of the variance) and also show that the Gaussian regularization can decay as fast as wished in the polynomial scale. On the other hand, we do impose a regularity property on the limit $a$ as well as on the sequence of matrices for which we assume that adding a polynomially small matrix is enough to obtain convergence to the Brown measure.

A corollary of our general results is the following.

Corollary 6. Let $G_N$ be a sequence of polynomially vanishing Gaussian matrices and let $T_N$ be as in (1). Then $L_N^{T_N+G}$ converges weakly, in probability, toward the uniform measure on the unit circle in $\mathbb{C}$.

In Figure 1 we give a simulation of the setup in Corollary 6 for various $N$.

We will now define a class of matrices $T_{b,N}$ for which, if $b$ is chosen correctly, adding a polynomially vanishing Gaussian matrix $G_N$ is not sufficient to regularize $T_{b,N} + G_N$. Let $b$ be a positive integer, and define $T_{b,N}$ to be an $N$-by-$N$ block diagonal matrix in which each $(b+1)$-by-$(b+1)$ block on the diagonal equals $T_{b+1}$ (as defined in (1)). If $b+1$ does not divide $N$ evenly, a block of zeros is inserted at the bottom of the diagonal. Thus, every entry of $T_{b,N}$ is zero except for entries on the superdiagonal (the superdiagonal is the list of entries with coordinates $(i, i+1)$ for $1 \leq i \leq N - 1$), and the superdiagonal of $T_{b,N}$ is equal to $(1,1,\ldots,1,0,1,1,\ldots,1,0,\ldots,1,1,\ldots,1,0,0,\ldots,0)$.
Recall that the spectral radius of a matrix is the maximum absolute value of the eigenvalues. Also, we will use \( \|A\| = \text{tr}(A^*A)^{1/2} \) to denote the Hilbert-Schmidt norm.

**Proposition 7.** Let \( b = b(N) \) be a sequence of positive integers such that \( b(N) \gtrless \log N \) for all \( N \), and let \( T_{b,N} \) be as defined above. Let \( R_N \) be an \( N \)-by-\( N \) matrix satisfying \( \|R_N\| \leq g(N) \), where for all \( N \) we assume that \( g(N) < \frac{1}{3b\sqrt{N}} \). Then

\[
\rho(T_{b,N} + R_N) \leq (Ng(N))^{1/b} + o(1),
\]

where \( \rho(M) \) denotes the spectral radius of a matrix \( M \) and \( o(1) \) denotes a small quantity tending to zero as \( N \to \infty \).

Note that \( T_{b,N} \) converges in \( * \)-moments to a unitary Haar element of \( \mathcal{A} \) (by a computation similar to (2)) if \( b(N)/N \) goes to zero, which is a regular element. The Brown measure of the unitary Haar element is uniform measure on the unit circle; thus, in the case where \( (Ng(N))^{1/b} < 1 \), Proposition 7 shows that \( T_{b,N} + R_N \) does not converge to the Brown measure for \( T_{b,N} \).

**Corollary 8.** Let \( R_N \) be an iid Gaussian matrix where each entry has mean zero and variance one. Let \( b = b(N) \gtrless \log N \) be a sequence of integers, and let \( \gamma > 5/2 \) be a constant. Then, with probability tending to 1 as \( N \to \infty \), we have

\[
\rho(T_{b,N} + \exp(-\gamma b)R_N) \leq \exp\left(-\gamma + \frac{2\log N}{b}\right) + o(1),
\]

where \( \rho \) denotes the spectral radius and where \( o(1) \) denotes a small quantity tending to zero as \( N \to \infty \). Note in particular that the bound on the spectral radius is strictly less than \( \exp(-1/2) < 1 \) in the limit as \( N \to \infty \), due to the assumptions on \( \gamma \) and \( b \).

Corollary 8 follows from Proposition 7 by noting that, with probability tending to 1, all entries in \( R_N \) are at most \( C\log N \) in absolute value for some constant \( C \), and then checking that the hypotheses of Proposition 7 are satisfied for \( g(N) = \exp(-\gamma b)CN(\log N)^{1/4} \). There are two instances of Corollary 8 that are particularly interesting: when \( b = N - 1 \), we see that an exponentially decaying Gaussian perturbation does not regularize \( T_N = T_{N-1,N} \); and when \( b = \log(N) \), we see that polynomially decaying Gaussian perturbation does not regularize \( T_{\log N,N} \) (see Figure 2).

