J Theor Probab (2007) 20: 381–395
DOI 10.1007/s10959-007-0097-7

Berry–Esseen for Free Random Variables

Vladislav Kargin

Received: 8 May 2006 / Revised: 6 October 2006 /
Published online: 26 April 2007

Abstract An analogue of the Berry–Esseen inequality is proved for the speed of convergence of free additive convolutions of bounded probability measures. The obtained rate of convergence is of the order $n^{-1/2}$, the same as in the classical case. An example with binomial measures shows that this estimate cannot be improved without imposing further restrictions on convolved measures.

Keywords Berry–Esseen inequality · Free probability · Central limit theorem · Speed of convergence

1 Introduction

In the setting of non-commutative probability theory, an interesting new operation on probability measures was defined. If two probability measures, μ_1 and μ_2, are represented as spectral measures of two free self-adjoint operators, X_1 and X_2, then the sum $X_1 + X_2$ has the spectral probability measure μ_3, which is called *free convolution* of μ_1 and μ_2, and denoted as $\mu_1 \boxplus \mu_2$. The theory of additive free convolutions was invented by Voiculescu [9, 10], and an excellent account of further development of free probability theory can be found in [5]. Free convolution has many properties similar to the usual convolution of probability measures. In particular, an analogue of the Central Limit Theorem holds, which says that multiple free convolutions of a probability measure with itself converge to the Wigner semicircle law. In this paper we investigate the speed of convergence and establish an inequality similar to the classical Berry–Esseen inequality.

V. Kargin is grateful to Diana Bloom for her help with editing this paper.

V. Kargin (✉)
Courant Institute of Mathematical Sciences, 109-20 71st Road, Apt. 4A, New York,
NY 11375, USA
e-mail: kargin@cims.nyu.edu

Without reference to operator theory, the free convolution can be defined as follows. Suppose μ_1 and μ_2 are two probability measures compactly supported on the real line. Define the *Cauchy transform* of μ_i as

$$G_i(z) = \int_{-\infty}^{\infty} \frac{d\mu_i(t)}{z-t}.$$

Each of $G_i(z)$ is well-defined and univalent for large enough z and we can define its functional inverse, which is well-defined in a neighborhood of 0. Let us call this inverse the *K-function* of μ_i and denote it as $K_i(z)$:

$$K_i(G_i(z)) = G_i(K_i(z)) = z.$$

Then, we define $K_3(z)$ by the following formula:

$$K_3(z) = K_1(z) + K_2(z) - \frac{1}{z}. \tag{1}$$

It turns out that $K_3(z)$ is the K-function of a probability measure, μ_3, which is the *free convolution* of μ_1 and μ_2.

Let us turn to issues of convergence. Let $d(\mu_1, \mu_2)$ denote the Kolmogorov distance between the probability measures μ_1 and μ_2. That is, if $\mathcal{F}_1$ and $\mathcal{F}_2$ are the distribution functions corresponding to measures μ_1 and μ_2, respectively, then

$$d(\mu_1, \mu_2) =: \sup_{x\in\mathbb{R}} |\mathcal{F}_1(x) - \mathcal{F}_2(x)|.$$

Let μ be a probability measure with the zero mean and unit variance and let m_3 be its third absolute moment. Then the classical Berry–Esseen inequality says that

$$d(\mu^{(n)}, \nu) \le C m_3 \frac{1}{\sqrt{n}},$$

where ν is the standard Gaussian measure and $\mu^{(n)}$ is the normalized n-time convolution of measure μ with itself:

$$\mu^{(n)}(du) = \mu * \cdots * \mu(\sqrt{n}du).$$

This inequality was proved by Berry [3] and Esseen [4] for a more general situation of independent but not necessarily identical measures. A simple example with binomial measures shows that in this inequality the order of $n^{-1/2}$ cannot be improved without further restrictions.

We aim to derive a similar inequality when the usual convolution of measures is replaced by free convolution. Namely, let

$$\mu^{(n)}(du) = \mu \boxplus \cdots \boxplus \mu(\sqrt{n}du)$$

and let ν denote the standard semicircle distribution. It is known that $\mu^{(n)}$ converges weakly to ν [6, 8, 9, 11]. We are interested in the speed of this convergence and we prove that if μ is supported on $[-L, L]$, then

$$d(\mu^{(n)}, \nu) \leq CL^3 \frac{1}{\sqrt{n}}. \tag{2}$$

An example shows that the rate of $n^{-1/2}$ cannot be improved without further restrictions, similar to the classical case.

The main tool in our proof of inequality (2) is Bai's theorem [1] that relates the supremum distance between two probability measures to a distance between their Cauchy transforms. To estimate the distance between Cauchy transforms, we use the fact that as n grows, the K-function of $\mu^{(n)}$ approaches the K-function of the semicircle law. Therefore, the main problem in our case is to investigate whether the small distance between K-functions implies a small distance between the Cauchy transforms themselves. We approach this problem using the Lagrange formula for functional inverses.

