Spectral law of the sum of random matrices

Florent Benaych-Georges
benaych@dma.ens.fr

May 5, 2005

Abstract

The spectral distribution of a matrix is the uniform distribution on its spectrum (with multiplicity). It has been proved in [5] that if for each $d \geq 1$, A_d, B_d are independant unitarily invariant $d{\times}d$ hermitian random matrices whose spectral distributions converge in probability to probability measures μ, ν, then under the hypothesis of boundness of the sequence $\mathbb{E}\left(\operatorname{tr}|A_d|\right)$ (where $\mathbb{E}$ designs the expectation and tr the normalized trace), the spectral distribution of $A_d + B_d$ converges in probability to the free convolution $\mu \boxplus \nu$ of μ and ν. We will show that this result stays true without the assumption of boundness of $\mathbb{E}\left(\operatorname{tr}|A_d|\right)$.

1 Preliminaries on the convergence in probability of random distributions

Recall that a sequence $(X_d)_{d\geq 1}$ of random variables with values in a topological space $(T, \mathcal{T})$ metrizable with the distance d is said to converge weakly to a point a of T if for each $\varepsilon > 0$, the probability of the event $\{\mathrm{d}(X_d, a) > \varepsilon\}$ tends to zero as d tends to infinity. This definition does not depend on the choice of the distance d that defines $\mathcal{T}$.

When ρ is a probability measure on the real line, its Stieltjes transform f_ρ is the analytic function defined on the upper half plane by

$$f_\rho(z) = \int_{u\in\mathbb{R}} \frac{\mathrm{d}\rho(u)}{u-z}.$$

Then we can define a distance on the set of probability measures on the real line by

$$(\rho_1, \rho_2) \mapsto \sup_{\Im z \geq 1} |f_{\rho_1}(z) - f_{\rho_2}(z)|,$$

where $\Im z$ denotes the imaginary part of a complex number z. This distance defines the topology of weak convergence (see in [1] or [6]). The Stieltjes transform of the spectral distribution of an hermitian matrix is the normalized trace of its resolvant $\mathfrak{R}_z(M) = (M-z)^{-1}$. So the spectral distribition of a sequence (M_d) of random matrices converges in probability to a probability measure ρ on the real line if and only if for each $\varepsilon > 0$, the probability of the event

$$\left\{ \sup_{\Im z \geq 1} |\mathrm{tr}\, \mathfrak{R}_z (M_d) - f_\rho(z)| > \varepsilon \right\}$$

tends to zero as d tends to infinity.

We will need the following lemma. When F is a real Borel function on the real line, for any hermitian matrix M, we will denote by $F(M)$ the matrix defined by the functionnal calculus, and for any probability measure ρ, we will denote by $F(\rho)$ the image measure of ρ by F, that is the distribution of the random variable $F(X)$, when X is a random variable with distribution ρ.

Lemma 1.1 *Let (M_d) be a sequence of hermitian random matrices whose spectral distribution converges in probability to a probability measure ρ. Then for any real function F continuous at ρ-almost every point of the real line, the spectral distribution of $F(M_d)$ converges in probability to $F(\rho)$.*

Proof It suffices to see that the spectral distribution of $F(M_d)$ is the image of the spectral distribution of M_d by F, and that the application $\mu \mapsto F(\mu)$ on the set of probability measures on the real line is weakly continuous at ρ because F is continuous at ρ-almost every point of the real line (see [3]). $\square$

2 Spectral distribution of the sum of random matrices

The main result of the article is the following theorem.

Theorem 2.1 *Let, for each $d \geq 1$, A_d, B_d be two independant unitarily invariant hermitian $d{\times}d$ random matrices whose spectral distributions converge to two probability measures μ, ν on the real line. Then the spectral law of $A_d + B_d$ converges to the free convolution $\mu \boxplus \nu$ of μ and ν.*

Remark

1. The reader can find the definition and properties of the free convolution $\boxplus$ in [2] (or in [4], but restricted to the compactly supported measures).
2. The proof is based on approximation by random matrices whose rank is bounded with a big probability. This idea can be used to prove the same kind of result (wich are already known when the measures are compactly supported) for the product, or the commutator of random matrices.

