The Annals of Probability
2013, Vol. 41, No. 5, 3241–3260
DOI: 10.1214/12-AOP756
© Institute of Mathematical Statistics, 2013

AN INEQUALITY FOR THE DISTANCE BETWEEN DENSITIES OF FREE CONVOLUTIONS

BY V. KARGIN

University of Cambridge

This paper contributes to the study of the free additive convolution of probability measures. It shows that under some conditions, if measures μ_i and ν_i, $i = 1, 2$, are close to each other in terms of the Lévy metric and if the free convolution $\mu_1 \boxplus \mu_2$ is sufficiently smooth, then $\nu_1 \boxplus \nu_2$ is absolutely continuous, and the densities of measures $\nu_1 \boxplus \nu_2$ and $\mu_1 \boxplus \mu_2$ are close to each other. In particular, convergence in distribution $\mu_1^{(n)} \to \mu_1$, $\mu_2^{(n)} \to \mu_2$ implies that the density of $\mu_1^{(n)} \boxplus \mu_2^{(n)}$ is defined for all sufficiently large n and converges to the density of $\mu_1 \boxplus \mu_2$. Some applications are provided, including: (i) a new proof of the local version of the free central limit theorem, and (ii) new local limit theorems for sums of free projections, for sums of $\boxplus$-stable random variables and for eigenvalues of a sum of two N-by-N random matrices.

1. Introduction. Free convolution is a binary operation on the set of probability measures on the real line that converts this set into a commutative semigroup. In contrast to the usual convolution, this operation is nonlinear relative to taking convex combinations of measures. The study of properties of free convolution is motivated by its numerous applications to operator algebras [11, 21, 24], random matrices [10, 17, 19, 22], representations of the symmetric group [8] and quantum physics [9, 27].

Starting with work by Voiculescu [21], it was noted that free convolution has strong smoothing properties. Let $\mu_1 \boxplus \mu_2$ denote the free convolution of probability measures μ_1 and μ_2. In [6], it was proved that $\mu_1 \boxplus \mu_2$ has an atom at $x \in \mathbb{R}$ if and only if there are $y \in \mathbb{R}$ and $z \in \mathbb{R}$ such that $x = y + z$, and $\mu_1(\{y\}) + \mu_2(\{z\}) > 1$. In [1], it was shown that $\mu_1 \boxplus \mu_2$ can have a singular component if and only if one of the measures is concentrated on one point, and the other has a singular component (so that the resulting free convolution is simply a translation of the measure with the singular component). Moreover, in the same paper it was shown that the density of the absolutely continuous part of the free convolution measure is analytic wherever the density is positive and finite.

Some quantitative versions of the smoothing property of free convolution have also been given. In particular, in [23] it was shown that if μ_1 is absolutely continuous with density $f_{\mu_1} \in L^p(\mathbb{R})$ ($p \in (1, \infty]$), then the free convolution of μ_1

Received August 2011; revised January 2012.
MSC2010 subject classifications. 46L54, 60B20.
Key words and phrases. Free probability, free convolution, convergence of measures.

with an arbitrary other measure μ_2 is absolutely continuous with density $f_{\mu_1 \boxplus \mu_2} \in L^p(\mathbb{R})$, and $\|f_{\mu_1 \boxplus \mu_2}\|_p \le \|f_{\mu_1}\|_p$. In particular, the supremum of the density $f_{\mu_1 \boxplus \mu_2}$ is less than or equal to the supremum of the density of f_{μ_1}.

Another important property of free convolution is that it is continuous with respect to weak convergence of measures. In particular, by a result in [4], if $\mu_1^{(N)} \to \mu_1$ and $\mu_2^{(N)} \to \mu_2$ as N grows to infinity (where $\to$ denotes convergence in distribution), then $\mu_1^{(N)} \boxplus \mu_2^{(N)} \to \mu_1 \boxplus \mu_2$. In fact, Theorem 4.13 in [4] says that $d_L(\mu_1 \boxplus \mu_2, \nu_1 \boxplus \nu_2) \le d_L(\mu_1, \nu_1) + d_L(\mu_2, \nu_2)$, where d_L denotes the Lévy distance on the set of probability measures on $\mathbb{R}$.

The main result of this paper establishes a strengthened version of this property. If distances $d_L(\mu_1, \nu_1)$ and $d_L(\mu_2, \nu_2)$ are sufficiently small, and if $\mu_1 \boxplus \mu_2$ is sufficiently smooth, then $\nu_1 \boxplus \nu_2$ is absolutely continuous and the distance between the densities of $\mu_1 \boxplus \mu_2$ and $\nu_1 \boxplus \nu_2$ can be bounded in terms of the Lévy distances between the original measures.

In particular, this result shows that the convergence in distribution $\mu_1^{(N)} \to \mu_1$ and $\mu_2^{(N)} \to \mu_2$ implies the convergence of the probability densities of $\mu_1^{(N)} \boxplus \mu_2^{(N)}$ to the density of $\mu_1 \boxplus \mu_2$.

We prove this result under an assumption imposed on the measures μ_1 and μ_2, which we call the smoothness of the pair (μ_1, μ_2) at a point of its support x. This assumption holds at a generic point x if $\mu_1 = \mu_2 = \mu$, and the density of $\mu \boxplus \mu$ is absolutely continuous and positive at x. In the case when $\mu_1 \neq \mu_2$, this assumption should be checked directly. We envision that in applications μ_1 and μ_2 are fixed measures for which this assumption can be directly checked, and $\mu_1^{(N)}$ and $\mu_2^{(N)}$ are (perhaps random) measures for which it can be checked that they are close to μ_1 and μ_2 in the Lévy distance.

In order to formulate our main result precisely, we introduce several definitions. Let μ_1 and μ_2 be two probability measures on $\mathbb{R}$ with the Stieltjes transforms $m_{\mu_1}(z)$ and $m_{\mu_2}(z)$, where the Stieltjes transform of a probability measure μ is defined by the formula

$$m_\mu(z) := \int_{\mathbb{R}} \frac{\mu(dx)}{x - z}.$$

Then, the free convolution $\mu_1 \boxplus \mu_2$ is defined as a probability measure on $\mathbb{R}$ with the Stieltjes transform $m_{\mu_1 \boxplus \mu_2}(z)$, which satisfies the following system of equations:

$$\begin{aligned} m_{\mu_1 \boxplus \mu_2}(z) &= m_{\mu_1}(\omega_1(z)), \\ m_{\mu_1 \boxplus \mu_2}(z) &= m_{\mu_2}(\omega_2(z)), \\ z - \frac{1}{m_{\mu_1 \boxplus \mu_2}(z)} &= \omega_1(z) + \omega_2(z). \end{aligned} \tag{1}$$

Here $\omega_1(z)$ and $\omega_2(z)$ are analytic functions in $\mathbb{C}^+ := \{z : \Im z > 0\}$, that map $\mathbb{C}^+$ to itself, that have the property $\Im \omega_j(z) \ge \Im z$, and such that $\omega_j(z) = z + o(z)$ as $z \to$

∞ in the sector $\Im z > \kappa|\Re z|$, where κ is an arbitrary positive constant [7]. Functions $\omega_1(z)$ and $\omega_2(z)$ are called the *subordination functions* for the pair (μ_1, μ_2).

The definition of free convolution by the system (1) is equivalent to the standard definition through R-transforms ([25] and [16]) if one sets $\omega_1(z) = z - R_{\mu_2}(-m_{\mu_1 \boxplus \mu_2}(z))$, and similarly for $\omega_2(z)$.

The subordination functions $\omega_j(z)$ depend not only on z but also on the pair (μ_1, μ_2). In particular, some properties of the measures μ_1 and μ_2 are encoded in the functions ω_j. A proper but more cumbersome notation would be $\omega_j(\mu_1, \mu_2, z)$ where $j = 1, 2$. In the cases when we need to compare the subordination functions for pairs (μ_1, μ_2) and (ν_1, ν_2), we will denote them by $\omega_{\mu,j}(z)$ and $\omega_{\nu,j}(z)$, respectively.

The system (1) implies the following system of equations for ω_j:

$$\begin{aligned}\frac{1}{z - \omega_1(z) - \omega_2(z)} &= m_{\mu_1}(\omega_1(z)),\\ \frac{1}{z - \omega_1(z) - \omega_2(z)} &= m_{\mu_2}(\omega_2(z)).\end{aligned} \tag{2}$$

Note that the analytic solutions of the system (2) that satisfy the asymptotic condition at infinity are unique in $\mathbb{C}^+$. (This follows from the facts that the solutions are unique in the area $\Im z \geq \eta_0$ for sufficiently large η_0 and that the analytic continuation in a simply-connected domain is unique.)