The rest of the paper is organized as follows: Sect. 2 contains the formulation and the proof of the main result. It consists of several subsections. In Sect. 2.1 we formulate the result and outline its proof. Section 2.2 evaluates how fast the K-function of $\mu^{(n)}$ approaches the K-function of the semicircle law. Section 2.3 provides useful estimates on behavior of the Cauchy transform of the semicircle law and related functions. Section 2.4 introduces a functional equation for the Cauchy transforms and concludes the proof by estimating how fast the Cauchy transform of $\mu^{(n)}$ converges to the Cauchy transform of the semicircle law. An example in Sect. 3 shows that the rate of $n^{-1/2}$ cannot be improved.

2 Formulation and Proof of the Main Result

2.1 Formulation and Outline of the Proof

Let the *semicircle law* be the probability measure on the real line that has the following cumulative distribution function:

$$\Phi(x) = \frac{1}{\pi} \int_{-\infty}^{x} \sqrt{4 - t^2} \chi_{[-2,2]}(t) dt,$$

where $\chi_{[-2,2]}(t)$ is the characteristic function of the interval $[-2, 2]$.

Theorem 1 *Suppose that μ is a probability measure that has zero mean and unit variance, and is supported on interval $[-L, L]$. Let $\mu^{(n)}$ be the normalized n-time free convolution of measure μ with itself: $\mu^{(n)}(du) = \mu \boxplus \cdots \boxplus \mu(\sqrt{n} du)$. Let $\mathcal{F}_n(x)$ denote the cumulative distribution function of $\mu^{(n)}$. Then for large enough n the following bound holds:*

$$\sup_x |\mathcal{F}_n(x) - \Phi(x)| \leq CL^3 n^{-1/2},$$

where C is an absolute constant.

Remark $C = 2^{16}$ will do, although this constant is far from the best possible.

Proof First, we quote one of Bai's results:

Theorem 2 [1] *Let measures with distribution functions $\mathcal{F}$ and $\mathcal{G}$ be supported on a finite interval $[-B, B]$ and let their Cauchy transforms be $G_{\mathcal{F}}(z)$ and $G_{\mathcal{G}}(z)$, respectively. Let $z = u + iv$. Then*

$$\sup_x |\mathcal{F}(x) - \mathcal{G}(x)| \leq \frac{1}{\pi(1-\kappa)(2\gamma - 1)} \left[\int_{-A}^{A} |G_{\mathcal{F}}(z) - G_{\mathcal{G}}(z)| du + \frac{1}{v} \sup_x \int_{|y| \leq 2vc} |\mathcal{G}(x+y) - \mathcal{G}(x)| dy \right], \tag{3}$$

where $A > B$, $\kappa = \frac{4B}{\pi(A-B)(2\gamma-1)} < 1$, $\gamma > 1/2$, and c and γ are related by the following equality

$$\gamma = \frac{1}{\pi} \int_{|u|<c} \frac{1}{1+u^2} du.$$

Note that if $\mathcal{G}(x)$ is the semicircle distribution then $|\mathcal{G}'(x)| \leq \pi^{-1}$. Therefore $|\mathcal{G}(x+y) - \mathcal{G}(x)| \leq |y|/\pi$. Integrating this inequality, we obtain:

$$\frac{1}{v} \sup_x \int_{|y| \leq 2vc} |\mathcal{G}(x+y) - \mathcal{G}(x)| dy \leq \frac{4c^2}{\pi} v. \tag{4}$$

Hence, the main question is how fast v can be made to approach zero if the first integral in (3) is also required to approach zero.

Let G_Φ and G_n be the Cauchy transforms of the semicircle law and $\mu^{(n)}$, respectively. Assume for the moment that the following lemma holds:

Lemma 3 *Suppose $v = 2^{10} L^3 / \sqrt{n}$. Then for all sufficiently large n, we have the following estimate*:

$$\int_{-8}^{8} |G_n(u+iv) - G_\Phi(u+iv)| du \leq \frac{2^{10} L^3}{\sqrt{n}}.$$

Let us apply Bai's theorem using Lemma 3 and inequality (4). Let $\Phi(x)$ and $\mathcal{F}_n(x)$ denote the cumulative distribution functions of the semicircle law and $\mu^{(n)}$, respectively. The semicircle law is supported on $[-2, 2]$, and for any fixed interval I that includes $[-2, 2]$, we can find such n_0 that $\mu^{(n)}$ is supported on I for all $n \geq n_0$ (see [2]). Suppose that n is so large that $\mu^{(n)}$ is supported on $[-2^{5/4}, 2^{5/4}]$. Then we can take $A = 8$, $B = 2^{5/4}$, $c = 6$, and calculate $\gamma = 0.895$ and $\kappa = 0.682$. Then Bai's theorem gives the following estimate:

$$\sup_x |\mathcal{F}_n(x) - \Phi(x)| \leq 1.268(2^{10} L^3 n^{-1/2} + 46937 L^3 n^{-1/2}) \leq 2^{16} L^3 n^{-1/2}.$$

□

Thus, the main task is to prove Lemma 3. Here is the plan of the proof. First, we estimate how close the K-functions of $\mu^{(n)}$ and Φ are to each other. Then we note that the Cauchy transforms of $\mu^{(n)}$ and Φ can be found from their functional equations:

$$K_n(G_n(z)) = z,$$

and

$$K_\Phi(G_\Phi(z)) = z,$$

where K_n and K_Φ denote the K-functions of $\mu^{(n)}$ and Φ. From the previous step we know that $K_n(z)$ is close to $K_\Phi(z)$. Our goal is to show that this implies that $G_n(z)$ is close to $G_\Phi(z)$.

