Contents

Chapter 2

Non-Classical Convolution of Measures

Given a probability measure μ on $\mathbb{R}$ (with the Borel σ-field $B(\mathbb{R})$), let $\widehat{\mu}$ denote its Fourier transform $\widehat{\mu}(t) = \int_{\mathbb{R}} e^{itx}\mu(dx)$, $t \in \mathbb{R}$, and C_μ^* its (classical) cumulant transform $C_\mu^*(t) = \ln \widehat{\mu}(t)$, $t \in \mathbb{R}$.

Let μ_1 and μ_2 be two probability measures on $\mathbb{R}$ with Fourier transforms $\widehat{\mu}_1$ and $\widehat{\mu}_2$. The classical convolution of μ_1 and μ_2 is the probability measure $\mu_1 * \mu_2$ on $\mathbb{R}$ given by

$$\mu_1 * \mu_2(B) = \int_{\mathbb{R}} \mu_1(B-x)\mu_2(dx) = \int_{\mathbb{R}} \mu_2(B-x)\mu_1(dx),\ B \in B(\mathbb{R}). \tag{2.1}$$

Therefore $\widehat{\mu_1 * \mu_2}(t) = \widehat{\mu}_1(t)\ \widehat{\mu}_2(t)$ for each t in $\mathbb{R}$. In other words, if X_1 and X_2 are real random variables with distributions μ_1 and μ_2, respectively, and independent in the stochastic (tensor) sense, then

$$C^*_{\mu_1 * \mu_2}(t) = C^*_{\mu_1}(t) + C^*_{\mu_2}(t) \text{ for each } t \text{ in } \mathbb{R}. \tag{2.2}$$

That is, the classical cumulant transform linearizes the classical convolution for the stochastic or tensor independence.

There are analogous cumulant transforms that linearize each of the additive convolutions considered in Chapter 1 with respect to the five non-classical independencies. They are transformations of the Cauchy type rather than the Fourier type.

The goal of this chapter is to show a parallelism between classical and non-classical convolutions from the analytic point of view. We will begin in Section 2.1 with a collection of results on the Herglotz and Pick-Nevanlinna theory of analytic functions. In Section 2.2 and 2.3 we show the key role these functions play in the study of the Cauchy and other transforms used in the characterization of infinite divisibility with respect to the non-classical independencies. Section 2.4 presents the analytic approach of non-classical additive convolutions while

Section 2.5 briefly deals with the free multiplicative convolution and the S-transform. Finally, Section 2.6 presents an introduction to divisibility with respect to the non-classical independencies, and some first criteria for infinite divisibility of probability measures with compact support.

Throughout the chapter several examples are worked out, with special consideration to the non-classical Gaussian distributions introduced in Chapter 1.

2.1 Herglotz or Pick functions

Let $\mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ and $\mathbb{C}^- = \{z \in \mathbb{C} : \mathrm{Im}(z) < 0\}$. A Herglotz function is an analytic function $H : \mathbb{C}^+ \to \mathbb{C}^+ \cup \mathbb{R}$ and by reflection it is extended to an analytic function on $\mathbb{C}\backslash\mathbb{R}$. Such functions are also called Pick or Pick-Nevanlinna functions. In this section we collect some properties as well as some first examples of these functions.

There are several equivalent representations for Herglotz functions. The canonical representation is as follows.

Theorem 1 *H is a Herglotz function if and only if*

$$H(z) = \gamma + \psi z + \int_{\mathbb{R}} \left(\frac{1}{t-z} - \frac{t}{t^2+1} \right) \rho(dt), \qquad z \in \mathbb{C}^{\pm}, \tag{2.3}$$

where $\gamma \in \mathbb{R}$, $\psi \geq 0$ and ρ is a measure on $\mathbb{R}$ satisfying

$$\int_{\mathbb{R}} \frac{1}{t^2+1} \rho(dt) < \infty.$$

Moreover, the triplet (γ, ψ, ρ) is uniquely determined by H, using

$$\gamma = \mathrm{Re}(H(i)),\ \psi = \lim_{y\to\infty} \frac{H(iy)}{iy} \tag{2.4}$$

and the Stieltjes inversion formula

$$\rho((t_0,t_1]) = \frac{1}{\pi} \lim_{\delta\to 0+} \lim_{y\to 0+} \int_{t_0+\delta}^{t_1+\delta} \mathrm{Im}(H(x+iy))dx, \qquad t_0 < t_1. \tag{2.5}$$

Remark 2 *The measure ρ is usually called the spectral measure, but we do not use that name in the present work. We rather call it the ρ-measure of H.*

An alternative representation of a Herglotz function is the Nevanlinna representation:

$$H(z) = \gamma + \psi z + \int_{\mathbb{R}} \left(\frac{1+tz}{t-z} \right) \sigma(dt), \tag{2.6}$$

with γ and ψ as in (2.4) and where σ is the finite measure

$$\sigma(u) = \sigma((-\infty,u]) = \int_{-\infty}^{u} \frac{1}{t^2+1} \rho(dt). \tag{2.7}$$

Elementary examples of Herglotz functions and their canonical representations are the following:

(i) $H(z) = cz$ for $c \geq 0$. In this case $\gamma = 0$, $\psi = c$ and $\rho = 0$.