We will prove Proposition 7 in Section 5. The proof of our main results (Theorems 4 and 5) borrows from the methods of 8.

We finally introduce some notation that will be used throughout. For any \( N \)-by-\( N \) matrix \( C_N \), let

\[
\tilde{C}_N = \begin{pmatrix} 0 & C_N \\ C_N^* & 0 \end{pmatrix}.
\]

We denote by \( G_C \) the Cauchy-Stieltjes transform of the spectral measure of the matrix \( \tilde{C}_N \), that is,

\[
G_C(z) = \frac{1}{2N} \text{tr}(z - \tilde{C}_N)^{-1}, \quad z \in \mathbb{C}_+.
\]

The following useful estimate is immediate from the definition and the resolvent identity:

\[
|G_C(z) - G_D(z)| \leq \frac{\|C - D\|_{\text{op}}}{|3z|^2}.
\]
Figure 2. The eigenvalues of $T_{\log N, N} + N^{-3-1/2} G_N$, where $G_N$ is iid complex Gaussian with mean 0 and variance 1 entries. The spectral radius is roughly 0.07, and the bound from Corollary 8 is $\exp(-1) \approx 0.37$.

2. Proof of Theorem 4

We keep throughout this section the notation and assumptions of Theorem 4. The following is a crucial simple observation.

Proposition 9. For all complex numbers $\xi$, and all $z$ so that $\Im z \geq N^{-\delta}$ with $\delta < \kappa/4$,

$$E|\Im G_{B_N+\xi}(z)| \leq E|\Im G_{A_N+\xi}(z)| + 1.$$  

Proof. Noting that

$$E\|B_N - A_N\|_{op}^k = E\|G_N\|_{op}^k \leq C_k N^{-\kappa k/2},$$

the conclusion follows from (4) and H"older’s inequality. \hfill \Box

We continue with the proof of Theorem 4. Let $\nu_{A_N}$ denote the empirical measure of the eigenvalues of the matrix $A_N - z$. We have that, for smooth test functions $\psi$,

$$\int \Delta \psi(z) \left( \int \log |x| \, d\nu_{A_N}^z(x) \right) dz = \frac{1}{2\pi} \int \psi(z) \, dL_{A_N}^z(z).$$

In particular, the convergence of $L_{A_N}^z$ toward $\nu_a$ implies that

$$E \int \Delta \psi(z) \left( \int \log |x| \, d\nu_{A_N}^z(x) \right) dz$$

$$\to \int \psi(z) \, d\nu_a(z) = \int \Delta \psi(z) \left( \int \log x \, d\nu_a^z(x) \right) dz.$$  

On the other hand, since $x \to \log x$ is bounded continuous on compact subsets of $(0, \infty)$, it also holds that for any continuous bounded function $\zeta: \mathbb{R}_+ \to \mathbb{R}$
compactly supported in \((0, \infty)\),
\[
\mathbf{E} \int \Delta \psi(z) \left( \int \zeta(x) \log x dv_{A_N}^z(x) \right) \, dz + \int \Delta \psi(z) \left( \int \zeta(x) \log x dv_{E_N}^z(x) \right) \, dz.
\]
Together with the fact that \(a\) is regular and that \(A_N\) is uniformly bounded, one concludes therefore that
\[
\lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} \mathbf{E} \int_0^{e\varepsilon} \log |x| \, dv_{A_N}^z(x) \, dx = 0.
\]

Our next goal is to show that the same applies to \(B_N\). In the following, we let \(\nu_{B_N}^z\) denote the empirical measure of the eigenvalues of \(B_N - z\).

**Lemma 10.**
\[
\lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} \int \mathbf{E} \left[ \int_0^{e\varepsilon} \log |x|^{-1} \, dv_{B_N}^z(x) \right] \, dx = 0.
\]

Because \(\mathbf{E} \|B_N - A_N\|_{op}^k \to 0\) for any \(k > 0\), we have for any fixed smooth \(w\) compactly supported in \((0, \infty)\) that
\[
\mathbf{E} \left| \int \Delta \psi(z) \left( \int w(x) \log x dv_{A_N}^z(x) \right) \, dx - \int \Delta \psi(z) \left( \int w(x) \log x dv_{B_N}^z(x) \right) \, dx \right| \xrightarrow{N \to \infty} 0.
\]

Theorem 13 follows at once from Lemma 10.