By Theorem 3.3 in [1], the limits $\omega_j(x) = \lim_{\eta \downarrow 0} \Im \omega_j(x + i\eta)$ exist, and we make the following definition.

DEFINITION 1.1. A pair of probability measures on the real line (μ_1, μ_2) is said to be *smooth* at x if the following two conditions hold:

(i) $\Im \omega_j(x) > 0$ for $j = 1, 2$, and
(ii)

$$k_\mu(x) := \frac{1}{m'_{\mu_1}(\omega_1(x))} + \frac{1}{m'_{\mu_2}(\omega_2(x))} - \big(x - \omega_1(x) - \omega_2(x)\big)^2 \neq 0. \tag{3}$$

Inequality (3) is a technical condition and holds for a generic point $x \in \mathbb{R}$.

Condition (i) is somewhat stronger than the condition that $\mu_1 \boxplus \mu_2$ is Lebesgue absolutely continuous at x. Indeed, if $\Im \omega_j(x) > 0$ for $j = 1, 2$, then the limit

$$\lim_{\eta \to 0} m_{\mu_1 \boxplus \mu_2}(x + i\eta) = \lim_{\eta \to 0} m_{\mu_1}(\omega_1(z))$$

exists and is finite. By using results in [1], we can infer from this fact that $\mu_1 \boxplus \mu_2$ is Lebesgue absolutely continuous at x.

In the converse direction, we have only that if $\mu_1 = \mu_2 = \mu$, and $\mu \boxplus \mu$ is absolutely continuous with positive density at x, then condition (i) in the definition of smoothness is satisfied; see Proposition 1.4 below.

The fact that smoothness is strictly stronger than absolute continuity of $\mu_1 \boxplus \mu_2$ can be seen from the following example. If μ_1 is a point mass at 0, that is, $\mu_1 = \delta_0$, and if μ_2 is absolutely continuous at x, then $\mu_1 \boxplus \mu_2 = \mu_2$ is absolutely continuous at x, but the pair (μ_1, μ_2) is not smooth at x. Indeed, $m_{\delta_0} = -z^{-1}$, and system (2) implies that $\omega_2(z) = z$. Hence, $\Im\omega_2(x) = 0$ for every x.

On the other hand smoothness holds for many examples that we consider below.

Next, let us recall the following standard definition.

DEFINITION 1.2. The *Lévy distance* between probability measures μ and ν is

$$d_L(\mu, \nu) = \sup_x \inf\{s \geq 0 : F_\nu(x - s) - s \leq F_\mu(x) \leq F_\nu(x + s) + s\},$$

where $F_\mu(t)$ and $F_\nu(t)$ are the cumulative distribution functions of μ and ν.

It is well known that $\mu^{(N)} \to \mu$ in distribution (i.e., the cumulative distribution function of $\mu^{(N)}$ weakly converges to the cumulative distribution function of μ) if and only if $d_L(\mu^{(N)}, \mu) \to 0$; see, for example, Theorem III.1.2 on page 314 and Exercise III.1.4 on page 316 in [18].

Here is the main result of this paper.

THEOREM 1.3. *Assume that a pair of probability measures (μ_1, μ_2) is smooth at x. Then, there are some $s_{\mu,0} >$ and $c_\mu > 0$, which depend only on (μ_1, μ_2), such that for all pairs of probability measures (ν_1, ν_2) with $d_L(\mu_j, \nu_j) < s \leq s_{\mu,0}$ for both $j = 1, 2$, it is true that $\nu_1 \boxplus \nu_2$ is absolutely continuous in a neighborhood of x, and*

$$\left|f_{\nu_1 \boxplus \nu_2}(x) - f_{\mu_1 \boxplus \mu_2}(x)\right| < c_\mu s,$$

where $f_{\nu_1 \boxplus \nu_2}$ and $f_{\mu_1 \boxplus \mu_2}$ are the densities of $\nu_1 \boxplus \nu_2$ and $\mu_1 \boxplus \mu_2$, respectively.

This theorem will be proved as a corollary to Proposition 2.4 below. The assumptions of the theorem are sufficient but possibly not necessary. Of course, it is necessary to require that $\mu_1 \boxplus \mu_2$ be absolutely continuous at x so that the density $f_{\mu_1 \boxplus \mu_2}(x)$ is well defined. In addition, a simple example shows that absolute continuity alone is not sufficient. Indeed, if $\mu_1 = \nu_1 = \delta_0$ is a point mass at zero, and μ_2 is absolutely continuous, then $\delta_0 \boxplus \mu_2$ is absolutely continuous, but $\delta_0 \boxplus \nu_2$ is not necessarily so, even if ν_2 is close to μ_2 in the Lévy distance. However, it is not clear if the assumption of absolute continuity of $\mu_1 \boxplus \mu_2$ implies the statement of the theorem once this degenerate case is ruled out.

The constant c_μ in the theorem can be bounded in terms of $\Im\omega_{\mu,j}(x)$ and $|k_\mu(x)|$ from (3). In particular, if $\Im\omega_{\mu,j}(x)$ and $|k_\mu(x)|$ are uniformly bounded away from zero for all $x \in (a, b)$, then $\sup_{x \in (a,b)} |f_{\nu_1 \boxplus \nu_2}(x) - f_{\mu_1 \boxplus \mu_2}(x)| < cs$ for some $c > 0$.

The main ideas of the proof of Theorem 1.3 are as follows. Let $m_{\nu_j}(z)$ and $m_{\nu_1 \boxplus \nu_2}(z)$ denote the Stieltjes transforms of ν_j and $\nu_1 \boxplus \nu_2$, respectively, and let

$\omega_{\nu,j}$ denote the subordination functions for the pair (ν_1, ν_2). First, we prove that the smallness of $d_L(\mu_j, \nu_j)$ implies that the differences $|m_{\nu_j} - m_{\mu_j}|$ are small, and that the differences between the derivatives of m_{ν_j} and m_{μ_j} are also small. Then we show that this fact, together with system (2), implies that the differences between the corresponding subordination functions are small. At this stage we need the assumption of smoothness. Finally, we check that if both the Stieltjes transforms and the subordination functions of pairs (μ_1, μ_2) and (ν_1, ν_2) are close to each other, then the Stieltjes transforms of $\mu_1 \boxplus \mu_2$ and $\nu_1 \boxplus \nu_2$ are close to each other uniformly on the half-line $\Re z = x$, $\Im z > 0$. This fact implies that the densities of $\mu_1 \boxplus \mu_2$ and $\nu_1 \boxplus \nu_2$ at x are close to each other.

Before discussing applications of Theorem 1.3, let us mention some results which are helpful in checking the assumptions of this theorem.

PROPOSITION 1.4. *If $\mu \boxplus \mu$ is (Lebesgue) absolutely continuous in a neighborhood of x, and the density of $\mu \boxplus \mu$ is positive at x, then $\Im\omega_j(x) > 0$ for $j = 1, 2$.*

Another important case is when one of the probability measures has the semicircle distribution with the density $f_{\text{sc}}(x) = \frac{1}{2\pi}\sqrt{(4 - x^2)_+}$. Since such a measure, μ_{sc}, is absolutely continuous, $\mu_{\text{sc}} \boxplus \mu$ is also absolutely continuous, for an arbitrary μ.

PROPOSITION 1.5. *If the density of $\mu_{\text{sc}} \boxplus \mu$ is positive at x, and*

$$\left|m_{\mu_{\text{sc}} \boxplus \mu}(x)\right| \neq 1,$$

then $\Im\omega_j(x) > 0$ for $j = 1, 2$.

The proofs of Propositions 1.4 and 1.5 will be given in Section 3.

Now let us turn to applications. Theorem 1.3 can be applied to derive some old and new results about sums of free random variables and about eigenvalues of large random matrices.

Recall that if $X_1, \ldots, X_n$ are free, identically distributed self-adjoint random variables with finite variance σ^2, then [15, 20] $S_n := (X_1 + \cdots + X_n)/(\sigma\sqrt{n})$ converges in distribution to a random variable X with the standard semicircle law.

In terms of free convolutions, it means that if μ is a probability measure with variance σ^2, and if

$$\mu_n(dx) := \underbrace{\mu \boxplus \cdots \boxplus \mu}_{n \text{ times}}(\sigma\sqrt{n}\,dx),$$

then $\mu_n \to \mu_{\text{sc}}$.

Bercovici and Voiculescu in [5] showed that the convergence in this limit law holds in a stronger sense. Namely, assuming in addition that support of μ is bounded, they showed that μ_n has a density for all sufficiently large n and that the sequence of these densities converges uniformly to the density of the semicircle law. Recently, this result was generalized in [26] to the case of μ_n with unbounded support and finite variance. Results in [5] and [26] can be considered as local limit versions of the free CLT.

