



HAL
open science

Circular law for non-central random matrices

Djalil Chafai

► **To cite this version:**

Djalil Chafai. Circular law for non-central random matrices. *Journal of Theoretical Probability*, 2010, 23 (4), pp.945-950. 10.1007/s10959-010-0285-8. hal-00169167v3

HAL Id: hal-00169167

<https://hal.science/hal-00169167v3>

Submitted on 5 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Circular law for non-central random matrices

Djalil CHAFAÏ

Preprint, June 2008. Revised March 2010.

Abstract

Let $(X_{jk})_{j,k \geq 1}$ be an infinite array of i.i.d. complex random variables, with mean 0 and variance 1. Let $\lambda_{n,1}, \dots, \lambda_{n,n}$ be the eigenvalues of $(\frac{1}{\sqrt{n}}X_{jk})_{1 \leq j,k \leq n}$. The strong circular law theorem states that with probability one, the empirical spectral distribution $\frac{1}{n}(\delta_{\lambda_{n,1}} + \dots + \delta_{\lambda_{n,n}})$ converges weakly as $n \rightarrow \infty$ to the uniform law over the unit disc $\{z \in \mathbb{C}; |z| \leq 1\}$. In this short note, we provide an elementary argument that allows to add a deterministic matrix M to $(X_{jk})_{1 \leq j,k \leq n}$ provided that $\text{Tr}(MM^*) = O(n^2)$ and $\text{rank}(M) = O(n^\alpha)$ with $\alpha < 1$. Conveniently, the argument is similar to the one used for the non-central version of Wigner's and Marchenko-Pastur theorems.

AMS 2010 Mathematical Subject Classification: 15B52.

Keywords: Random matrices; Circular law.

1 Introduction

For any square $n \times n$ matrix \mathbf{A} with complex entries, let the complex eigenvalues $\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A})$ of \mathbf{A} be labeled so that $|\lambda_1(\mathbf{A})| \geq \dots \geq |\lambda_n(\mathbf{A})|$. The *empirical spectral distribution* of \mathbf{A} is the discrete probability measure $\mu_{\mathbf{A}} := \frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k(\mathbf{A})}$. We denote by $s_1(\mathbf{A}) \geq \dots \geq s_n(\mathbf{A})$ the *singular values* of \mathbf{A} , i.e. the eigenvalues of the positive semi-definite Hermitian matrix $\sqrt{\mathbf{A}\mathbf{A}^*}$ where \mathbf{A}^* is the conjugate-transpose of \mathbf{A} . The *operator norm* is $s_1(\mathbf{A}) = \max_{\|x\|_2=1} \|\mathbf{A}x\|_2$ and the square *Hilbert-Schmidt* norm is $\|\mathbf{A}\|^2 := s_1(\mathbf{A})^2 + \dots + s_n(\mathbf{A})^2 = \text{Tr}(\mathbf{A}\mathbf{A}^*) = \sum_{j,k=1}^n |\mathbf{A}_{j,k}|^2$. Weyl's inequality $|\lambda_1(\mathbf{A})|^2 + \dots + |\lambda_n(\mathbf{A})|^2 \leq s_1(\mathbf{A})^2 + \dots + s_n(\mathbf{A})^2$ ensures that the second moment of $\mu_{\mathbf{A}}$ is always bounded above by $\frac{1}{n}\|\mathbf{A}\|^2$. The following result was recently obtained by Tao and Vu [19, Corollary 1.15].

Theorem 1.1 (Circular law for central random matrices). *Let $(X_{jk})_{j,k \geq 1}$ be i.i.d. complex random variables. Let $(M_{jk})_{j,k \geq 1}$ be deterministic complex numbers. For every integer $n \geq 1$, set $\mathbf{X}_n = (X_{jk})_{1 \leq j,k \leq n}$ and $\mathbf{M}_n = (M_{jk})_{1 \leq j,k \leq n}$. If*

- $\mathbb{E}[|X_{1,1}|^2] = 1$ and $\mathbb{E}[X_{1,1}] = 0$
- $\|\mathbf{M}_n\|^2 = O(n^2)$ and $\text{rank}(\mathbf{M}_n) = O(n^\alpha)$ for some $\alpha < 1$

then with probability one, $\mu_{\frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n)}$ tends weakly as $n \rightarrow \infty$ to the uniform distribution on the unit disc $\{z \in \mathbb{C}; |z| \leq 1\}$ (known as the circular law).