**Proof of Lemma 10.** Note first that by [5, Theorem 3.3] (or its generalization in [3, Proposition 16] to the complex case) there exists a constant \(C\) so that for any \(z\), the smallest singular value \(\sigma_N^z\) of \(B_N + zI\) satisfies
\[
P(\sigma_N^z \leq x) \leq C \left( \frac{N}{x} \right)^{\frac{1}{2} + \kappa'},
\]
with \(\beta = 1\) or 2 according to whether we are in the real or the complex case. Therefore, for any \(\zeta > 0\), uniformly in \(z\),
\[
\mathbf{E} \left[ \int_0^{N^{-\zeta}} \log |x|^{-1} \, dv_{B_N}^z(x) \right] \leq \mathbf{E}[\log(\sigma_N^z)^{-1}1_{\sigma_N^z \leq N^{-\zeta}}] = C \left( \frac{N^{\frac{1}{2} + \kappa' - \zeta}}{x} \right)^{\beta} \log(N) + \int_0^{N^{-\zeta}} \frac{1}{x} C \left( \frac{N^{\frac{1}{2} + \kappa'}}{x} \right)^{\beta} \, dx
\]
goes to zero as \(N\) goes to infinity as soon as \(\zeta > \frac{1}{2} + \kappa'\). We fix hereafter such a \(\zeta\) and we may and shall restrict the integration from \(N^{-\zeta}\) to \(\varepsilon\). To compare the integral for the spectral measure of \(A_N\) and \(B_N\), observe that for any probability measure \(P\), with \(P_{\gamma}\) the Cauchy law with parameter \(\gamma\),
\[
P([a, b]) \leq (P * P_{\gamma}) ([a - \eta, b + \eta]) + P_{\gamma}([-\eta, \eta] \cup [b + \eta, \infty)) \leq (P * P_{\gamma}) ([a - \eta, b + \eta]) + \frac{\gamma}{\eta},
\]
whereas for \(b - a > \eta\),
\[
P([a, b]) \geq (P * P_{\gamma}) ([a + \eta, b - \eta]) - \frac{\gamma}{\eta}.
\]
Recall that
\[
(P * P_{\gamma}) ([a, b]) = \int_a^b |3G(x + i\gamma)| \, dx.
\]
Set $\gamma = N^{-\kappa/5}$, $\kappa'' = \kappa/2$ and $\eta = N^{-\kappa''/5}$. We have, whenever $b - a \geq 4\eta$, 
\[
\mathbb{E}\nu_{B_N}^z([a, b]) \leq \int_{a - \eta}^{b + \eta} \mathbb{E}[3G_{B_n+z}(x + i\gamma)]dx + N^{-(\kappa - \kappa'')/5}
\]
\[
\leq (b - a + 2N^{-\kappa''/5}) + \nu_{A_N}^z \ast P_{N^{-\kappa''/5}}([a - N^{-\kappa/10}, b + N^{-\kappa/10}]) + N^{-\kappa/10}
\]
\[
\leq (b - a + 2N^{-\kappa''/5}) + \nu_{A_N}^z ([(a - 2N^{-\kappa/10})_+, (b + 2N^{-\kappa/10})]) + 2N^{-\kappa/10},
\]
where the first inequality is due to (6) and (8), the second is due to Proposition 9 and the last uses (7) and (8). Therefore, if $b - a = CN^{-\kappa/10}$ for some fixed $C$ larger than 4, we deduce that there exists a finite constant $C'$ which depends only on $C$ so that 
\[
\mathbb{E}\nu_{B_N}^z([a, b]) \leq C'(b - a) + \nu_{A_N}^z([(a - 2N^{-\kappa/10})_+, (b + 2N^{-\kappa/10})])
\]
As a consequence, as we may assume without loss of generality that $\kappa' > \kappa/10$, 
\[
\mathbb{E}[\int_{N^{-\zeta}}^{\varepsilon} \log |x|^{-1}d\nu_{B_N}^z(x)]
\]
\[
\leq \sum_{k=0} [N^{{\kappa''/10}+\varepsilon}] \log(N^{-\zeta} + 2CkN^{-\kappa/10})^{-1} \cdot \mathbb{E}[\nu_{B_N}^z(\left[N^{-\zeta} + 2CkN^{-\kappa/10}, N^{-\zeta} + 2C(k + 1)N^{-\kappa/10}\right])].
\]
We need to pay special attention to the first term that we bound by noticing that 
\[
\log(N^{-\zeta})^{-1}\mathbb{E}[\nu_{B_N}^z(\left[N^{-\zeta}, N^{-\zeta} + 2CN^{-\kappa/10}\right])]
\]
\[
\leq \frac{10C}{\kappa} \log(N^{-\kappa/10})^{-1} \mathbb{E}[\nu_{B_N}^z(\left[0, 2(C + 1)N^{-\kappa/10}\right])]
\]
\[
\leq \frac{10C}{\kappa} \log(N^{-\kappa/10})^{-1}(2C'N^{-\kappa/10} + \nu_{A_N}^z(\left[0, (C + 2)N^{-\kappa/10}\right]))
\]
\[
\leq \frac{20C'\zeta}{\kappa} \log(N^{-\kappa/10})^{-1}N^{-\kappa/10} + C'' \int_0^{2(C + 2)N^{-\kappa/10}} \log |x|^{-1}d\nu_{A_N}^z(x).
\]
For the other terms, we have 
\[
\sum_{k=1} [N^{{\kappa''/10}+\varepsilon}] \log(N^{-\zeta} + 2Ck\kappa^{-\kappa/10})^{-1} \cdot \mathbb{E}[\nu_{B_N}^z(\left[N^{-\zeta} + 2Ck\kappa^{-\kappa/10}, N^{-\zeta} + 2C(k + 1)\kappa^{-\kappa/10}\right])]
\]
\[
\leq 2C' \sum_{k=1} [N^{{\kappa''/10}+\varepsilon}] \log(Ck\kappa^{-\kappa/10})^{-1}C\kappa^{-\kappa/10}
\]
\[
+ \sum_{k=1} [N^{{\kappa''/10}+\varepsilon}] \log(Ck\kappa^{-\kappa/10})^{-1}\nu_{A_N}^z(\left[2C(k - 1)\kappa^{-\kappa/10}, 2C(k + 2)\kappa^{-\kappa/10}\right]).
\]
Finally, we can sum up all these inequalities to find that there exists a finite constant $C'''$ so that 
\[
\mathbb{E}\left[\int_{N^{-\zeta}}^{\varepsilon} \log |x|^{-1}d\nu_{B_N}^z(x)\right] \leq C''' \int_0^{\varepsilon} \log |x|^{-1}d\nu_{A_N}^z(x) + C''' \int_0^{\varepsilon} \log |x|^{-1}dx
\]
and therefore goes to zero when $N \to \infty$ and then $\varepsilon$ goes to zero. This proves the claim. \qed
3. Proof of Theorem 5