In the first application (Theorem 4.1), we give a short proof of the easier part of the results in [5] and [26] by using Theorem 1.3. (A more difficult part of these results concerns the uniformity of the convergence on $\mathbb{R}$.)

In the second application (Theorem 4.2), we prove an analogous local limit result for the sums $S_n = X_{1,n} + \cdots + X_{n,n}$, where $X_{i,n}$ are free projection operators with parameters $p_{i,n}$ such that $\sum_{i=1}^{n} p_{i,n} \to \lambda$ and $\max_i p_{i,n} \to 0$ as $n \to \infty$. The classical analogue of this situation is the sum of independent indicator random variables, and the classical result states that the sums converge in distribution to the Poisson random variable with parameter λ. A local version of this result is absent in the classical case because the Poisson random variable is discrete, and it does not make sense to talk about convergence of densities. In the free probability case, the limit of the spectral distributions of S_n is the Marchenko–Pastur distribution, which is absolutely continuous with bounded density for $\lambda > 1$. We show that in this case the spectral measures of S_n have a density for all sufficiently large n and that the sequence of these densities converges uniformly to the density of the Marchenko–Pastur law.

In the third application (Theorem 4.3), we show that a similar local limit result holds for sums of free $\boxplus$-stable random variables.

The fourth application (Theorem 4.4) is of a different kind and is concerned with eigenvalues of large random matrices. Let $H_N = A_N + U_N B_N U_N^*$, where A_N and B_N are N-by-N Hermitian matrices, and U_N is a random unitary matrix with the Haar distribution on the unitary group $\mathcal{U}(N)$. Let $\lambda_1^{(A)} \geq \cdots \geq \lambda_N^{(A)}$ be the eigenvalues of A_N. Similarly, let $\lambda_k^{(B)}$ and $\lambda_k^{(H)}$ be ordered eigenvalues of matrices B_N and H_N, respectively. Define the *spectral point measures* of A_N by $\mu_{A_N} := N^{-1} \sum_{k=1}^{N} \delta_{\lambda_k^{(A)}(H)}$, and define the spectral point measures of B_N and H_N similarly.

Assume that $\mu_{A_N} \to \mu_\alpha$ and $\mu_{B_N} \to \mu_\beta$, and that the support of μ_{A_N} and μ_{B_N} is uniformly bounded. Let the pair (μ_α, μ_β) be smooth at x.

Define $\mathcal{N}_I := N\mu_{H_N}(I)$, the number of eigenvalues of H_N in interval I, and let $\mathcal{N}_\eta(x) := \mathcal{N}_{(x-\eta,x+\eta]}$. Finally, assume that $\eta = \eta(N)$ and $\frac{1}{\sqrt{\log(N)}} \ll \eta(N) \ll 1$.

Then, by using the author's previous results from [14], and Theorem 1.3, it is shown that

$$\frac{\mathcal{N}_\eta(x)}{\eta N} \to f_{\mu_\alpha \boxplus \mu_\beta}(x)$$

with probability 1, where $f_{\mu_\alpha \boxplus \mu_\beta}$ denotes the density of $\mu_\alpha \boxplus \mu_\beta$. This result generalizes the main result in [17] where it was proved that $\mu_{H_N} \to \mu_\alpha \boxplus \mu_\beta$. It can be interpreted as a local limit law for eigenvalues of a sum of random Hermitian matrices.

The rest of the paper is organized as follows. Section 2 is concerned with the proof of the main theorem, Section 3 contains proofs of Propositions 1.4 and 1.5, Section 4 contains applications, and Section 5 concludes.

2. Proof of Theorem 1.3.

Let $F_\mu(x)$ and $F_\nu(x)$ denote the cumulative distribution functions of the measures μ and ν, respectively.

LEMMA 2.1. *Suppose that $d_L(\mu,\nu) = s$. Assume that $h(x)$ is a C^1 real-valued function, such that $\int_{-\infty}^{\infty} |h(u)|\,du < \infty$ and $\int_{-\infty}^{\infty} |h'(u)|\,du < \infty$. Assume in addition that $h(u)$ has a finite number of zeros. Then,*

$$\Delta := \int_{\mathbb{R}} \big|h(u)[F_\nu(\eta u) - F_\mu(\eta u)]\big|\,du \le cs \max\{1, \eta^{-1}\}, \tag{4}$$

where $c > 0$ depends only on h.

PROOF. Since h is a continuous function with a finite number of zeros, we can decompose the set on which $h(u)$ is nonzero into a finite number of intervals I_k on which $h(u)$ has a constant sign. Note that it suffices to estimate the integral on each of these intervals. Consider the case when $h(u) > 0$ on an interval I_k. The treatment of the case $h(u) < 0$ is similar.

By using the definition of the Lévy distance, we obtain the following estimate:

$$\begin{aligned}
&\big|F_\nu(\eta u) - F_\mu(\eta u)\big| \\
&\quad \le \max\{F_\mu(\eta u + s) - F_\mu(\eta u), F_\nu(\eta u + s) - F_\nu(\eta u), \\
&\qquad\qquad F_\mu(\eta u) - F_\mu(\eta u - s), F_\nu(\eta u) - F_\nu(\eta u - s)\} + s.
\end{aligned}$$

It suffices to estimate

$$\int_{I_k} h(u)\{F_\mu(\eta u + s) - F_\mu(\eta u) + s\}\,du,$$

since the other cases are similar.

First of all, note that

$$\int_{I_k} h(u) s\,du \le s \int_{-\infty}^{\infty} \big|h(u)\big|\,du \le cs. \tag{5}$$

Next, let $\widetilde{I}_k = I_k + s/\eta$. Then,

$$\int_{I_k} h(u) F_\mu(\eta u + s)\,du = \int_{\widetilde{I}_k} h(t - s/\eta) F_\mu(\eta t)\,dt$$

and therefore,

$$\begin{aligned}(6)\qquad &\int_{I_k} h(u)\big[F_\mu(\eta u+s)-F_\mu(\eta u)\big]\,du\\ &\quad\le \int_{I_k\cap\widetilde{I}_k}\big[h(t-s/\eta)-h(t)\big]F_\mu(\eta t)\,dt\\ &\qquad+\int_{I_k\triangle\widetilde{I}_k}\max\big(|h(t-s/\eta)|,|h(t)|\big)F_\mu(\eta t)\,dt.\end{aligned}$$

For the first integral in this estimate, we can use the fact that

$$h(t-s/\eta)-h(t)=-\int_{t-s/\eta}^{t}h'(\xi)\,d\xi$$

and therefore,

$$\begin{aligned}\left|\int_{I_k\cap\widetilde{I}_k}\big[h(t-s/\eta)-h(t)\big]F_\mu(\eta t)\,dt\right| &\le \int_{\mathbb{R}}\int_{t-s/\eta}^{t}|h'(\xi)|F_\mu(\eta t)\,d\xi\,dt\\ (7)\qquad &=\int_{\mathbb{R}}|h'(\xi)|\left(\int_{\xi}^{\xi+s/\eta}F_\mu(\eta t)\,dt\right)d\xi\\ &\le \frac{s}{\eta}\int_{\mathbb{R}}|h'(\xi)|\,d\xi.\end{aligned}$$

For the second integral, we note that

$$\begin{aligned}\int_{I_k\triangle\widetilde{I}_k}\max\big(|h(t-s/\eta)|,|h(t)|\big)F_\mu(\eta t)\,dt &\le \sup|h(t)|\,|I_k\triangle\widetilde{I}_k|\\ (8)\qquad &\le 2\sup|h(t)|s/\eta.\end{aligned}$$

By using estimates (5), (6), (7) and (8), we obtain

$$\Delta\le cs\max\{1,\eta^{-1}\},$$

where c depends only on function h. □

Now, let $m_\mu(z)$ and $m_\nu(z)$ denote the Stieltjes transforms of the probability measures μ and ν, respectively.

LEMMA 2.2. *Let $d_L(\mu,\nu)=s$ and $z=x+i\eta$, where $\eta>0$. Then*:

(a) $|m_\mu(z)-m_\nu(z)|<cs\eta^{-1}\max\{1,\eta^{-1}\}$ *where c is a positive constant*, *and*
(b) $|\frac{d^r}{dz^r}(m_\mu(z)-m_\nu(z))|<c_r s\eta^{-1-r}\max\{1,\eta^{-1}\}$ *where c_r are positive constants*.