The aim of this note is to provide an alternative and elementary argument which reduces theorem 1.1 to the central case where $\mathbf{M}_n \equiv 0$ for every n . Conveniently, the approach is close in spirit to the one used by Bai [3] for the derivation of Wigner's and Marchenko-Pastur theorems for non-central random matrices.

This note was motivated by the study of random Markov matrices, including the Dirichlet Markov Ensemble [8, 7], for which a circular law theorem is conjectured. The initial version of this note was written before the apparition of [19], and provided for the first time a non-central version of the circular law theorem. The initial version was based on potential theoretic tools. For convenience, the present version makes use instead of the replacement principle borrowed from [19].

Theorem 1.1 belongs to a sequence of works by many authors, including Mehta [13], Girko [10], Silverstein [12], Bai [2], Edelman [9] Śniady [17], Bai and Silverstein [4], Pan and Zhou [14], Götze and Tikhomirov [11], and Tao and Vu [18].

Remark 1.2 (Constant case). *Consider the case where the entries of \mathbf{M}_n are all equal to 1 in theorem 1.1. We have then $\text{rank}(\mathbf{M}_n) = 1$ and $s_1(\mathbf{M}_n) = n$. Suppose additionally that $X_{1,1}$ has finite fourth moment. Then, by Bai and Yin theorem [5], with probability one, $\lim_{n \rightarrow \infty} s_1(\frac{1}{\sqrt{n}}\mathbf{X}_n) = 2$, and thus $\frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n)$ is a random bounded perturbation of the rank one symmetric matrix $\frac{1}{\sqrt{n}}\mathbf{M}_n$ which has spectrum*

$$\lambda_n(\mathbf{M}_n) = \dots = \lambda_2(\mathbf{M}_n) = 0 \quad \text{and} \quad \lambda_1(\mathbf{M}_n) = \sqrt{n}.$$

From this observation, Silverstein [16] has shown, via perturbation techniques such as Bauer-Fike and Gerschgorin theorems, that with probability one,

$$\left| \lambda_2 \left(\frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n) \right) \right| \leq 2 + o(1) \quad \text{and} \quad \left| \lambda_1 \left(\frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n) \right) - \sqrt{n} \right| \leq 2 + o(1).$$

See also the work of Andrew [1]. Also, with probability one, as $n \rightarrow \infty$, the spectral radius $|\lambda_1(\frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n))|$ blows up while $\mu_{\frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n)}$ remains weakly localized.

2 Reduction to the central case

In order to show that theorem 1.1 reduces to the central case where $\mathbf{M}_n \equiv 0$ for every $n \geq 1$, it suffices to check the assumptions of the replacement principle of theorem 3.1 with $\mathbf{A}_n := \frac{1}{\sqrt{n}}\mathbf{X}_n$ and $\mathbf{B}_n := \frac{1}{\sqrt{n}}(\mathbf{X}_n + \mathbf{M}_n)$. By the strong law of large numbers and the assumption on $\|\mathbf{M}_n\|$, with probability one,

$$\frac{1}{n}\|\mathbf{A}_n\|^2 + \frac{1}{n}\|\mathbf{B}_n\|^2 = O(1).$$

Next, by theorem (3.4) and the first Borel-Cantelli lemma, for all $z \in \mathbb{C}$, with probability one, the random matrices $\mathbf{A}_n - z\mathbf{I}_n$ and $\mathbf{B}_n - z\mathbf{I}_n$ are invertible for large enough n . Let us define, for large enough n , the quantity

$$\Delta_{n,z} := \frac{1}{n} \log |\det(\mathbf{A}_n - z\mathbf{I}_n)| - \frac{1}{n} \log |\det(\mathbf{B}_n - z\mathbf{I}_n)|.$$