If we introduce an extra parameter, t, then we can include these functional equations in a parametric family:

$$K_t(G_t(z)) = z. \tag{5}$$

Parameter $t = 0$ corresponds to Φ and $t = 1$ to $\mu^{(n)}$. Next, we fix z and consider G_t as a function of t. We develop this function in a power series in t:

$$G_t = G_\Phi + \sum_{k=1}^{\infty} c_k t^k,$$

where c_k are functions of z. Then we estimate I_k for each $k \geq 1$, where

$$I_k = \int_{-8}^{8} |c_k(u + iv)| du.$$

Then,

$$\int_{-8}^{8} |G_n(u + iv) - G_\Phi(u + iv)| du \leq \sum_{k=1}^{\infty} I_k,$$

and our estimates of I_k allow us to prove the claim of Lemma 3.

2.2 Speed of Convergence of K-functions

In this section, we start with some preliminary facts about functional inverses of holomorphic functions. Then we derive an estimate for the speed of convergence of the K-functions of $\mu^{(n)}$ and the semicircle law.

First, recall the Lagrange formula for the functional inversion. By a function holomorphic in a domain D we mean a function which is bounded and complex-differentiable in D.

Lemma 4 (Lagrange's inversion formula) *Suppose f is a function of a complex variable, which is holomorphic in a neighborhood of z_0 and suppose that $f(z_0) = 0$ and $f'(z_0) \neq 0$. Then the functional inverse of $f(z)$ is well defined in a neighborhood of* 0 *and the Taylor series of the inverse is given by the following formula*:

$$f^{-1}(u) = z_0 + \sum_{k=1}^{\infty} \left[\frac{1}{k} \mathrm{res}_{z=z_0} \frac{1}{f(z)^k} \right] u^k,$$

where $\mathrm{res}_{z=z_0}$ *denotes the Cauchy residual at point* z_0.

For proof see Theorems II.3.2 and II.3.3 in [7], or Sect. 7.32 in [12]. We also need the following modification of the Lagrange formula.

Lemma 5 *Suppose* G *is a function of a complex variable, which is holomorphic in a neighborhood of* $z_0 = \infty$ *and has the expansion*

$$G(z) = \frac{1}{z} + \frac{a_1}{z^2} + \cdots,$$

converging for all sufficiently large z. *Define* $g(z) = G(1/z)$. *Then the functional inverse of* $G(z)$ *is well defined in a neighborhood of* 0 *and the Laurent series of the inverse is given by the following formula*:

$$G^{-1}(w) = \frac{1}{w} + a_1 - \sum_{n=1}^{\infty} \left[\frac{1}{n} \frac{1}{2\pi i} \oint_{\partial\gamma} \frac{dz}{z^2 g(z)^n} \right] w^n,$$

where γ *is a closed disc around* 0 *in which* $g(z)$ *has only one zero.*

Proof Let γ be a disc around 0 in which $g(z)$ has only one zero. This disk exists because $g(0) = 0$, and $g(z)$ is analytical in a neighborhood of 0 and has a non-zero derivative at 0. Let

$$r_w = \frac{1}{2} \inf_{z \in \partial\gamma} |g(z)|.$$

Then $r_w > 0$ by our assumption about γ. We can apply Rouché's theorem and conclude that the equation $g(z) - w = 0$ has only one solution inside γ if $|w| \le r_w$. Let us fix such w that $|w| \le r_w$. Inside γ, the function

$$\frac{g'(z)}{z(g(z) - w)}$$

has a pole at $z = 1/G^{-1}(w)$ with the residual $G^{-1}(w)$ and a pole at $z = 0$ with the residual $-1/w$. Consequently, we can write:

$$G^{-1}(w) = \frac{1}{2\pi i} \oint_{\partial\gamma} \frac{g'(z)dz}{z(g(z) - w)} + \frac{1}{w}.$$

The integral can be re-written as follows:

$$\oint_{\partial\gamma} \frac{g'(z)dz}{z(g(z) - w)} = \oint_{\partial\gamma} \frac{g'(z)}{zg(z)} \frac{1}{1 - \frac{w}{g(z)}} dz = \sum_{n=0}^{\infty} \oint_{\partial\gamma} \frac{g'(z)dz}{zg(z)^{n+1}} w^n.$$

For $n=0$ we calculate

$$\frac{1}{2\pi i}\oint_{\partial\gamma}\frac{g'(z)dz}{zg(z)}=a_1.$$

Indeed, the only pole of the integrand is at $z=0$ and it has order two. The corresponding residual can be computed from the series expansion for $g(z)$:

$$\operatorname{res}_{z=0}\frac{g'(z)dz}{zg(z)}=\frac{d}{dz}\frac{z^2(1+2a_1z+\cdots)}{z(z+a_1z^2+\cdots)}\bigg|_{z=0}=\frac{d}{dz}\frac{1+2a_1z+\cdots}{1+a_1z+\cdots}\bigg|_{z=0}=a_1.$$

For $n>0$ we integrate by parts:

$$\frac{1}{2\pi i}\oint_{\partial\gamma}\frac{g'(z)dz}{zg(z)^{n+1}}=-\frac{1}{2\pi i}\frac{1}{n}\oint_{\partial\gamma}\frac{dz}{z^2g(z)^n}.$$ □

Our first application of the Lagrange formula is the following estimate on the convergence radius of power series for K-functions.