PROOF. (a) By integration by parts,

$$m_\mu(z)=\int_{\mathbb{R}}\frac{F_\mu(\lambda)}{(\lambda-z)^2}\,d\lambda.$$

Hence, setting $u = (\lambda - x)/\eta$,

$$\Im m_\mu(z) = \frac{2}{\eta} \int_{\mathbb{R}} F_\mu(x + \eta u) \frac{u\, du}{(1 + u^2)^2},$$

$$\Re m_\mu(z) = \frac{1}{\eta} \int_{\mathbb{R}} F_\mu(x + \eta u) \frac{(u^2 - 1)\, du}{(1 + u^2)^2},$$

and similar formulas hold for $\Im m_\nu(z)$ and $\Re m_\nu(z)$. Since $u(1 + u^2)^{-2}$ and $(u^2 - 1)(1 + u^2)^{-2}$ satisfy the assumptions of Lemma 2.1, Claim (a) follows. Claim (b) can be derived similarly by writing

$$\begin{aligned} \frac{d^r}{dz^r} m_\mu(z) &= (r + 1)! \int_{\mathbb{R}} \frac{F_\mu(\lambda)\, d\lambda}{(\lambda - x - i\eta)^{r+2}} \\ &= \frac{(r + 1)!}{\eta^{r+1}} \int_{\mathbb{R}} \frac{1}{(u - i)^{r+2}} F_\mu(\eta u + x)\, du, \end{aligned}$$

separating imaginary and real parts of the integrand, and applying Lemma 2.1. □

LEMMA 2.3. *Assume that the pair (μ_1, μ_2) is smooth at x. Suppose that (ν_1, ν_2) is another pair of probability measures such that $d_L(\mu_j, \nu_j) < s$ for $j = 1, 2$. Let $\Re z = x$ and $\Im z \geq 0$. Then*

$$\left| \frac{1}{z - \omega_{\mu,1}(z) - \omega_{\mu,2}(z)} - m_{\nu_j}\big(\omega_{\mu,j}(z)\big) \right| \leq c_\mu s$$

for $j = 1, 2$. Here $c_\mu > 0$ depends only on (μ_1, μ_2) and x.

That is, if we substitute $\omega_{\mu,j}(z)$ in the system for $\omega_{\nu,j}(z)$, then the equalities will be satisfied up to a quantity of order s.

PROOF OF LEMMA 2.3. The functions $\omega_{\mu,j}(z)$ satisfy equations (2), which implies that it is enough to show that

$$\big| m_{\nu_j}\big(\omega_{\mu,j}(z)\big) - m_{\mu_j}\big(\omega_{\mu,j}(z)\big) \big| < cs$$

for $j = 1, 2$. Note that $\min_{j=1,2}\{\Im(\omega_{\mu,j}(x))\} > 0$ by the assumption of smoothness of (μ_1, μ_2). In addition, $\Im(\omega_{\mu,j}(x + i\eta)) \geq \eta$ for all $\eta > 0$. Hence, by continuity of $\omega_{\mu,j}(x + i\eta)$ in η, we have $\kappa_j := \inf_{\eta \geq 0} \omega_{\mu,j}(x + i\eta) > 0$. Then, by Lemma 2.2,

$$\big| m_{\nu j}\big(\omega_{\mu,j}(z)\big) - m_{\mu j}\big(\omega_{\mu,j}(z)\big) \big| < cs \min\{\kappa_j^{-1}, \kappa_j^{-2}\}. \qquad \square$$

PROPOSITION 2.4. *Assume that a pair of probability measures (μ_1, μ_2) is smooth at x. Then there are some $s_{\mu,0} >$ and $c_\mu > 0$ that depend only on (μ_1, μ_2) and x, such that for all pairs of probability measures (ν_1, ν_2) with $d_L(\mu_j, \nu_j) < s \leq s_{\mu,0}$ for $j = 1, 2$, the limits $\omega_{\nu,j}(x) := \lim_{\eta \downarrow 0} \omega_{\nu,j}(x + i\eta)$ exist, and it is true that*

$$\big| \omega_{\nu,j}(x) - \omega_{\mu,j}(x) \big| \leq c_\mu s$$

for $j = 1, 2$.

COROLLARY 2.5. *Assume that the assumptions of Proposition 2.4 hold and that $d_L(\mu_j, \nu_j) < s \le s_{\mu,0}$ for $j = 1, 2$. Then, $\nu_1 \boxplus \nu_2$ is absolutely continuous in a neighborhood of x, and*

$$\left|f_{\nu_1\boxplus\nu_2}(x) - f_{\mu_1\boxplus\mu_2}(x)\right| < c_\mu s,$$

where $f_{\mu_1\boxplus\mu_2}$ and $f_{\nu_1\boxplus\nu_2}$ denote the densities of $\mu_1 \boxplus \mu_2$ and $\nu_1 \boxplus \nu_2$, respectively.

PROOF. Since $m_{\nu_1\boxplus\nu_2}(z) = (z - \omega_{\nu,1}(z) - \omega_{\nu,2}(z))^{-1}$, Proposition 2.4 implies that the limit $m_{\nu_1\boxplus\nu_2}(x) := \lim_{\eta\downarrow 0} m_{\nu_1\boxplus\nu_2}(x + i\eta)$ exists and

$$\left|m_{\nu_1\boxplus\nu_2}(x) - m_{\mu_1\boxplus\mu_2}(x)\right| < c_\mu s. \tag{9}$$

By [1], $\nu_1 \boxplus \nu_2$ has no singular component. Hence, inequality (9) and the absolute continuity of $\mu_1 \boxplus \mu_2$ in a neighborhood of x imply that for all sufficiently small s, the measure $\nu_1 \boxplus \nu_2$ is absolutely continuous in a neighborhood of x with the density $f_{\nu_1\boxplus\nu_2}(x) = \pi^{-1}\Im(m_{\nu_1\boxplus\nu_2}(x))$, and

$$\left|f_{\nu_1\boxplus\nu_2}(x) - f_{\mu_1\boxplus\mu_2}(x)\right| < c_\mu s. \qquad \square$$

PROOF OF PROPOSITION 2.4. Let $F(\omega) : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by the formula

$$F : \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \to \begin{pmatrix} c(z - \omega_1 - \omega_2)^{-1} - m_{\nu_1}(\omega_1) \\ (z - \omega_1 - \omega_2)^{-1} - m_{\nu_2}(\omega_2) \end{pmatrix}.$$

Let us use the norm $\|(x_1, x_2)\| = (|x_1|^2 + |x_2|^2)^{1/2}$. By Lemma 2.3, $\|F(\omega_{\mu,1}(z), \omega_{\mu,2}(z))\| \le c_\mu s$ for all $z = x + i\eta$ and $\eta \ge 0$.

The derivative of F with respect to ω is

$$F' = \begin{pmatrix} (z - \omega_1 - \omega_2)^{-2} & (z - \omega_1 - \omega_2)^{-2} - m'_{\nu_1}(\omega_1) \\ (z - \omega_1 - \omega_2)^{-2} - m'_{\nu_2}(\omega_2) & (z - \omega_1 - \omega_2)^{-2} \end{pmatrix}.$$

The determinant of this matrix is

$$\left[m'_{\nu_1}(\omega_1) + m'_{\nu_2}(\omega_2)\right](z - \omega_1 - \omega_2)^{-2} - m'_{\nu_1}(\omega_1)m'_{\nu_2}(\omega_2).$$

By the assumption of smoothness and by Lemma 2.2, this is close (i.e., the difference $< cs$ for some $c > 0$) to

$$\left[m'_{\mu_1}(\omega_1) + m'_{\mu_2}(\omega_2)\right](z - \omega_1 - \omega_2)^{-2} - m'_{\mu_1}(\omega_1)m'_{\mu_2}(\omega_2)$$

at $(\omega_1, \omega_2) = (\omega_{\mu,1}(z), \omega_{\mu,2}(z))$ for all $z = x + i\eta$ with $\eta \ge 0$. The latter expression is nonzero by (3). In addition, the assumption of smoothness shows that $(z - \omega_{\mu,1}(z) - \omega_{\mu,2}(z))^{-2}$ is bounded for $z = x + i\eta$ with $\eta \ge 0$. Hence, the entries of the matrix $[F']^{-1}$ are bounded at $(\omega_{\mu,1}(z), \omega_{\mu,2}(z))$, and the bound does not depend on η. This shows that the operator norm of $[F']^{-1}$ is bounded uniformly in η.

Similarly, an explicit calculation of F'', the assumption of smoothness of (μ_1, μ_2) and Lemma 2.2 imply that for all $z = x + i\eta$ with $\eta \ge 0$, the operator norm of F'' is bounded (uniformly in η) for all (ω_1, ω_2) in a neighborhood of $(\omega_{\mu,1}(z), \omega_{\mu,2}(z))$.