From the assumptions, it is clear that \((A_N + E_N)\) converges in \(\ast\)-moments to the regular element \(a\). By Theorem 4, it follows that \(L^{A+E+G}_N\) converges (weakly, in probability) towards \(\nu_a\). We can now remove \(E_N\). Indeed, by (4) and (5), we have for any \(\chi < \kappa'/2\) and all \(\xi \in \mathbb{C}\),

\[
|G_{A+G+\xi}^N(z) - G_{A+G+E+\xi}^N(z)| \leq \frac{N^{-\chi}}{|z|^2},
\]

and therefore for \(\exists \ z \geq N^{-\chi/2}\),

\[
|\exists G_{A+G+\xi}^N(z)| \leq |\exists G_{A+G+E+\xi}^N(z)| + 1.
\]

Again by [5, Theorem 3.3] (or its generalization in [3, Proposition 16]) to the complex case, for any \(z\), the smallest singular value \(\sigma^x_N\) of \(A_N + G_N + z\) satisfies

\[
P(\sigma^x_N \leq x) \leq C \left( N^{\frac{3}{2} + \kappa'} x \right)^{\beta}
\]

with \(\beta = 1\) or 2 according to whether we are in the real or the complex case. We can now rerun the proof of Theorem 4, replacing \(A_N\) by \(A'_N = A_N + E_N + G_N\) and \(B_N\) by \(A'_N - E_N\). \(\square\)