Proof of the theorem Let $\varepsilon > 0$. We will show that

$$\mathrm{P}\left\{\sup_{\Im z\geq 1}\left|\ \mathrm{tr}\left(\mathfrak{R}_z(A_d+B_d)\right)-f_{\mu\boxplus\nu}(z)\right|\ >\ \varepsilon\right\}\overset{d\to\infty}{\longrightarrow}0,$$

where P designs the probability measure of the probability space where the random matrices are defined.

Let us define, for each positive number t,

$$\begin{array}{rcl} F_t\ :\ \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} x & \text{if } |x|\leq t, \\ 0 & \text{elseif.}\end{cases}\end{array}$$

For every probability measure μ, $F_t(\mu)$ converges weakly to μ when t tends to infinity, so, by continuity of $\boxplus$, there exists $t\in]0,+\infty[$ such that t and $-t$ are not atoms of the measures μ and ν, and such that

$$\sup_{\Im z\geq 1}\left|f_{F_t(\mu)\boxplus F_t(\nu)}(z)-f_{\mu\boxplus\nu}(z)\right|\ <\ \frac{\varepsilon}{3} \tag{1}$$

$$\mu\Big(\mathbb{R}\backslash[-t-1,t+1]\Big)+\nu\Big(\mathbb{R}\backslash[-t-1,t+1]\Big)\ <\ \frac{\varepsilon}{12}. \tag{2}$$

We will now use the notations $F_t = F$ and $G(x) = x - F(x)$.

Using triangular inequality and inequality (1), we have

$$\begin{aligned} & \mathrm{P}\left\{\sup_{\Im z\geq 1}\left|\ \mathrm{tr}\left(\mathfrak{R}_z(A_d+B_d)\right)-f_{\mu\boxplus\nu}(z)\right|\ >\ \varepsilon\right\} \\ \leq\ & \mathrm{P}\left\{\sup_{\Im z\geq 1}\left|\ \mathrm{tr}\left[\mathfrak{R}_z(A_d+B_d)-\mathfrak{R}_z[F(A_d)+F(B_d)]\right]\right|\ >\ \tfrac{\varepsilon}{3}\right\} \\ +\ & \mathrm{P}\left\{\sup_{\Im z\geq 1}\left|\ \mathrm{tr}\,\mathfrak{R}_z[F(A_d)+F(B_d)]-f_{F(\mu)\boxplus F(\nu)}(z)\right|\ >\ \tfrac{\varepsilon}{3}\right\} \end{aligned} \tag{3}$$

But forall M, N hermitian matrices, for all complex number z of the upper half plane,

$$\mathfrak{R}_z(M+N)-\mathfrak{R}_z(M) = -\mathfrak{R}_z(M+N)N\mathfrak{R}_z(M).$$

Applying this to $M = F(A_d)+F(B_d)$, $N = G(A_d)+G(B_d)$, what gives $M+N = A_d+B_d$, we have

$$\begin{gathered}\mathfrak{R}_z(A_d+B_d)-\mathfrak{R}_z[F(A_d)+F(B_d)] = \\ -\mathfrak{R}_z(A_d+B_d)\big[G(A_d)+G(B_d)\big]\mathfrak{R}_z[F(A_d)+F(B_d)]\end{gathered}$$

But for every complex $d{\times}d$ matrix T, $|\,\mathrm{tr}\,T|\leq\frac{1}{d}||T||\,\mathrm{rg}\,T$, where $||.||$ designs the operator norm associated with the canonical hermitian norm, and rg designs the rank. We apply this with

$$T = -\mathfrak{R}_z(A_d+B_d)\big[G(A_d)+G(B_d)\big]\mathfrak{R}_z[F(A_d)+F(B_d)].$$