If we set $\mu_{n,z} := \mu_{\sqrt{(\mathbf{A}_n - z\mathbf{I}_n)(\mathbf{A}_n - z\mathbf{I}_n)^*}}$ and $\nu_{n,z} := \mu_{\sqrt{(\mathbf{B}_n - z\mathbf{I}_n)(\mathbf{B}_n - z\mathbf{I}_n)^*}}$ then

$$\Delta_{n,z} = \int_0^\infty \log(t) d(\mu_{n,z} - \nu_{n,z})(t).$$

By the strong law of large numbers and the assumption on $\|\mathbf{M}_n\|$, for all $z \in \mathbb{C}$, with probability one, there exists $a > 0$ such that

$$\max(s_1(\mathbf{A}_n - z\mathbf{I}_n), s_1(\mathbf{B}_n - z\mathbf{I}_n)) \leq n^a$$

for large enough n . On the other hand, by theorem (3.4) and the first Borel-Cantelli lemma, for all $z \in \mathbb{C}$, with probability one, there exists $b > 0$ such that

$$\min(s_n(\mathbf{A}_n - z\mathbf{I}_n), s_n(\mathbf{B}_n - z\mathbf{I}_n)) \geq n^{-b}$$

for large enough n . Therefore, with $\alpha_n := n^{-b}$ and $\beta_n := n^a$, and large enough n ,

$$\Delta_{n,z} = \int_{\alpha_n}^{\beta_n} \log(t) d(\mu_{n,z} - \nu_{n,z})(t).$$

Let $F_{n,z}$ and $G_{n,z}$ be the cumulative distribution functions of the real probability measures $\mu_{n,z}$ and $\nu_{n,z}$. By lemma 3.3 and the assumption on $\text{rank}(\mathbf{M}_n)$, for almost all $z \in \mathbb{C}$, with probability one, there exists $\varepsilon > 0$ such that

$$\|F_{n,z} - G_{n,z}\|_\infty = O(n^{-\varepsilon}).$$

Therefore, by lemma 3.2, we obtain, for almost all $z \in \mathbb{C}$, with probability one,

$$|\Delta_{n,z}| \leq (\log(\beta_n) - \log(\alpha_n)) \|F_{n,z} - G_{n,z}\|_\infty = o(1).$$

3 Tools

This section gathers some tools used in our proof of theorem 1.1. By Green's theorem, for any complex polynomial P and smooth compactly supported $f : \mathbb{C} \rightarrow \mathbb{R}$,

$$\int_{\mathbb{C}} f d\mu = \frac{1}{2\pi} \int_{\mathbb{C}} \Delta f \log |P| dx dy$$

where $\mu := \delta_{\lambda_1} + \dots + \delta_{\lambda_n}$ is the counting measure of the roots $\lambda_1, \dots, \lambda_n$ of P in \mathbb{C} . Used for characteristic polynomials of random matrices, this identity provides, via dominated convergence arguments, the following theorem, see [19, Theorem 2.1].

Theorem 3.1 (Replacement principle). *Let $(\mathbf{A}_n)_{n \geq 1}$ and $(\mathbf{B}_n)_{n \geq 1}$ be two sequences of complex random matrices where $\mathbf{A}_n, \mathbf{B}_n$ are $n \times n$, without any assumptions. If*

- with probability one $\frac{1}{n} \|\mathbf{A}_n\|^2 + \frac{1}{n} \|\mathbf{B}_n\|^2 = O(1)$
- for almost all $z \in \mathbb{C}$, with probability one, the random matrices $\mathbf{A}_n - z\mathbf{I}_n$ and $\mathbf{B}_n - z\mathbf{I}_n$ are invertible for large enough n

- for almost all $z \in \mathbb{C}$, with probability one,

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \log |\det(\mathbf{A}_n - z\mathbf{I}_n)| - \frac{1}{n} \log |\det(\mathbf{B}_n - z\mathbf{I}_n)| \right) = 0$$

then with probability one, $\mu_{\mathbf{A}_n} - \mu_{\mathbf{B}_n}$ tends weakly to zero as $n \rightarrow \infty$.

The following lemma is a special case of the integration by parts formula for the Lebesgue-Stieltjes integral (with atoms). We give a short proof for convenience.