4. Proof of Corollary 6

We apply Theorem 5 with \(A_N = T_N\), \(E_N\) the \(N\)-by-\(N\) matrix with

\[
E_N(i, j) = \begin{cases} 
\delta_N = N^{-(1/2 + \kappa')}, & i = 1, j = N, \\
0, & \text{otherwise},
\end{cases}
\]

where \(\kappa' > \kappa\). We check the assumptions of Theorem 5. We take \(a\) to be a unitary Haar element in \(A\), and recall that its Brown measure \(\nu_a\) is the uniform measure on \(\{z \in \mathbb{C} : |z| = 1\}\). We now check that \(a\) is regular. Indeed, \(\int x^k d\nu_a^z(x) = 0\) if \(k\) is odd by symmetry, while for \(k\) even,

\[
\int x^k d\nu_a^z(x) = \frac{\varphi(([z - a](z - a)^*)^{k/2})}{\sum_{j=1}^{k/2} (|z|^2 + 1)^{k/2 - j}} = \sum_{j=1}^{k/2} (|z|^2 + 1)^{k/2 - j} \left( \begin{array}{c} k \\ j \end{array} \right) \left( \begin{array}{c} 2j \\ j \end{array} \right),
\]

and one therefore verifies that for \(k\) even,

\[
\int x^k d\nu_a^z(x) = \frac{1}{2\pi} \int (|z|^2 + 1 + 2|z| \cos \theta)^{k/2} d\theta.
\]

It follows that

\[
\int_0^\epsilon \log x d\nu_a^z(x) = \frac{1}{4\pi} \int_0^{2\pi} \log(|z|^2 + 1 + 2|z| \cos \theta) 1_{\{|z|^2 + 1 + 2|z| \cos \theta < \epsilon\}} d\theta \xrightarrow{\epsilon \to 0} 0,
\]

proving the required regularity.

Further, we claim that \(L^{A+E}_N\) converges to \(\nu_a\). Indeed the eigenvalues \(\lambda\) of \(A_N + E_N\) are such that there exists a non-vanishing vector \(u\) so that

\[
u_N \delta_N = \lambda u_1, u_{i-1} = \lambda u_i,
\]

that is

\[
\lambda^N = \delta_N.
\]
In particular, all the $N$-roots of $\delta_N$ are (distinct) eigenvalues; that is, the eigenvalues $\lambda_j^N$ of $A_N$ are

$$\lambda_j^N = |\delta_N|^{1/N}e^{2i\pi j/N}, \quad 1 \leq j \leq N.$$ 

Therefore, for any bounded continuous $g$ function on $\mathbb{C}$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} g(\lambda_j^N) = \frac{1}{2\pi} \int g(\theta) d\theta,$$

as claimed. \hfill \Box

5. Proof of Proposition 7

In this section we will prove the following proposition:

**Proposition 11.** Let $b = b(N)$ be a sequence of positive integers, and let $T_{b,N}$ be as in Proposition 7. Let $R_N$ be an $N$-by-$N$ matrix satisfying $\|R_N\| \leq g(N)$, where for all $N$ we assume that $g(N) < \frac{1}{3b\sqrt{N}}$. Then

$$\rho(T_{b,N} + R_N) \leq \left( O\left( \sqrt{Nb} \left( 2N^{1/4} g^{1/2} \right)^b \right) \right)^{1/(b+1)} + \left( b^2 Ng \right)^{1/(b+1)}.$$

Proposition 7 follows from Proposition 11 by adding the assumption that $b(N) \geq \log(N)$ and then simplifying the upper bound on the spectral radius.

**Proof of Proposition 11.** To bound the spectral radius, we will use the fact that $\rho(T_{b,N} + R_N) \leq \| (T_{b,N} + R_N)^k \|^{1/k}$ for all integers $k \geq 1$, where $\| \cdot \|$ denotes the Hilbert-Schmidt norm. Our general plan will be to bound $\| (T_{b,N} + R_N)^k \|$ and then take a $k$-th root of the bound. We will take $k = b+1$, which allows us to take advantage of the fact that $T_{b,N}$ is a $(b+1)$-step nilpotent. In particular, we make use of the fact that for any positive integer $a$,

$$(9) \quad \|T_{b,N}^a\| = \begin{cases} (b - a + 1)^{1/2} \left[ \frac{N}{b+1} \right]^{1/2} & \text{if } 1 \leq a \leq b \\ 0 & \text{if } b + 1 \leq a. \end{cases}$$

We may write

$$\| (T_{b,N} + R_N)^{b+1} \| \leq \sum_{\lambda \in \{0,1\}^{b+1}} \left\| \prod_{i=1}^{b+1} T_{b,N}^{\lambda_i} R_N^{1-\lambda_i} \right\|$$

$$= \sum_{\ell=0}^{b+1} \sum_{\lambda \in \{0,1\}^{b+1}} \left( \sum_{\lambda \text{ has } \ell \text{ ones}} \left\| \prod_{i=1}^{b+1} T_{b,N}^{\lambda_i} R_N^{1-\lambda_i} \right\| \right).$$

When $\ell$ is large, we will make use of the following lemma.