(ii) $H(z) = -1/z$, for which $\gamma = \psi = 0$ and $\rho = \delta_0$.

(iii) $H(z) = \ln(z)$, with the branch $(-\pi/2, 3\pi/2)$. In this case $\gamma = \psi = 0$ and $\rho(dt) = 1_{(-\infty,0)}(t)dt$.

(iv) $H(z) = \sqrt{z}$, where the square root is taken positive on the right half-axis. The canonical representation is

$$H(z) = \frac{1}{\sqrt{2}} + \int_{-\infty}^{0} (\frac{1}{t-z} - \frac{t}{t^2+1}) \frac{\sqrt{t}}{\pi} dt, \ z \in \mathbb{C}^{\pm}.$$

The class of Herglotz functions is closed under addition and composition. In particular, if H is a Herglotz function, $\ln(H(z))$ and $-1/H(z)$ are also Herglotz functions. Then, the representation (2.3) for $\ln(H(z))$ trivially gives an exponential representation for H. Since the imaginary part of $\ln(z)$ is uniformly bounded, the ρ-measure of $\ln(H(z))$ is absolutely continuous with respect to the Lebesgue measure. More precisely,

$$H(z) = \exp\left(c + \int_{-\infty}^{\infty} \left(\frac{1}{t-z} - \frac{t}{1+t^2}\right) \xi(t) \mathrm{d}t\right) \tag{2.8}$$

where $c = \log |H(\mathrm{i})| \in \mathbb{R}, \xi \in L^1\left(\mathbb{R}, (1+\lambda^2)^{-1} \mathrm{d}\lambda\right)$, with ξ the non-zero measurable function

$$\xi(\lambda) = \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \operatorname{Im}\left(\log\left(H\left(\lambda + \mathrm{i}\varepsilon\right)\right)\right) = \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \arg(H(\lambda + i\varepsilon)). \tag{2.9}$$

The main feature of (2.8) is the absolute continuity of the ρ-measure.

For example the function $H(z) = \tan(z)$ is a Herglotz function having exponential representation (2.8) with $c = \ln \tanh(1)$ and $\xi(t) = 1_{\{\tanh(t)<0\}}$.

The choice of the representation of a Herglotz function depends on the context of the problem under consideration. For example, (2.6) is essential in Chapter 3 to obtain the Lévy-Khintchine representation of infinitely divisible distributions with respect to the different independencies. On the other hand, the Herglotz exponential representation (2.8) links free probability to the seemingly unrelated area of representation theory of symmetric groups. Although we will not consider this latter direction, in Section 2.2.3 we compute the exponential representation of the Cauchy transform of the Gaussian distributions for each of the independencies considered in Chapter 1.

Finally, another useful characterization of Herglotz functions is given in terms of the so-called Pick-matrices as follows. As in the Bochner theorem for Fourier transforms, the notion of positive definiteness is used, although in a different context.

Theorem 3 *Let D be a domain in $\mathbb{C}^+$. Given a function $H : D \to \mathbb{C}$ the following assertions are equivalent:*

i) H can be extended to a Herglotz function to all $\mathbb{C}^+$.

ii) For any $n \geq 1$ and $z_1, ..., z_n$ in D the matrix

$$\left[\frac{H(z_j) - \overline{H(z_k)}}{z_j - \overline{z_k}}\right]_{j,k}$$

is positive semi-definite.