Lemma 3.2 (Integration by parts). *If $a_1, \dots, a_n, b_1, \dots, b_n \in [\alpha, \beta] \subset \mathbb{R}$, and F_μ and F_ν are the cumulative distribution functions of $\mu = \frac{1}{n}(\delta_{a_1} + \dots + \delta_{a_n})$ and $\nu = \frac{1}{n}(\delta_{b_1} + \dots + \delta_{b_n})$ respectively, then for any smooth $f : [\alpha, \beta] \rightarrow \mathbb{R}$,*

$$\int_\alpha^\beta f(x) d\mu(x) - \int_\alpha^\beta f(x) d\nu(x) = \int_\alpha^\beta f'(x)(F_\mu(x) - F_\nu(x)) dx.$$

In particular, when f is non decreasing,

$$\left| \int_\alpha^\beta f(x) d\mu(x) - \int_\alpha^\beta f(x) d\nu(x) \right| \leq (f(\beta) - f(\alpha)) \|F_\mu - F_\nu\|_\infty.$$

Proof. One can assume by continuity that $a_1, \dots, a_n, b_1, \dots, b_n$ are all different. We reorder $a_1, \dots, a_n, b_1, \dots, b_n$ into $c_1 \leq \dots \leq c_{2n}$. For every $1 \leq k \leq 2n$, set $\varepsilon_k = +1$ if $c_k \in \{a_1, \dots, a_n\}$ and $\varepsilon_k = -1$ if $c_k \in \{b_1, \dots, b_n\}$. We have

$$\int_\alpha^\beta f(x) d\mu(x) - \int_\alpha^\beta f(x) d\nu(x) = \frac{1}{n} \sum_{k=1}^n (f(a_i) - f(b_i)) = \frac{1}{n} \sum_{k=1}^{2n} \varepsilon_k f(c_k).$$

By an Abel transform, we get by denoting $S_k = \varepsilon_1 + \dots + \varepsilon_k$,

$$\sum_{k=1}^{2n} \varepsilon_k f(c_k) = - \sum_{k=1}^{2n-1} S_k (f(c_{k+1}) - f(c_k)) + S_{2n} f(c_{2n}).$$

Since $F_\mu - F_\nu$ is constant and equal to S_k on $[c_k, c_{k+1}[$,

$$S_k (f(c_{k+1}) - f(c_k)) = \int_{c_k}^{c_{k+1}} f'(x) (F_\mu(x) - F_\nu(x)) dx.$$

It remains to notice that $S_{2n} = F_\mu(c_{2n}) - F_\nu(c_{2n}) = 0$. \square

The following lemma is a direct consequence of interlacing inequalities for singular values obtained by Thompson [20] in 1976. It was also obtained by Bai [3] and generalized by Benaych-Georges and Rao [6]. It is worthwhile to mention that it gives neither an upper bound for $s_1(\mathbf{B}), \dots, s_k(\mathbf{B})$ nor a lower bound for $s_{n-k+1}(\mathbf{B}), \dots, s_n(\mathbf{B})$ where $k := \text{rank}(\mathbf{A} - \mathbf{B})$, even in the case $k = 1$.

Lemma 3.3 (Rank inequality). *Let \mathbf{A} and \mathbf{B} be two $n \times m$ complex matrices. Let $F_{\sqrt{\mathbf{A}\mathbf{A}^*}}, F_{\sqrt{\mathbf{B}\mathbf{B}^*}}$ be the cumulative distribution functions of $\mu_{\sqrt{\mathbf{A}\mathbf{A}^*}}$ and $\mu_{\sqrt{\mathbf{B}\mathbf{B}^*}}$. Then*

$$\|F_{\sqrt{\mathbf{A}\mathbf{A}^*}} - F_{\sqrt{\mathbf{B}\mathbf{B}^*}}\|_\infty \leq \frac{1}{n} \text{rank}(\mathbf{A} - \mathbf{B}).$$

The following theorem is due to Tao and Vu [18, Theorem 2.1], and is inspired from the work of Rudelson and Vershynin [15].

Theorem 3.4 (Polynomial bounds for smallest singular values). *Let L be a probability distribution on \mathbb{C} with finite and non-zero variance. For every constants $A > 0$ and $C_1 > 0$, there exists constants $B > 0$ and $C_2 > 0$ such that for every $n \times n$ random matrix \mathbf{X} with i.i.d. entries of law L and every $n \times n$ deterministic matrix \mathbf{C} with $s_1(\mathbf{C}) \leq n^{C_1}$, we have*

$$\mathbb{P}(s_n(\mathbf{X} + \mathbf{C}) \leq n^{-B}) \leq C_2 n^{-A}.$$

Acknowledgments. The final form of this note benefited from the comments of anonymous referees. Part of this work was done during two visits to the LABORATOIRE JEAN DIEUDONNÉ in Nice, France. The author would like to thank Pierre DEL MORAL and Persi DIACONIS for their kind hospitality there, and also Neil O'CONNELL for his encouragements.