**Lemma 12.** If $\lambda \in \{0,1\}^k$ has $\ell$ ones and $\ell \geq (k+1)/2$, then

$$\left\| \prod_{i=1}^{k} T_{b,N}^{\lambda_i} R_N^{1-\lambda_i} \right\| \leq \left\| T_{b,N}^{\frac{k-\ell+1}{\ell+1}} \right\|^{k-\ell} \| R_N \|^{k-\ell}.$$
We will prove Lemma 12 in Section 5.1. Using Lemma 12 with \( k = b + 1 \) along with the fact that \( \|AB\| \leq \|A\| \|B\| \), we have

\[
\left\| (T_{b,N} + R_N)^{b+1} \right\| \leq \sum_{\ell=0}^{\left\lfloor \frac{b+2}{2} \right\rfloor} \left( \frac{b+1}{\ell} \right) \|T_{b,N}\| \|R_n\| \ell^{b-\ell+1}
\]

\[
+ \sum_{\ell=\left\lfloor \frac{b+2}{2} \right\rfloor}^{b+1} \left( \frac{b+1}{\ell} \right) \left\| T_{b,N} \left[ \frac{\ell+1}{\ell+2} \right] \right\| \ell^{b-\ell+2} \|R_N\| \ell^{b-\ell+1}
\]

(10)

where the second inequality comes from the assumption \( \|R_N\| \leq g = g(N) \).

We will bound (10) and (11) separately. To bound (10) note that

\[
\sum_{\ell=0}^{\left\lfloor \frac{b+2}{2} \right\rfloor} \left( \frac{b+1}{\ell} \right) \|T_{b,N}\| \ell^{b-\ell+1} \leq \sum_{\ell=0}^{\left\lfloor \frac{b+2}{2} \right\rfloor} \left( \frac{b+1}{\ell} \right) \left( \frac{N}{b+1} \right)^{\ell/2} \left( \frac{2^2}{b+1} \right) N^{(b+2)/4} g^{b/2}
\]

\[
= O \left( \sqrt{N} b (2N^{1/4} g^{1/2})^b \right).
\]

(12)

Next, we turn to bounding (11). We will use the following lemma to show that the largest term in the sum (11) comes from the \( \ell = b \) term. Note that when \( \ell = b + 1 \), the summand in (11) is equal to zero by (9).

**Lemma 13.** If \( \left\| T_{b,N} \left[ \frac{\ell+1}{\ell+2} \right] \right\| > 0 \) and \( \ell \leq b - 1 \) and

\[
g \leq \frac{2}{e^{3/2} N^{1/2} b},
\]

then

\[
\left( \frac{b+1}{\ell} \right) \left\| T_{b,N} \left[ \frac{\ell+1}{\ell+2} \right] \right\| \ell^{b-\ell+2} g^{b-\ell+1} \leq \left( \frac{b+1}{\ell+1} \right) \left\| T_{b,N} \left[ \frac{\ell+1}{\ell+2} \right] \right\| \ell^{b-\ell+1} g^{b-\ell}.
\]

We will prove Lemma 13 in Section 5.1.

Using Lemma 13 we have

\[
\sum_{\ell=\left\lfloor \frac{b+2}{2} \right\rfloor}^{b+1} \left( \frac{b+1}{\ell} \right) \left\| T_{b,N} \left[ \frac{\ell+1}{\ell+2} \right] \right\| \ell^{b-\ell+2} g^{b-\ell+1} \leq \frac{b}{2} (b+1) \left\| T_{b,N} \left[ \frac{\ell+1}{\ell+2} \right] \right\| g^{b-\ell+1}
\]

\[
\leq \frac{b}{2} (b+1) (b - \lfloor b/2 \rfloor + 1) \frac{N}{b+1} g
\]

\[
\leq b^2 N g.
\]

(13)

Combining (12) and (13) with (10) and (11), we may use the fact that

\[
(x + y)^{1/(b+1)} \leq x^{1/(b+1)} + y^{1/(b+1)}
\]
for positive $x, y$ to complete the proof of Proposition 11. It remains to prove Lemma 12 and Lemma 13, which we do in Section 5.1 below.