Lemma 6 *Suppose that the measure μ is supported on interval $[-L,L]$, and $K(z)$ denotes the functional inverse of its Cauchy transform $G(z)$. Then the Laurent series of $K(z)$ converge in the area $\Omega=\{z:0<|z|<(4L)^{-1}\}$.*

Proof Let us apply Lemma 5 to $G(z)$ with circle γ having radius $(2L)^{-1}$. We need to check that $g(z)=:G(1/z)$ has only one zero inside this circle. This holds because

$$g(z)=z(1+a_2z^2+a_3z^3+\cdots),$$

and inside $|z|\le(2L)^{-1}$ we can estimate:

$$|a_2z^2+a_3z^3+\cdots|\le L^2\Big(\frac{1}{2L}\Big)^2+L^3\Big(\frac{1}{2L}\Big)^3+\cdots=\frac{1}{2}, \tag{6}$$

and an application of Rouché's theorem shows that $g(z)$ has only one zero inside this circle.

Another consequence of the estimate (6) is that on the circle $|z|=(2L)^{-1}$

$$|g(z)|\ge|z|/2=1/(4L).$$

By Lemma 5 the coefficients in the series for the inverse of $G(z)$ are

$$b_k=\frac{1}{2\pi ik}\oint_{\partial\gamma}\frac{dz}{z^2g(z)^k},$$

and we can estimate them as

$$|b_k|\le\frac{2L}{k}(4L)^k. \tag{7}$$

This implies that the radius of convergence of power series for $K(z)$ is at least $(4L)^{-1}$. □

Let $K_n(z)$ denote the K-function of $\mu^{(n)}$. For the semicircle law the K-function is $K_\Phi(z) = z^{-1} + z$. Define $\varphi_n(z) = K_n(z) - z - z^{-1}$.

Lemma 7 *Suppose μ has zero mean and unit variance and is supported on* $[-L, L]$. *Then the function $\varphi_n(z)$ is holomorphic in $|z| \le \sqrt{n}/(8L)$ and*

$$|\varphi_n(z)| \le 32L^3 \frac{|z|^2}{\sqrt{n}}.$$

Proof The measure $\mu^{(n)}$ is the n-time convolution of the measure $\widetilde{\mu}(dx) =: \mu(\sqrt{n}dx)$ with itself. Therefore, $K_n(z) = nK_{\widetilde{\mu}}(z) - (n-1)z^{-1}$. Since $\widetilde{\mu}$ is supported on $[-L/\sqrt{n}, L/\sqrt{n}]$, we can estimate $K_{\widetilde{\mu}}(z) - \frac{1}{z} - \frac{1}{n}z$ inside the circle $|z| = \sqrt{n}/(8L)$ using the estimates for coefficients of $K_{\widetilde{\mu}}(z)$ in (7) and changing L to $L/\sqrt{n}$ in these estimates:

$$\begin{aligned}\left|K_{\widetilde{\mu}}(z) - \frac{1}{z} - \frac{1}{n}z\right| &= \sum_{k=2}^{\infty} b_k z^k \le \frac{2L}{\sqrt{n}} \sum_{k=2}^{\infty} \frac{1}{k}\left(\frac{4L}{\sqrt{n}}\right)^k |z|^k \\ &\le 32\left(\frac{L}{\sqrt{n}}\right)^3 |z|^2 \sum_{k=2}^{\infty} \frac{1}{k2^{k-1}} \le 32\left(\frac{L}{\sqrt{n}}\right)^3 |z|^2.\end{aligned}$$

Note that we used the assumption about the mean and variance of the measure μ in the first line by setting $b_1 = 0$ and $b_1 = 1/n$.

Using the summation formula (1) for K-functions, we further obtain:

$$\left|K_n(z) - \frac{1}{z} - z\right| \le 32\frac{L^3}{\sqrt{n}}|z|^2.$$

□

Lemma 7 shows that as n grows, the radius of the convergence area of $\varphi_n(z)$, and therefore of $K_n(z)$, grows proportionally to $\sqrt{n}$. In particular, the radius of convergence will eventually cover every bounded domain. Lemma 7 also establishes the rate of convergence of $K_n(z)$ to its limit $K_\Phi(z) = z^{-1} + z$.

2.3 Useful Estimates

Suppose $G_\Phi(z)$ is the Cauchy transform of the semicircle distribution.

Lemma 8 (1) $|G_\Phi(z)| \le 1$ *if* $\mathrm{Im} z > 0$;
(2) $|z - 2G_\Phi(z)| \ge 2\sqrt{\mathrm{Im} z}$ *if* $\mathrm{Im} z \in (0, 2)$.