It follows by the Newton–Kantorovich theorem [13] that if $s = \max_j d_L(\mu_j, \nu_j)$ is sufficiently small, then the solution of the equation $F(\omega) = 0$ exists for all z with $\Re z = x$ and $\Im z \geq 0$.

This solution must be $(\omega_{\nu,1}(z), \omega_{\nu,2}(z))$ by the following argument from [2]. A solution of equation $F(\omega) = 0$ satisfies the following pair of equations:

$$\omega_1 = z + h_2(\omega_2),$$
$$\omega_2 = z + h_1(\omega_1),$$

where

$$h_j(\omega) = -\omega - \frac{1}{m_{\nu_j}(\omega)}.$$

Note in particular that $\Im h_j(\omega) \geq 0$ for all $\omega \in \mathbb{C}^+$; see, for example, [4] or [15].

Hence, ω_1 is a fixed point of the function

$$f_z(\omega) = z + h_2\big(z + h_1(\omega)\big),$$

which maps $\mathbb{C}^+$ to $\mathbb{C}^+$. For every $z \in \mathbb{C}^+$, the function $f_z(\omega)$ is not a conformal automorphism because it maps $\mathbb{C}^+$ to a subset of $\mathbb{C}^+ + \Im z$, which is a proper subset of $\mathbb{C}^+$. In addition, it is analytic as a function of z and ω that maps $\mathbb{C}^+ \times \mathbb{C}^+$ to $\mathbb{C}^+$. Hence, by Theorem 2.4 in [2], for every $z \in \mathbb{C}^+$ the function $f_z(\omega)$ has a unique fixed point $\omega_1(z)$.

A similar argument holds for $\omega_2(z)$, and we conclude that equation $F(\omega) = 0$ has a unique solution in $\mathbb{C}^+ \times \mathbb{C}^+$, which necessarily coincides with $(\omega_{\nu,1}(z), \omega_{\nu,2}(z))$.

In addition, this solution satisfies the inequalities

$$|\omega_{\nu,j}(z) - \omega_{\mu,j}(z)| < c_\mu s, \qquad j = 1, 2, \tag{10}$$

for all z with $\Re z = E$ and $\Im z > 0$.

By Theorem 3.3 in [1], the limits

$$\omega_{\nu,j}(E) := \lim_{\eta \downarrow 0} \omega_{\nu,j}(x + i\eta)$$

and

$$\omega_{\mu,j}(E) := \lim_{\eta \downarrow 0} \omega_{\mu,j}(x + i\eta)$$

exist, and by taking the limits in (10), we find that

$$|\omega_{\nu,j}(x) - \omega_{\mu,j}(x)| \leq cs. \qquad \square$$

3. Proofs of Propositions 1.4 and 1.5. Recall that a function $f(x)$ is said to be Hölder continuous at x_0 if there exist positive constants α, C and ε such that $|x - x_0| < \varepsilon$ implies that $|f(x) - f(x_0)| < C|x - x_0|^{\alpha}$.

LEMMA 3.1. *Suppose that a probability measure μ has a density which is positive and Hölder continuous at x. Let $m_\mu(z)$ be the Stieltjes transform of μ. Then $|m_\mu(x + i\eta)| \le M < \infty$ for all $\eta > 0$.*

PROOF. The results of Sokhotskyi, Plemelj and Privalov ensure that the limit of $m_\mu(x + i\eta)$ exists when $\eta \downarrow 0$; see Theorems 14.1b and 14.1c in [12]. In particular this implies that $m_\mu(x + i\eta)$ is bounded for sufficiently small η. In addition, $|m_\mu(x + i\eta)| \le 1/\eta$ so it is bounded for large η. Since $m_\mu(x + i\eta)$ is continuous in the upper half-plane, $m_\mu(x + i\eta)$ is bounded for all η, and the claim of the lemma follows. □

PROOF OF PROPOSITION 1.4. Note that for the case $\mu_1 = \mu_2 = \mu$,

$$\omega_1(z) = \omega_2(z) = \bigl(z - m_{\mu\boxplus\mu}(z)^{-1}\bigr)/2. \tag{11}$$

Since by assumption $\mu \boxplus \mu$ is absolutely continuous in a neighborhood of x, and its density $f_{\mu\boxplus\mu}$ is positive at x, by the results in [1] $f_{\mu\boxplus\mu}$ is analytic and therefore uniformly Hölder continuous in a neighborhood of x. By Sokhotskyi, Plemelj and Privalov's results, the limit $m_{\mu\boxplus\mu}(x) = \lim_{\eta\downarrow 0} m_{\mu\boxplus\mu}(x + i\eta)$ exists and $\Im m_{\mu\boxplus\mu}(x) = \pi f_{\mu\boxplus\mu}(x) > 0$. Then it follows from (11) that the limits $\omega_j(x) = \lim_{\eta\downarrow 0} \omega_j(x + i\eta)$ exist. Moreover, since

$$\Im\omega_j(z) = \frac{1}{2}\left(\eta + \frac{\Im m_{\mu\boxplus\mu}(z)}{|m_{\mu\boxplus\mu}(z)|^2}\right)$$

and by Lemma 3.1, $|m_{\mu\boxplus\mu}(z)|^2$ is bounded uniformly in η, hence the fact that $\Im m_{\mu\boxplus\mu}(x) = \pi f_{\mu\boxplus\mu}(x) > 0$ implies that $\Im\omega_j(x) > 0$. This completes the proof of the proposition. □

LEMMA 3.2. *If μ_1 has the semicircle distribution, then*:

(i) $\omega_1(z) = z - \omega_2(z) + [z - \omega_2(z)]^{-1}$;
(ii) $m_{\mu_{\mathrm{sc}}\boxplus\mu}(z) = \omega_2(z) - z$;
(iii) $\omega_2(z)$ *satisfies the equation*

$$\omega_2(z) = z + \int \frac{\mu(dx)}{x - \omega_2(z)}.$$

PROOF. (i) If μ_1 has the semicircle distribution, then $m_{\mu_1}^{(-1)} = -(z + z^{-1})$; hence the first equation in system (2) implies

$$\omega_1 = -\left(\frac{1}{z - \omega_1 - \omega_2} + z - \omega_1 - \omega_2\right),$$

which simplifies to

$$\omega_1 = z - \omega_2 + \frac{1}{z - \omega_2}.$$

(ii) By using (i),

$$m_{\mu_{\mathrm{sc}} \boxplus \mu} = \frac{1}{z - \omega_1 - \omega_2} = -(z - \omega_2).$$

(iii) The second equation in system (2) becomes

$$-\big(z - \omega_2(z)\big) = \int \frac{\mu(dx)}{x - \omega_2(z)}.$$

□

PROOF OF PROPOSITION 1.5. From (ii) in Lemma 3.2,

$$\Im \omega_2(x) = \Im m_{\mu_{\mathrm{sc}} \boxplus \mu}(x) = \pi f_{\mu_{\mathrm{sc}} \boxplus \mu}(x) > 0.$$

From (i),

$$\begin{aligned} \Im \omega_1(x) &= \Im \omega_2(x)\left(-1 + \frac{1}{|x - \omega_2|^2}\right) \\ &= \Im \omega_2(x)\left(-1 + \frac{1}{|m_{\mu_{\mathrm{sc}} \boxplus \mu}(x)|^2}\right). \end{aligned}$$

Since $\Im \omega_2(x) > 0$, if $|m_{\mu_{\mathrm{sc}} \boxplus \mu}(x)|^2 < 1$, then $\Im \omega_1(x) > 0$, and we are done. Two remaining possibilities are $|m_{\mu_{\mathrm{sc}} \boxplus \mu}(x)|^2 = 1$ and $|m_{\mu_{\mathrm{sc}} \boxplus \mu}(x)|^2 > 1$. However, $|m_{\mu_{\mathrm{sc}} \boxplus \mu}(x)|^2 > 1$ is in fact not possible because this would imply that $\Im \omega_1(x) < 0$, which is ruled out by a general result of Biane. To sum up, the assumptions $f_{\mu_{\mathrm{sc}} \boxplus \mu}(x) > 0$ and $|m_{\mu_{\mathrm{sc}} \boxplus \mu}(x)|^2 \neq 1$ imply that $\Im \omega_j(x) > 0$. □

4. Applications. In the first application we re-prove an easier part of the free local limit theorem which was first demonstrated in [5] for bounded random variables and later generalized in [26] to the case of unbounded variables with finite variance. We will show the convergence of densities, but we will not investigate whether the convergence is uniform on $\mathbb{R}$.