5.1. Proofs of Lemma 12 and Lemma 13

Proof of Lemma 12. Using (9), it is easy to show that
\begin{equation}
\|T_{a, b, N}^c\| \leq \|T_{a - 1, b, N}^c\| \|T_{a, b, N}^c\| \quad \text{for integers } 3 \leq c + 2 \leq a \leq b.
\end{equation}

It is also clear from (9) that
\begin{equation}
\|T_{a, b, N}^c\| \leq \|T_{a - 1, b, N}^c\| \quad \text{for all positive integers } a.
\end{equation}

Let $\lambda \in \{0, 1\}^k$ have $\ell$ ones. Then, using the assumption that $\ell \geq k - \ell + 1$, we may write
\begin{equation}
\prod_{i=1}^k T_{a_i, b, N}^{\lambda_i} R_{1 - \lambda_i} = T_{a_1, b, N}^{b_1} T_{a_2, b, N}^{b_2} \cdots T_{a_k, b, N}^{a_k - \ell} R_{b_k - a_k - \ell}^{-1},
\end{equation}

where $a_i \geq 1$ for all $i$ and $b_i \geq 0$ for all $i$. Thus
\begin{equation}
\left\| \prod_{i=1}^k T_{a_i, b, N}^{\lambda_i} R_{1 - \lambda_i} \right\| \leq \|R_N\|^{k - \ell} \prod_{i=1}^{k - \ell + 1} \left\| T_{a_i, b, N}^{a_i - \lambda_i} \right\|.
\end{equation}

Applying (14) repeatedly, we may assume that two of the $a_i$ differ by more than 1, all without changing the fact that $\sum_{i=1}^{k - \ell + 1} a_i = \ell$. Thus, some of the $a_i$ are equal to $\left\lfloor \frac{\ell}{k - \ell + 1} \right\rfloor$ and some are equal to $\left\lceil \frac{\ell}{k - \ell + 1} \right\rceil$. Finally, applying (15), we have that
\begin{equation}
\prod_{i=1}^{k - \ell + 1} \left\| T_{a_i, b, N}^{a_i} \right\| \leq \left\| T_{a_i, b, N}^{\frac{\ell}{k - \ell + 1}} \right\|^{k - \ell + 1}.
\end{equation}

Proof of Lemma 13. Using (9) and rearranging, it is sufficient to show that
\begin{equation}
\frac{\ell + 1}{b - \ell + 1} \left( b - \left\lfloor \frac{\ell}{b - \ell + 2} \right\rfloor + 1 \right)^{1/2} \left( \frac{N}{b + 1} \right)^{1/2} g \leq \left( b - \left\lfloor \frac{\ell + 1}{b - \ell + 1} \right\rfloor + 1 \right) \left( b - \left\lfloor \frac{\ell + 2}{b - \ell + 2} \right\rfloor + 1 \right)^{1/2}.
\end{equation}

Using a variety of manipulations, it is possible to show that
\begin{equation}
\left( b - \left\lfloor \frac{\ell + 1}{b - \ell + 1} \right\rfloor + 1 \right)^{\frac{b - \ell + 1}{2}} \geq \exp \left( -\frac{(b - \ell + 2)(b - \ell + 1)}{(b + 2)(b - \ell + 2)} \right) \geq \exp(-3/2).
\end{equation}

Thus, it is sufficient to have
\begin{equation}
\frac{b}{2} N^{1/2} g \leq \exp(-3/2),
\end{equation}

which is true by assumption.
REFERENCES


UMPA, CNRS UMR 5669, ENS Lyon, 46 allée d’Italie, 69007 Lyon, France
E-mail address: aguionne@ens-lyon.fr
Current address: Department of Mathematics, Room 2-272, Massachusetts Institute of Technology, Cambridge, Massachusetts 02134

DEPARTMENT OF MATHEMATICS, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, Wisconsin 53706-1388
E-mail address: pmwood@math.wisc.edu

SCHOOL OF MATHEMATICS, University of Minnesota, 206 Church Street SE, Minneapolis, Minnesota 55455—and—FACULTY OF MATHEMATICS, Weizmann Institute, POB 26, Rehovot 76100, Israel
Current address: Faculty of Mathematics, Weizmann Institute, POB 26, Rehovot 76100, Israel
E-mail address: zeitouni@math.umn.edu