Proof $G_\Phi(z) = (z - \sqrt{z^2 - 4})/2$. If $z = u + iv$ and v is fixed, then the maximum of $|G_\Phi(z)|$ is reached for $u = 0$. Then $|G_\Phi(iv)| = (\sqrt{v^2 + 4} - v)/2$ and $\sup |G_\Phi(iv)| = 1$.

Next, $|z - 2G_\Phi(z)| = |\sqrt{u^2 - v^2 - 4 + i2uv}|$. If v is in $(0, 2)$ and fixed, the minimum of this expression is reached for $u = \pm\sqrt{4 - v^2}$ and equals $2\sqrt{v}$. □

Lemma 9 *If* $n \geq 64L^2$ *and* $\mathrm{Im} z > 0$, *then we have*:

$$|\varphi_n(G_\Phi(z))| \leq \frac{32L^3}{\sqrt{n}}.$$

Proof This Lemma follows directly from Lemmas 7 and 8. □

2.4 Functional Equation for the Cauchy Transform

Let $G_n(z)$ denote the Cauchy transform of $\mu^{(n)}$. Let us write the following functional equation:

$$G(t,z) + \frac{1}{G(t,z)} + t\varphi_n(G(t,z)) = z, \tag{8}$$

where t is a complex parameter. For $t = 0$ the solution is $G_\Phi(z)$, and for $t = 1$ the solution is $G_n(z)$. Assume that $\varphi_n(z)$ is not identically zero. (If it is, then $\mu^{(n)}$ is semicircle and $d(\mu^{(n)}, \nu) = 0$.) Let us write equation (8) as

$$t = \frac{zG - G^2 - 1}{G\varphi_n(G)}. \tag{9}$$

We can think about z as a fixed complex parameter and about t as a function of the complex variable G, i.e., $t = f(G)$. Suppose $\varphi_n(G_\Phi(z))$ does not equal zero for a given value of z. (This holds for all but a countable number of values of parameter z.) Then, as a function of G, f is holomorphic in a neighborhood of $G_\Phi(z)$. What we would like to do is to invert this function f and write $G = f^{-1}(t)$. In particular we would like to develop $f^{-1}(t)$ in a series of t around $t = 0$. Then we would be able to estimate $|f^{-1}(1) - f^{-1}(0)|$, which is equal to $|G_n(z) - G_\Phi(z)|$. To perform this inversion, we use the Lagrange formula in Lemma 4.

Assume that z is fixed, and let us write G instead of $G(z)$ and G_Φ instead of $G_\Phi(z)$. By Lemma 4, we can write the solution of (9) as

$$G = G_\Phi + \sum_{k=1}^{\infty} c_k t^k, \tag{10}$$

where

$$c_k = \frac{1}{k}\mathrm{res}_{G=G_\Phi}\left(\frac{G\varphi_n(G)}{zG - G^2 - 1}\right)^k. \tag{11}$$

We aim to estimate $I_k =: \int_{-8}^{8} |c_k(u+iv)|du$. In particular, we will show that for any $v \in (0,1)$, $I_1 = O(n^{-1/2})$. In addition, we will show that if $v = b/\sqrt{n}$ for a suitably chosen b, then $\sum_{k=2}^{\infty} I_k = o(n^{-1/2})$. This information is sufficient for a good estimate of $\int_{-8}^{8} |G_n(u+iv) - G_\Phi(u+iv)|du$.

Let us consider first the case of $k = 1$. Then

$$c_1 = \frac{G_\Phi\varphi_n(G_\Phi)}{G_2 - G_\Phi} = \frac{G_\Phi\varphi_n(G_\Phi)}{\sqrt{z^2-4}},$$

where G_2 denotes the root of the equation $G^2 - zG + 1 = 0$, which is different from G_Φ. Therefore, if $z = u + iv$, then we can calculate:

$$|c_1| = \frac{|G_\Phi||\varphi_n(G_\Phi)|}{[(u^2-4)^2 + 2(u^2+4)v^2 + v^4]^{1/4}}$$
$$\leq \frac{32L^3}{\sqrt{n}} \frac{1}{[(u^2-4)^2 + 2(u^2+4)v^2 + v^4]^{1/4}},$$

where the last inequality holds by Lemma 9 for all $n \geq 64L^2$.

Lemma 10 *For every* $v \in (0, 1)$,

$$\int_{-8}^{8} [(u^2-4)^2 + 2(u^2+4)v^2 + v^4]^{-1/4} du < 24.$$

Proof Let us make substitution $x = u^2 - 4$. Then we get:

$$J = \int_{-4}^{60} [x^2 + 2(x+8)v^2 + v^4]^{-1/4} \frac{dx}{\sqrt{x+4}} \leq \int_{-4}^{60} \frac{1}{(x^2+v^2)^{1/4}} \frac{dx}{\sqrt{x+4}}.$$