Let X_i be a sequence of self-adjoint identically-distributed free random variables in the sense of free probability theory. Define $S_n = (X_1 + \cdots + X_n)/\sqrt{n}$, and let μ and μ_n denote the spectral probability measures of X_i and S_n, respectively. It is known that

$$\mu_n(dx) = \underbrace{\mu \boxplus \cdots \boxplus \mu}_{n \text{ times}}\big(\sqrt{n}\, dx\big).$$

THEOREM 4.1. *Suppose μ has zero mean and unit variance. Let $I_\varepsilon = [-2 + \varepsilon, 2 - \varepsilon]$. Then for all sufficiently large n, μ_n is (Lebesgue) absolutely continuous everywhere on I, and the density $d\mu_n/dx$ uniformly converges on I_ε to the density of the standard semicircle law.*

Note that the results in [5] imply that for every closed interval J outside of $[2, -2]$, the measure $\mu_n(J) = 0$ for all sufficiently large n, provided that μ_1 has bounded support. In addition, the uniform convergence on I_ε can be strengthened to the uniform convergence on $\mathbb{R}$ as in the proof of Theorem 3.4(iii) in [26].

PROOF OF THEOREM 4.1. Let $\nu_{1,n}$ be the distribution of $(X_1 + \cdots + X_{[n/2]})/\sqrt{n}$ and $\nu_{2,n}$ be the distribution of $(X_{[n/2]+1} + \cdots + X_n)/\sqrt{n}$. By using the free CLT (Central limit theorem) from [15] (which generalizes the free CLT in [20]), we infer that both $\nu_{1,n}$ and $\nu_{2,n}$ converge weakly to $\widetilde{\mu}_{\mathrm{sc}}$, where $\widetilde{\mu}_{\mathrm{sc}}$ is the semicircle law with variance $1/2$. It is easy to calculate for the pair $(\widetilde{\mu}_{\mathrm{sc}}, \widetilde{\mu}_{\mathrm{sc}})$ that

$$\omega_{\widetilde{\mu},1} = \omega_{\widetilde{\mu},2} = \frac{3z + \sqrt{z^2 - 4}}{4}$$

and therefore $\Im \omega_{\widetilde{\mu},j}(x) > 0$ on I_ε. (This also follows from Proposition 1.4.) A calculation shows that the genericity condition (3) is satisfied for each $x \in I_\varepsilon$, and therefore the density of $\nu_{1,n} \boxplus \nu_{2,n}$ exists for all sufficiently large n, and converges to the density of $\widetilde{\mu}_{\mathrm{sc}} \boxplus \widetilde{\mu}_{\mathrm{sc}}$ at each $x \in I_\varepsilon$. A remark after Theorem 1.3 shows that the convergence is in fact uniform on I_ε. Since $\nu_{1,n} \boxplus \nu_{2,n} = \mu_n$, this implies that the density of μ_n converges uniformly on I_ε to the density of the standard semicircle law. □

In a similar fashion, it is possible to prove the local limit law for the convergence to the free Poisson distribution.

Let $\{X_{n,i}\}_{i=1}^n$ be freely independent self-adjoint random variables with the distribution $\mu_{n,i} = p_{n,i}\delta_1 + (1 - p_{n,i})\delta_0$. Let $S_n = X_{n,1} + \cdots + X_{n,n}$, and let μ_n denote the spectral probability measure of S_n. Then

$$\mu_n(dx) = \mu_{n,1} \boxplus \cdots \boxplus \mu_{n,n}(dx).$$

Recall that the *Marchenko–Pastur distribution* with parameter $\lambda \geq 1$ is a probability measure μ_{mp} on $\mathbb{R}$, with the density

$$f_{\mathrm{mp}}(x) = \frac{\sqrt{4x - (1 - \lambda + x)^2}}{2\pi x}$$

supported on the interval $[x_{\min}, x_{\max}] := [(1 - \sqrt{\lambda})^2, (1 + \sqrt{\lambda})^2]$. In the free probability literature, this distribution is called the *free Poisson distribution*.

THEOREM 4.2. *Assume that* $\sum_{i=1}^n p_{n,i} \to \lambda > 1$ *and* $\max_i p_{n,i} \to 0$ *as* $n \to \infty$. *Let* $I_\varepsilon = [x_{\min} + \varepsilon, x_{\max} - \varepsilon]$. *Then for all sufficiently large* n, μ_n *is* (*Lebesgue*) *absolutely continuous everywhere on* I_ε, *and the density* $d\mu_n/dx$ *uniformly converges on* I_ε *to the density of the Marchenko–Pastur law with parameter* λ.

The proof of this theorem is similar to the proof of the previous one. The first step is the weak convergence of μ_n. In the case when $p_{n,i} = \lambda/n$ for all i, a proof of weak convergence can be found on page 34 in [25]. The general case is a minor adaptation of this case, and we omit it. Next, we choose k_n so that

$$\sum_{i=1}^{k_n} p_{n,i} \le \lambda/2 < \sum_{i=1}^{k_n+1} p_{n,i}$$

and define $\nu_{1,n}$ and $\nu_{2,n}$ as the spectral probability measures of $X_{n,1} + \cdots + X_{n,k_n}$ and $X_{n,k_n+1} + \cdots + X_{n,n}$, respectively. It is easy to see that both $\nu_{1,n}$ and $\nu_{2,n}$ converge weakly to $\widetilde{\mu}_{\mathrm{mp}}$, the Marchenko–Pastur distribution with parameter $\lambda/2$. By using Proposition 1.4, we conclude that $\Im\omega_{\widetilde{\mu}_{\mathrm{mp}},j}(x) > 0$ on I_ε. Moreover, a direct calculation shows that

$$\omega_{\widetilde{\mu},1}(z) = \omega_{\widetilde{\mu},2}(z) = \tfrac{1}{4}\Big(z + \lambda - 1 + \sqrt{\big(z - (1+\lambda)\big)^2 - 4\lambda}\Big)$$

and

$$m'_{\widetilde{\mu}_{\mathrm{mp}}} = \frac{1-\lambda/2}{2z^2} + \frac{-z(1+\lambda/2) + (1-\lambda/2)^2}{2z^2\sqrt{(z-(1+\lambda/2))^2 - 2\lambda}}.$$

After some calculations the genericity condition (3) can be simplified to the following inequality:

$$\begin{aligned} f(x,\lambda) := {}& x^3 - \big(5 + \tfrac{5}{2}\lambda\big)x^2 + \big(7 + \tfrac{13}{2}\lambda + 2\lambda^2\big)x \\ & - \big(3 - 5\lambda + \tfrac{5}{4}\lambda^2 + \tfrac{1}{2}\lambda^3\big) \\ \neq {}& 0. \end{aligned} \tag{12}$$

Figure 1 shows the contour plot of $f(x,\lambda)$. It can be seen from this plot and can be checked formally that for $\lambda > 1$, there is only one $x = x(\lambda)$ that violates (12). Figure 2 shows the zero set of $f(x,\lambda)$ for $\lambda > 1$, compared with the bounds on the support of the Marchenko–Pastur distribution. It can be seen from this graph and can be checked formally that $x(\lambda) < t_{\min}(\lambda) = (1-\sqrt{\lambda})^2$. Consequently, if x is in the support of $\widetilde{\mu}_{\mathrm{mp}} \boxplus \widetilde{\mu}_{\mathrm{mp}}$, the genericity condition (3) holds, and the pair $(\widetilde{\mu}_{\mathrm{mp}}, \widetilde{\mu}_{\mathrm{mp}})$ is smooth at x. Hence, Theorem 1.3 applies, and the density of $\mu_n = \nu_{1,n} \boxplus \nu_{2,n}$ converges uniformly on I_ε to the density of $\widetilde{\mu}_{\mathrm{mp}} \boxplus \widetilde{\mu}_{\mathrm{mp}}$, that is, to the density of the Marchenko–Pastur distribution with parameter λ.