We have

$$\operatorname{rg} T \leq 2 \operatorname{rg} G(A_d) + 2 \operatorname{rg} G(B_d),$$

and, as T can be written $\mathfrak{R}_z(A_d + B_d) - \mathfrak{R}_z[F(A_d) + F(B_d)]$, we have

$$||T|| \leq \frac{2}{\Im z} \leq 2$$

when $\Im z \geq 1$, hence

$$\mathrm{P}\left\{ \sup_{\Im z \geq 1} \left| \operatorname{tr} \left[\mathfrak{R}_z(A_d + B_d) - \mathfrak{R}_z[F(A_d) + F(B_d)] \right] \right| \; > \; \tfrac{\varepsilon}{3} \right\}$$
$$\leq \mathrm{P}\left\{ \tfrac{2}{d}\big(\operatorname{rg} G(A_d) + \operatorname{rg} G(B_d)\big) \; > \; \tfrac{\varepsilon}{6} \right\}.$$

But denoting by μ_{A_d}, μ_{B_d} the spectral distributions of A_d, B_d, the event

$$\left\{ \frac{2}{d}\big(\operatorname{rg} G(A_d) + \operatorname{rg} G(B_d)\big) \; > \; \frac{\varepsilon}{6} \right\}$$

is equal to the event

$$\left\{ \mu_{A_d}(\mathbb{R}\backslash[-t,t]) + \mu_{B_d}(\mathbb{R}\backslash[-t,t]) > \frac{\varepsilon}{12} \right\},$$

whose probability tends to zero when d tends to infinity (by inequality (2)). Hence,

$$\lim_{d\to\infty} \mathrm{P}\left\{ \sup_{\Im z \geq 1} \left| \operatorname{tr} \left[\mathfrak{R}_z(A_d + B_d) - \mathfrak{R}_z(F(A_d) + F(B_d)) \right] \right| \; > \; \frac{\varepsilon}{3} \right\} = 0 \quad (4)$$

Moreover, by Theorem 2.1 of [5] and the lemma 1.1 (wich can be applied because t and $-t$ are not atoms of the measures μ and ν), we have

$$\lim_{d\to\infty} \mathrm{P}\left\{ \sup_{\Im z \geq 1} \left| \operatorname{tr} \mathfrak{R}_z(F(A_d) + F(B_d)) - f_{F(\mu)\boxplus F(\nu)}(z) \right| \; > \; \frac{\varepsilon}{3} \right\} = 0 \quad (5)$$

Now we can conclude : according to inequality (3), and using inequalities (4) and (5), we have

$$\lim_{d\to\infty} \mathrm{P}\left\{ \sup_{\Im z \geq 1} \left| \operatorname{tr} \Big(\mathfrak{R}_z(A_d + B_d) \Big) - f_{\mu\boxplus\nu}(z) \right| \; > \; \varepsilon \right\} \leq 0 + 0 = 0.$$

□

References

[1] Akhiezer, N.I. *The classical moment problem*, Moscou, 1961

[2] Bercovici, H., Voiculescu, D. *Free convolution of measures with unbounded supports* Indiana Univ. Math. J. **42** (1993) 733-773

[3] Billingsley, P. *Convergence of probability measures* Wiley, 1968

[4] Hiai, F., Petz, D. *The semicircle law, free random variables, and entropy* Amer. Math. Soc., Mathematical Surveys and Monographs Volume **77**, 2000

[5] Pastur, L., Vasilchuk, V. *On the law of addition of random matrices* Commun. Math. Phys. 214 249-286 (2000)

[6] Pastur, L., Lejay, A. *Matrices alatoires : statistique asymptotique des valeurs propres* Seminaire de Probabilits XXXVI, Lecutre notes in M. **1801**, Springer 2002