Now we divide the interval of integration in two parts and write:

$$J \leq \int_{-4}^{-2} \cdots + \int_{-2}^{60} \cdots \leq \frac{1}{\sqrt{2}} \int_{-4}^{-2} \frac{dx}{\sqrt{x+4}} + \frac{2}{\sqrt{2}} \int_{0}^{60} \frac{dx}{x^{1/2}} = 2 + \frac{4}{\sqrt{2}}\sqrt{60} < 24.$$

□

Corollary 11 *For every* $v \in (0, 1)$ *and all* $n \geq 64L^2$, *it is true that*

$$I_1 =: \int_{-8}^{8} |c_1(u+iv)| du \leq \frac{768L^3}{\sqrt{n}}.$$

Now we estimate c_k in (10) for $k \geq 2$. Define function $f_k(G)$ by the formula

$$f_k(G) =: \left(\frac{G\varphi_n(G)}{G_2 - G} \right)^k,$$

where G_2 denotes the root of the equation $G^2 - zG + 1 = 0$, which is different from G_Φ. Then formula (11) implies that kc_k equal to the coefficient before $(G - G_\Phi)^{k-1}$ in the expansion of $f_k(G)$ in power series of $(G - G_\Phi)$. To estimate this coefficient, we will use the Cauchy inequality:

$$|kc_k| \leq \frac{M_k(r)}{r^{k-1}},$$

where $M_k(r)$ is the maximum of $|f_k(G)|$ on the circle $|G - G_\Phi| = r$.

We will use $r = \sqrt{v}$ and our first goal is to estimate $M_k(\sqrt{v})$.

Lemma 12 *Let $z = u + iv$ and suppose that $v \in (0, 1)$. If $n \geq 256L^2$, then*

$$M_k(v) \leq \left[\frac{512L^3}{\sqrt{n}} \frac{1}{((u^2-4)^2 + 2u^2v^2)^{1/4}}\right]^k.$$

Proof Note that $|G| \leq |G_\Phi| + \sqrt{v}$ and therefore $|G| \leq 2$ provided that $v \in (0, 1)$. Then Lemma 7 implies that if $n \geq 256L^2$, then $\varphi_n(G)$ is well defined and $|\varphi_n(G)| \leq 128L^3/\sqrt{n}$. It remains to estimate $|G_2 - G|$ from below. If we write $G = G_\Phi + e^{i\theta}\sqrt{v}$, then we have

$$\begin{aligned} |G_2 - G| &= |\sqrt{z^2-4} - e^{i\theta}\sqrt{v}| \geq |\sqrt{z^2-4}| - \sqrt{v} \\ &= ((u^2-4)^2 + 2(u^2+4)v^2 + v^4)^{1/4} - \sqrt{v} \\ &> ((u^2-4)^2 + 2(u^2+4)v^2)^{1/4} - \sqrt{v} > 0. \end{aligned}$$

From the concavity of function $t^{1/4}$ it follows that for positive A and B the following inequality holds:

$$[8(A+B)]^{1/4} - A^{1/4} \geq B^{1/4}.$$

Using v^2 as A, and $[(u^2-4)^2 + 2u^2v^2]/8$ as B, we can write this inequality as follows:

$$((u^2-4)^2 + 2(u^2+4)v^2)^{1/4} - \sqrt{v} \geq \frac{1}{8^{1/4}}((u^2-4)^2 + 2u^2v^2)^{1/4} > 0.$$

Therefore

$$M_k(v) \leq \left[\frac{512L^3}{\sqrt{n}} \frac{1}{((u^2-4)^2 + 2u^2v^2)^{1/4}}\right]^k. \qquad \square$$

Corollary 13 *For every $v \in (0, 1)$, $k \geq 2$, and all $n \geq 256L^2$, it is true that*

$$|kc_k(u+iv)| \leq v^{-\frac{k-1}{2}} \left[\frac{512L^3}{\sqrt{n}} \frac{1}{((u^2-4)^2 + 2u^2v^2)^{1/4}}\right]^k.$$

Now we want to estimate integrals of $|c_k(u+iv)|$ when u changes from -8 to 8. The cases of $k = 2$ and $k > 2$ are slightly different and we treat them separately.

Let

$$I_k =: \int_{-8}^{8} |c_k(u+iv)|du.$$

Lemma 14 *If $v \in (0, 1)$ and $n \geq 256L^2$, then i)*

$$I_2 \leq \frac{\log(60/v)}{\sqrt{v}} \frac{2^{19}L^6}{n},$$

and ii) if $k > 2$, then

$$I_k \leq \frac{12}{k} v^{3/2} \left(\frac{512L^3}{v\sqrt{n}}\right)^k.$$

Proof Using Corollary 13, we write:

$$I_k =: \int_8^8 |c_k(u+iv)|du \le \frac{1}{k}\frac{1}{v^{(k-1)/2}}\left(\frac{512L^3}{\sqrt{n}}\right)^k \int_{-8}^8 \frac{du}{((u^2-4)^2+2u^2v^2)^{k/4}}.$$