References

- [1] A. L. Andrew, *Eigenvalues and singular values of certain random matrices*, J. Comput. Appl. Math. **30** (1990), no. 2, 165–171.
- [2] Z. D. Bai, *Circular law*, Ann. Probab. **25** (1997), no. 1, 494–529.
- [3] ———, *Methodologies in spectral analysis of large-dimensional random matrices, a review*, Statist. Sinica **9** (1999), no. 3, 611–677, With comments by G. J. Rodgers and J. W. Silverstein; and a rejoinder by the author.
- [4] Z. D. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Mathematics Monograph Series 2, Science Press, Beijing, 2006.
- [5] Z. D. Bai and Y. Q. Yin, *Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix*, Ann. Probab. **21** (1993), no. 3, 1275–1294.
- [6] F. Benaych-Georges and R. N. Rao, *The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices*, preprint <http://arxiv.org/abs/0910.2120>, 2009.
- [7] D. Chafaï, *Aspects of large random Markov kernels*, Stochastics **81** (2009), no. 3-4, 415–429. MR MR2549497
- [8] ———, *The Dirichlet Markov Ensemble*, Journal of Multivariate Analysis **101** (2010), 555–567.
- [9] A. Edelman, *The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law*, J. Multivariate Anal. **60** (1997), no. 2, 203–232.

- [10] V. L. Girko, *The circular law*, Teor. Veroyatnost. i Primenen. **29** (1984), no. 4, 669–679.
- [11] F. Götze and A. Tikhomirov, *The Circular Law for Random Matrices*, to appear in the Annals of Probability [arXiv:0709.3995](#) [[math.PR](#)], 2010.
- [12] C.-R. Hwang, *A brief survey on the spectral radius and the spectral distribution of large random matrices with i.i.d. entries*, Random matrices and their applications (Brunswick, Maine, 1984), Contemp. Math., vol. 50, Amer. Math. Soc., Providence, RI, 1986, pp. 145–152. MR [MR841088](#) (87m:60080)
- [13] M. L. Mehta, *Random matrices and the statistical theory of energy levels*, Academic Press, New York, 1967.
- [14] G. Pan and W. Zhou, *Circular law, extreme singular values and potential theory*, Journal of Multivariate Analysis **101** (2010), 645–656.
- [15] M. Rudelson and R. Vershynin, *The Littlewood-Offord problem and invertibility of random matrices*, Adv. Math. **218** (2008), no. 2, 600–633. MR [MR2407948](#)
- [16] J. W. Silverstein, *The spectral radii and norms of large-dimensional non-central random matrices*, Comm. Statist. Stochastic Models **10** (1994), no. 3, 525–532.
- [17] P. Śniady, *Random regularization of Brown spectral measure*, J. Funct. Anal. **193** (2002), no. 2, 291–313.
- [18] T. Tao and V. Vu, *Random matrices: the circular law*, Commun. Contemp. Math. **10** (2008), no. 2, 261–307. MR [MR2409368](#) (2009d:60091)
- [19] ———, *Random matrices: Universality of ESDs and the circular law*, preprint [arXiv:0807.4898](#) [[math.PR](#)] to appear in the Annals of Probability, 2010.
- [20] R. C. Thompson, *The behavior of eigenvalues and singular values under perturbations of restricted rank*, Linear Algebra and Appl. **13** (1976), no. 1/2, 69–78, Collection of articles dedicated to Olga Taussky Todd.

Djalil CHAFAÏ **E-mail:** [djalil\(at\)chafai.net](mailto:djalil(at)chafai.net)

LABORATOIRE D'ANALYSE ET DE MATHÉMATIQUES APPLIQUÉES

UMR 8050 CNRS UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE

5 BOULEVARD DESCARTES, F-77454 CEDEX 2, CHAMPS-SUR-MARNE, FRANCE.