Similar results can be established for other limit theorems, except that it is more difficult to check the genericity condition (3) for a point in the support of the limit distribution. Here is one more theorem of this type. Let measures μ and ν be called equivalent ($\mu \sim \nu$) if there exist such real a and b, with $b > 0$, that for every Borel set $S \subset \mathbb{R}$, $\mu(S) = \nu(bS + a)$. Recall that a measure μ is called $\boxplus$-stable,

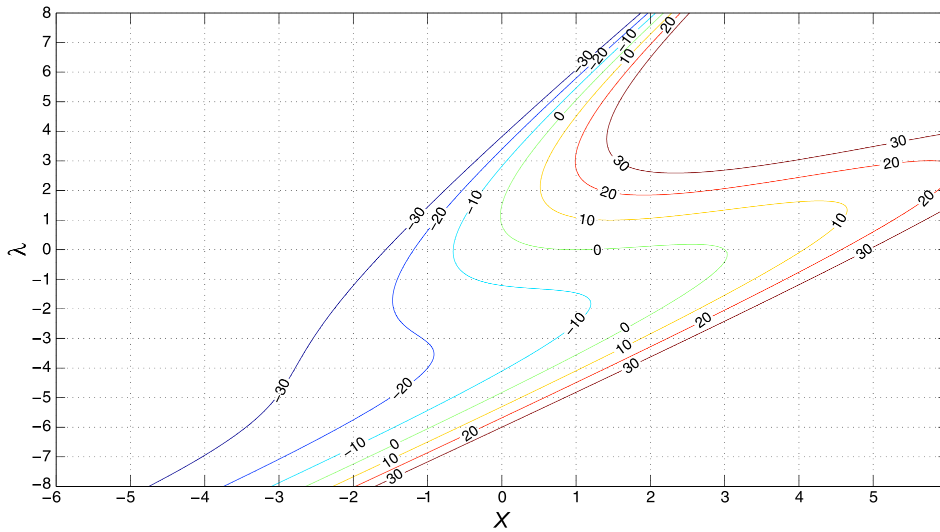

FIG. 1. *Contour plot of the right-hand side of* (12).

if $\mu \boxplus \mu \sim \mu$. The measure ν belongs to the domain of attraction of a $\boxplus$-stable law μ, if there exist measures ν_n equivalent to ν such that

$$\underbrace{\nu_n \boxplus \nu_n \boxplus \cdots \boxplus \nu_n}_{n \text{ times}} \to \mu.$$

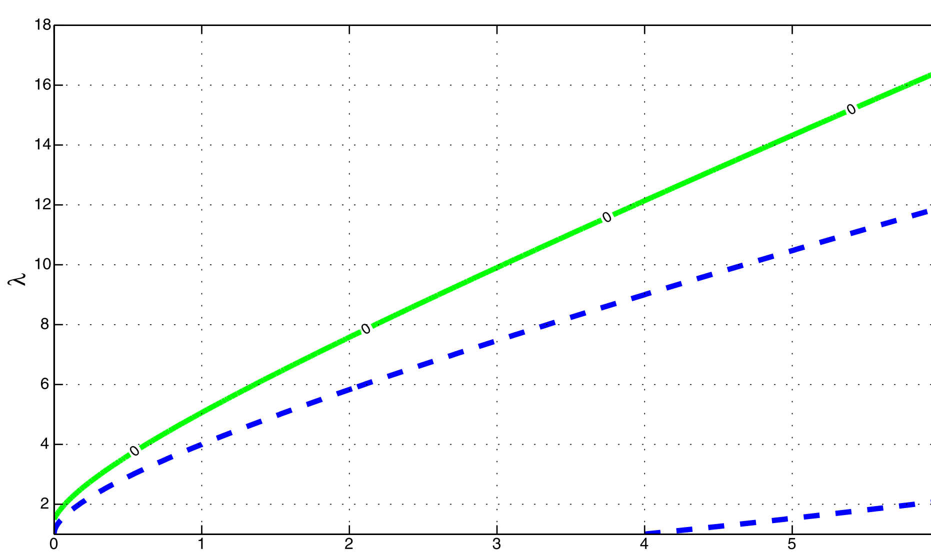

FIG. 2. *The zero set of the right-hand side of* (12) *compared with the support bounds for* $x(\lambda)$.

Clearly, in this case there exists a sequence of real constants $b_n > 0$ and a_n such that

$$\mu_n := \underbrace{\nu \boxplus \nu \boxplus \cdots \boxplus \nu}_{n \text{ times}}(b_n \cdot + a_n) \to \mu. \tag{13}$$

(More about the $\boxplus$-stability of probability measures and its relation to the classical stability of probability measures can be found in [3].)

THEOREM 4.3. *Suppose that a $\boxplus$-stable distribution μ is not equivalent to δ_0 and that ν belongs to the domain of attraction of μ. Let a_n, b_n and μ_n be defined as in* (13), *and let J be a bounded closed interval such that the density of μ is strictly positive on J. Then μ_n is* (*Lebesgue*) *absolutely continuous on J for all sufficiently large n, and there exist such real $\kappa_n > 0$ and ξ_n that the density of $\mu_n(\kappa_n \cdot + \xi_n)$ converges to the density of μ at* (*Lebesgue*) *almost all $E \in J$.*

PROOF. Let $J \subset I$, where the inclusion is strict, and I is a bounded, closed interval such that density of μ is strictly positive on I. (Interval I exists because by the results of Biane in [3] μ is absolutely continuous with analytical density.)

First, note that μ_n is (Lebesgue) absolutely continuous on $\mathbb{R}$ for all sufficiently large n. Indeed, for even $n = 2k$, the definition of μ_n implies that $\mu_{2k} = \mu_k \boxplus \mu_k(s_k^{-1} \cdot - t_k)$ for some constants t_k and $s_k > 0$. For large k, μ_k is close in the Lévy metric to μ, which is known to be absolutely continuous. Hence, μ_k has no atoms with weight $\geq 1/2$. This implies that μ_{2k} has no atoms at all. In addition, by results of Belinschi, μ_{2k} has no singular component. Therefore, μ_{2k} is absolutely continuous on $\mathbb{R}$ if k is sufficiently large. The argument for the odd $n = 2k + 1$ is similar if we write $\mu_{2k+1} = \mu_{k+1} \boxplus \mu_k(s_k \cdot + t_k)$.

In the second and final step, we note that there exists a sequence of constants $\kappa_n > 0$ and ξ_n such that the density of $\mu_n(\kappa_n \cdot + \xi_n)$ converges to the density of μ at (Lebesgue) almost all $x \in I$. Indeed, by the stability of μ, $\mu \boxplus \mu = \mu(s \cdot + t)$ and μ has positive analytic density on I; therefore, by Proposition 1.4 $\Im \omega_{\mu,j}(x) > 0$ at all $x \in (I - t)/s$. For almost all points x, the genericity condition (3) holds, since otherwise $k_\mu(x)$ (in the genericity condition) would be exactly 0 which is not possible. For even $n = 2k$, we have $\mu_k \boxplus \mu_k = \mu_{2k}(s_k \cdot + t_k)$, where $s_k > 0$ and t_k are certain real constants. Hence, by Theorem 1.3 the weak convergence $\mu_k \to \mu$ implies that the density of $\mu_k \boxplus \mu_k \equiv \mu_{2k}(s_k \cdot + t_k)$ converges to the density of $\mu \boxplus \mu \equiv \mu(s \cdot + t)$ at almost all points of $(I - t)/s$. It follows that for $\kappa_{2k} = s/s_k > 0$ and $\xi_{2k} = t - (s/s_k)t_k$, the density of $\mu_{2k}(\kappa_{2k} \cdot + \xi_{2k})$ converges to the density of μ at almost all points of I. The case of μ_{2k+1} can be handled similarly by considering $\mu_k \boxplus \mu_{k+1}$. $\square$

Our next application is of a different kind and answers a question that arises in the theory of large random matrices.

Let $H_N = A_N + U_N B_N U_N^*$, where A_N and B_N are N-by-N Hermitian matrices, and U_N is a random unitary matrix with the Haar distribution on the unitary group $\mathcal{U}(N)$.

Let $\lambda_1^{(A)} \geq \cdots \geq \lambda_N^{(A)}$ be the eigenvalues of A_N. Similarly, let $\lambda_k^{(B)}$ and $\lambda_k^{(H)}$ be ordered eigenvalues of matrices B_N and H_N, respectively.

Define the *spectral point measures* of A_N by $\mu_{A_N} := N^{-1} \sum_{k=1}^{N} \delta_{\lambda_k^{(A)}(H)}$, and define the spectral point measures of B_N and H_N similarly. Let $\mathcal{N}_I := N\mu_{H_N}(I)$ denote the number of eigenvalues of H_N in interval I, and let $\mathcal{N}_\eta(x) := \mathcal{N}_{(x-\eta,x+\eta]}$.

Let the notation $g_1(N) \ll g_2(N)$ mean that $\lim_{N\to\infty} g_2(N)/g_1(N) = +\infty$.

THEOREM 4.4. *Assume that*:

(1) $\mu_{A_N} \to \mu_\alpha$ *and* $\mu_{B_N} \to \mu_\beta$;
(2) $\operatorname{supp}(\mu_{A_N}) \cup \operatorname{supp}(\mu_{B_N}) \subseteq [-K, K]$ *for all* N;
(3) *the pair* (μ_α, μ_β) *is smooth at* x;
(4) $\frac{1}{\sqrt{\log(N)}} \ll \eta(N) \ll 1$.