After substitution $x = u^2 - 4$, the integral in the right-hand side of the inequality can be re-written as

$$J_k =: \int_{-4}^{60} \frac{1}{(x^2+2(x+4)v^2)^{k/4}}\frac{dx}{\sqrt{x+4}}.$$

We divide the interval of integration into two portions and write the following inequality:

$$J_k \le \int_{-4}^{-1} \frac{dx}{\sqrt{x+4}} + \int_{-1}^{60} \frac{dx}{(x^2+2(x+4)v^2)^{k/4}} \le 2\sqrt{3} + \int_{-1}^{60} \frac{dx}{(x^2+v^2)^{k/4}}.$$

If we use substitution $s = x/v$, then we can write:

$$\int_{-1}^{60} \frac{dx}{(x^2+v^2)^{k/4}} = \int_{-1/v}^{60/v} \frac{ds}{v^{\frac{k}{2}-1}(1+s^2)^{k/4}} \le \frac{2}{v^{\frac{k}{2}-1}} \int_0^{60/v} \frac{ds}{(1+s^2)^{k/4}}.$$

We again separate the interval of integration in two parts and write:

$$\frac{2}{v^{\frac{k}{2}-1}} \int_0^{60/v} \frac{ds}{(1+s^2)^{k/4}} \le \frac{2}{v^{\frac{k}{2}-1}}\left[\int_0^1 ds + \int_1^{60/v} \frac{ds}{s^{k/2}}\right].$$

Here we have two different cases. If $k = 2$, then we evaluate the integrals as $1 + \log(60/v)$. Therefore,

$$J_2 \le 2\sqrt{3} + 2 + 2\log(60/v) \le 4\log(60/v).$$

Hence,

$$I_2 \le \frac{\log(60/v)}{\sqrt{v}}\frac{2^{19}L^6}{n}.$$

If $k > 2$, then we have:

$$\frac{2}{v^{\frac{k}{2}-1}}\left[\int_0^1 ds + \int_1^{60/v} \frac{ds}{s^{k/2}}\right] = \frac{2}{v^{\frac{k}{2}-1}}\left[1 + \frac{1}{-\frac{k}{2}+1}\left(\left(\frac{60}{v}\right)^{-\frac{k}{2}+1} - 1\right)\right]$$
$$\le \frac{2}{v^{\frac{k}{2}-1}}\left(1 + \frac{1}{\frac{k}{2}-1}\right) \le \frac{6}{v^{\frac{k}{2}-1}}.$$

Therefore,

$$J_k \le 2\sqrt{3} + \frac{6}{v^{\frac{k}{2}-1}} \le \frac{12}{v^{\frac{k}{2}-1}},$$

and

$$I_k \leq \frac{1}{k}\frac{1}{v^{(k-1)/2}}\left(\frac{512L^3}{\sqrt{n}}\right)^k \frac{12}{v^{\frac{k}{2}-1}} \leq \frac{12}{k}v^{3/2}\left(\frac{512L^3}{v\sqrt{n}}\right)^k. \qquad \square$$

Corollary 15 *If $v = 1024L^3/\sqrt{n}$, and $n \geq 256L^2$ then*

$$I_2 \leq 2^{14}L^{9/2}\frac{\log(\frac{15}{256L^3}\sqrt{n})}{n^{3/4}}.$$

In particular, $I_2 = o(n^{-1/2})$ as $n \to \infty$.

Now we address the case when $k > 2$.

Corollary 16 *Suppose $v = 1024L^3/\sqrt{n}$, and $n \geq 256L^2$. Then*

$$\sum_{k=3}^{\infty} I_k \leq \frac{3}{2}v^{3/2} = 1536L^{9/2}\frac{1}{n^{3/4}}.$$

In particular, $\sum_{k=2}^{\infty} I_k = o(n^{-1/2})$ as $n \to \infty$.

Joining results of Corollaries 11, 15, and 16, we get the following result.

Lemma 17 *Suppose $v = 1024L^3/\sqrt{n}$. Then for all sufficiently large n, we have the following estimate*:

$$\int_{-8}^{8} |G_n(u+iv) - G_\Phi(u+iv)|du \leq \frac{2^{10}L^3}{\sqrt{n}}.$$

Proof From formula (10) we have

$$|G_n - G_\Phi| \leq \sum_{k=1}^{\infty} |c_k|.$$

Since the series has only positive terms, we can integrate it term by term and write:

$$\begin{aligned}\int_{-8}^{8} |G_n(u+iv) - G_\Phi(u+iv)|du &\leq \sum_{k=1}^{\infty}\int_{-8}^{8} |c_k(u+iv)|du \\ &\leq \frac{768L^3}{\sqrt{n}} + o(n^{-1/2}) \leq \frac{2^{10}L^3}{\sqrt{n}}\end{aligned}$$

for all sufficiently large n. $\square$

Lemma 17 is identical to Lemma 3 and its proof completes the proof of Theorem 1.

3 Example

Consider a binomial measure: $\mu\{-1/p\} = p$ and $\mu\{1/q\} = q \equiv 1 - p$. This is a zero-mean measure with the variance equal to $(pq)^{-1}$. Let $\mu^{(n)}(dx) = \mu \boxplus \cdots \boxplus \mu(\sqrt{\frac{n}{pq}}dx)$ and let $\mathcal{F}_n(x)$ be the distribution function corresponding to $\mu^{(n)}$.

Proposition 18 *If $p \neq q$, then there exist such positive constants C_1 and C_2 that*

$$C_1 n^{-1/2} \leq \sup_x |\mathcal{F}_n(x) - \Phi(x)| \leq C_2 n^{-1/2}$$

for every n.

Proof From the Voiculescu addition formula and the Stieltjes inversion formula, it is easy to compute the density of the distribution of $\mu^{(n)}$:

$$f_n(x) = \frac{1}{2\pi} \frac{\sqrt{4 - x^2 + 2\frac{p-q}{\sqrt{pqn}}x - \frac{1}{pqn}}}{(1 + \frac{x}{\sqrt{nq/p}})(1 - \frac{x}{\sqrt{np/q}})},$$

if the square root is real, and if not, $f_n(x) = 0$. We compare this distribution with the semicircle distribution, which has the following density:

$$\phi(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\chi_{[-2,2]}(x).$$

More precisely, we seek to estimate

$$\sup_x \left| \int_{-\infty}^{x} (f_n(t) - \phi(t))dt \right|.$$

The support of f_n is $[-2\sqrt{1 - n^{-1}} + cn^{-1/2}, 2\sqrt{1 - n^{-1}} + cn^{-1/2}]$, where $c = (p - q)/\sqrt{pq}$. Suppose in the following that $p > q$ and introduce the new variable $u = x + 2\sqrt{1 - n^{-1}} - c/\sqrt{n}$. Then,

$$2\pi f_n(u) = \sqrt{4u\sqrt{1 - n^{-1}} - u^2} \sim \sqrt{4u - u^2},$$

where the asymptotic equivalence is for u fixed and $n \to \infty$ and we omit all terms that are $o(n^{-1/2})$. Similarly,

$$\begin{aligned} 2\pi\phi(x) &= \sqrt{4 - (u - 2(1 - n^{-1}) + cn^{-1/2})^2} \sim \sqrt{4u - u^2 + 4cn^{-1/2} - 2cun^{-1/2}} \\ &= \sqrt{4u - u^2}\sqrt{1 + 2c\frac{1}{\sqrt{n}u}\frac{2 - u}{4 - u}}. \end{aligned}$$

Consequently,

$$\phi(u) - f_n(u) \sim \sqrt{4u - u^2}\left[\sqrt{1 + 2c\frac{1}{\sqrt{n}u}\frac{2-u}{4-u}} - 1\right]$$
$$\sim \sqrt{4u - u^2}c\frac{1}{\sqrt{n}u}\frac{2-u}{4-u} = c\frac{1}{\sqrt{nu}}\frac{2-u}{\sqrt{4-u}}.$$

After integrating we get:

$$\left|\int_0^x (f_n(u) - \phi(u))du\right| \sim c\frac{1}{\sqrt{n}}f(x),$$

where $f(x)$ is a continuous positive bounded function. From this expression it is clear that $\sup_x |\int_{-\infty}^x (f_n(t) - \phi(t))dt|$ has the order of $n^{-1/2}$ provided that $p \neq q$. □

This example shows that the rate of $n^{-1/2}$ in Theorem 1 cannot be improved without further restrictions on measures. It would be interesting to extend Theorem 1 to measures with unbounded support or relate the constant in the inequality to moments of the convolved measures, similar to the classical case.

References

1. Bai, Z.D.: Convergence rate of expected spectral distributions of large random matrices. Part I. Wigner matrices. Ann. Probab. **21**, 625–648 (1993)
2. Bercovici, H., Voiculescu, D.: Superconvergence to the central limit and failure of the Cramer theorem for free random variables. Probab. Theory Rel. Fields **102**, 215–222 (1995)
3. Berry, A.C.: The accuracy of the Gaussian approximation to the sum of independent variables. Trans. Am. Math. Soc. **49**, 122–136 (1941)
4. Esseen, C.G.: Fourier analysis of distribution functions. Acta Math. **77**, 1–125 (1945)
5. Hiai, F., Petz, D.: The Semicircle Law, Free Random Variables and Entropy, 1 edn. Mathematical Surveys and Monographs, vol. 77. American Mathematical Society, Providence (2000)
6. Maassen, H.: Addition of freely independent random variables. J. Funct. Anal. **106**, 409–438 (1992)
7. Markushevich, A.I.: Theory of Functions of a Complex Variable, 2 edn. Chelsea, New York (1977)
8. Pata, V.: The central limit theorem for free additive convolution. J. Funct. Anal. **140**, 359–380 (1996)
9. Voiculescu, D.: Symmetries of some reduced free product C^*-algebras. In: Lecture Notes in Mathematics, vol. 1132, pp. 556–588. Springer, New York (1983)
10. Voiculescu, D.: Addition of certain non-commuting random variables. J. Funct. Anal. **66**, 323–346 (1986)
11. Voiculescu, D.: Lectures on free probability theory. In: Bernard, P. (ed.) Lectures on Probability Theory and Statistics. Lecture Notes in Mathematics, vol. 1738, pp. 279–349. Springer, Berlin (1998)
12. Whittaker, E.T., Watson, G.N.: A Course of Modern Analysis, 4 edn. Cambridge University Press, Cambridge (1927)