Then

$$\frac{\mathcal{N}_\eta(x)}{2\eta N} \to f_{\mu_\alpha \boxplus \mu_\beta}(x)$$

with probability 1, *where* $f_{\mu_\alpha\boxplus\mu_\beta}$ *denotes the density of* $\mu_\alpha \boxplus \mu_\beta$.

Previously, it was shown by Pastur and Vasilchuk in [17] that assumption (1) together with a weaker version of assumption (2) implies that $\mu_{H_N} \to \mu_\alpha \boxplus \mu_\beta$ with probability 1. Theorem 4.4 says that the convergence of μ_{H_N} to $\mu_\alpha \boxplus \mu_\beta$ holds on the level of densities, so it can be seen as a local limit law for the eigenvalues of the sum of random Hermitian matrices.

PROOF OF THEOREM 4.4. In Theorem 2 in [14], it was shown that the following claim holds. Suppose that $\eta = \eta(N)$ and $1/\sqrt{\log N} \ll \eta(N) \ll 1$. Assume that the measure $\mu_{A_N} \boxplus \mu_{B_N}$ is absolutely continuous, and its density is bounded by a constant T_N. Then, for all sufficiently large N,

$$P\left\{\sup_x \left|\frac{\mathcal{N}_\eta(x)}{2N\eta} - f_{\boxplus,N}(x)\right| \geq \delta\right\} \leq \exp\left(-c\delta^2 \frac{(\eta N)^2}{(\log N)^2}\right), \tag{14}$$

where $c > 0$ depends only on $K_N := \max\{\|A_N\|, \|B_N\|\}$ and T_N. Here $f_{\boxplus,N}$ denotes the density of $\mu_{A_N} \boxplus \mu_{B_N}$.

This statement can be modified so that the supremum in the inequality holds for x in an interval, provided that the density of $\mu_{A_N} \boxplus \mu_{B_N}$ is bounded by a constant T_N in the following interval:

$$P\left\{\sup_{x\in(a,b)} \left|\frac{\mathcal{N}_\eta(x)}{2N\eta} - f_{\boxplus,N}(x)\right| \geq \delta\right\} \leq \exp\left(-c\delta^2 \frac{(\eta N)^2}{(\log N)^2}\right). \tag{15}$$

Since assumptions (1) and (3) hold, we can use Theorem 1.3 and infer that $f_{\boxplus,N}(x) \to f_{\mu_\alpha \boxplus \mu_\beta}(x)$. In particular, the sequence of densities $f_{\boxplus,N}(x)$ is uniformly bounded by a constant T. This fact and assumption (2) imply that the positive constant c in (14) can be chosen independently of N. By using the Borel–Cantelli lemma, we can conclude that

$$\frac{\mathcal{N}_\eta(x)}{2N\eta} \to f_{\mu_\alpha \boxplus \mu_\beta}(x)$$

with probability 1. □

5. Conclusion. We have proved that if probability measures ν_1 and ν_2 are sufficiently close to probability measures μ_1 and μ_2 in the Lévy distance, and if $\mu_1 \boxplus \mu_2$ is sufficiently smooth at x, then $\nu_1 \boxplus \nu_2$ is absolutely continuous at x, and its density is close to the density of $\mu_1 \boxplus \mu_2$.

We have applied this result to derive several local limit law results for sums of free random variables and for eigenvalues of a sum of random Hermitian matrices.

Acknowledgment. I would like to thank Diana Bloom for her editorial help and an anonymous referee for helpful suggestions.

REFERENCES

[1] BELINSCHI, S. T. (2008). The Lebesgue decomposition of the free additive convolution of two probability distributions. *Probab. Theory Related Fields* **142** 125–150. MR2413268

[2] BELINSCHI, S. T. and BERCOVICI, H. (2007). A new approach to subordination results in free probability. *J. Anal. Math.* **101** 357–365. MR2346550

[3] BERCOVICI, H. and PATA, V. (1999). Stable laws and domains of attraction in free probability theory. *Ann. of Math.* (2) **149** 1023–1060. MR1709310

[4] BERCOVICI, H. and VOICULESCU, D. (1993). Free convolution of measures with unbounded support. *Indiana Univ. Math. J.* **42** 733–773. MR1254116

[5] BERCOVICI, H. and VOICULESCU, D. (1995). Superconvergence to the central limit and failure of the Cramér theorem for free random variables. *Probab. Theory Related Fields* **103** 215–222. MR1355057

[6] BERCOVICI, H. and VOICULESCU, D. (1998). Regularity questions for free convolution. In *Nonselfadjoint Operator Algebras, Operator Theory, and Related Topics* (H. Bercovici and C. Foias, eds.). *Operator Theory: Advances and Applications* **104** 37–47. Birkhäuser, Basel. MR1639647

[7] BIANE, P. (1998). Processes with free increments. *Math. Z.* **227** 143–174. MR1605393

[8] BIANE, P. (1998). Representations of symmetric groups and free probability. *Adv. Math.* **138** 126–181. MR1644993

[9] FEINBERG, J. and ZEE, A. (1997). Non-Gaussian non-Hermitian random matrix theory: Phase transition and addition formalism. *Nuclear Phys. B* **501** 643–669. MR1477381

[10] GUIONNET, A., KRISHNAPUR, M. and ZEITOUNI, O. (2011). The single ring theorem. *Ann. of Math.* (2) **174** 1189–1217. MR2831116

[11] HAAGERUP, U. and THORBJØRNSEN, S. (2005). A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math.* (2) **162** 711–775. MR2183281

[12] HENRICI, P. (1986). *Applied and Computational Complex Analysis. Vol.* 3. Wiley, New York. MR0822470

[13] KANTOROVIČ, L. V. (1948). Functional analysis and applied mathematics. *Uspekhi Matem. Nauk* (*N.S.*) **3** 89–185. MR0027947

[14] KARGIN, V. (2012). A concentration inequality and a local law for the sum of two random matrices. *Probab. Theory Related Fields*. To appear. Available at http://arxiv.org/abs/1010.0353.

[15] MAASSEN, H. (1992). Addition of freely independent random variables. *J. Funct. Anal.* **106** 409–438. MR1165862

[16] NICA, A. and SPEICHER, R. (2006). *Lectures on the Combinatorics of Free Probability*. *London Mathematical Society Lecture Note Series* **335**. Cambridge Univ. Press, Cambridge. MR2266879

[17] PASTUR, L. and VASILCHUK, V. (2000). On the law of addition of random matrices. *Comm. Math. Phys.* **214** 249–286. MR1796022

[18] SHIRYAEV, A. N. (1996). *Probability*, 2nd ed. *Graduate Texts in Mathematics* **95**. Springer, New York. MR1368405

[19] SPEICHER, R. (1993). Free convolution and the random sum of matrices. *Publ. Res. Inst. Math. Sci.* **29** 731–744. MR1245015

[20] VOICULESCU, D. (1985). Symmetries of some reduced free product C^*-algebras. In *Operator Algebras and Their Connections with Topology and Ergodic Theory* (*BuŞteni*, 1983). *Lecture Notes in Math.* **1132** 556–588. Springer, Berlin. MR0799593

[21] VOICULESCU, D. (1986). Addition of certain noncommuting random variables. *J. Funct. Anal.* **66** 323–346. MR0839105

[22] VOICULESCU, D. (1991). Limit laws for random matrices and free products. *Invent. Math.* **104** 201–220. MR1094052

[23] VOICULESCU, D. (1993). The analogues of entropy and of Fisher's information measure in free probability theory. I. *Comm. Math. Phys.* **155** 71–92. MR1228526

[24] VOICULESCU, D. (1996). The analogues of entropy and of Fisher's information measure in free probability theory. III. The absence of Cartan subalgebras. *Geom. Funct. Anal.* **6** 172–199. MR1371236

[25] VOICULESCU, D. V., DYKEMA, K. J. and NICA, A. (1992). *Free Random Variables*: *A Noncommutative Probability Approach to Free Products With Applications to Random Matrices*, *Operator Algebras and Harmonic Analysis on Free Groups*. *CRM Monograph Series* **1**. Amer. Math. Soc., Providence, RI. MR1217253

[26] WANG, J.-C. (2010). Local limit theorems in free probability theory. *Ann. Probab.* **38** 1492–1506. MR2663634

[27] ZEE, A. (1996). Law of addition in random matrix theory. *Nuclear Phys. B* **474** 726–744. MR1404225

STATISTICAL LABORATORY
DEPARTMENT OF PURE MATHEMATICS
AND MATHEMATICAL STATISTICS
UNIVERSITY OF CAMBRIDGE, CAMBRIDGE
UNITED KINGDOM
E-MAIL: v.kargin@statslab.cam.ac.uk