2.2 The Cauchy transform

2.2.1 Definition and properties

Given a finite measure μ on $\mathbb{R}$, its Cauchy transform G_μ is defined as

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-t}\mu(dt), \qquad z \in \mathbb{C}\backslash\mathbb{R}. \tag{2.10}$$

The Cauchy and the Fourier transforms are related by the expression

$$G_\mu(z) = \begin{cases} i\int_{-\infty}^{0} e^{-itz}\widehat{\mu}(t)dt, & \operatorname{Im}(z) > 0 \\ -i\int_{0}^{\infty} e^{-itz}\widehat{\mu}(t)dt, & \operatorname{Im}(z) < 0. \end{cases} \tag{2.11}$$

Writing $z = x + iy$ one easily obtains the following decomposition of $G_\mu(z)$ in its real and imaginary parts:

$$G_\mu(x+iy) = \int_{\mathbb{R}} \frac{x-t}{(x-t)^2+y^2}\mu(dt) - i\int_{\mathbb{R}} \frac{y}{(x-t)^2+y^2}\mu(dt). \tag{2.12}$$

Moreover, $H_\mu(z) = -G_\mu(z)$ is a Herglotz function with representation (2.3) with $\psi = 0$, $\rho = \mu$ and

$$\gamma = \int_{\mathbb{R}} \frac{t}{1+t^2}\mu(dt). \tag{2.13}$$

From the above considerations we easily obtain the following properties.

Proposition 4 *Let μ be a finite measure on $\mathbb{R}$. Then*
i) $G_\mu(\mathbb{C}^\pm) \subset \mathbb{C}^\mp$ and $G_\mu(\overline{z}) = \overline{G_\mu(z)}$.
ii) $|G_\mu(z)| \leq \dfrac{\mu(\mathbb{R})}{|\operatorname{Im}(z)|}$.
iii)

$$-\operatorname{Im}(G_\mu(z)) = \operatorname{Im}(z)\int_{\mathbb{R}} \frac{1}{|z-t|^2}\mu(\mathrm{d}t)$$

and $\operatorname{Im}(z)\operatorname{Im}G_\mu(z) < 0$.
iv) $\lim_{y\to\infty} y\,|G_\mu(iy)| < \infty$.
v) $\lim_{y\to\infty} iyG_\mu(iy) = \mu(\mathbb{R})$. In particular, if μ is a probability measure

$$\lim_{y\to\infty} iyG_\mu(iy) = 1.$$

As a special case of the Stieltjes inversion formula (2.5), one can recover the distribution μ from the Cauchy transform G_μ as follows:

$$\mu((t_0,t_1]) = -\frac{1}{\pi}\lim_{\delta\to 0+}\lim_{y\to 0+}\int_{t_0+\delta}^{t_1+\delta} \mathrm{Im}(G_\mu(x+iy))dx, \quad t_0 < t_1. \tag{2.14}$$

In particular, if μ is absolutely continuous with respect to Lebesgue measure with "density" f_μ then

$$f_\mu(x) = -\frac{1}{\pi}\lim_{y\to 0+}\mathrm{Im}\, G_\mu(x+iy). \tag{2.15}$$

Thus, there is a one-to-one correspondence between finite measures on $\mathbb{R}$ and their Cauchy transforms. The following result characterizes the analytic functions that are Cauchy transforms of probability measures. It is analogous to the Bochner theorem for Fourier transforms. Let us introduce the notation Γ_α for the Herglotz region

$$\Gamma_\alpha = \{z = x+iy : y > 0,\ x < \alpha y\}, \text{ for } \alpha > 0. \tag{2.16}$$

Proposition 5 *Let $G : \mathbb{C}^+ \to \mathbb{C}^-$ be an analytic function. The following three statements are equivalent:*

i) There exists a probability measure μ on $\mathbb{R}$ such that $G_\mu = G$ in $\mathbb{C}^+$.

ii) For each $\alpha > 0$

$$\lim_{|z|\to\infty,\ z\in\Gamma_\alpha} zG(z) = 1.$$

iii) $\lim_{y\to+\infty} iyG(iy) = 1.$

Proof. Equivalence between (i) and (iii) follows from Proposition 4(v) and the representation (2.3). It is trivial that (iii) follows from (ii), since $iy \in \Gamma_\alpha$ for each $y > 0$. It remains to prove that a Cauchy transform satisfies (ii). Now observe that the function $t^2 - ((x-t)^2 + y^2)(\alpha^2+1)$ has a maximum at $t = (\alpha^2 + 1/\alpha^2)x$, and one has the inequality $|t/(z-t)| \le |\alpha + i| = \sqrt{\alpha^2+1}$ for each $t \in \mathbb{R}$. Then, for each $T > 0$, $z = x+iy \in \Gamma_\alpha$ and writing $C = \sqrt{\alpha^2+1}$

$$\begin{aligned}
|zG_\mu(z) - 1| &= \left|\int_{\mathbb{R}} \frac{t}{z-t}\mu(dt)\right| \\
&\le C\mu(\{t : |t| \ge T\}) + \int_{(-T,T)} \left|\frac{t}{z-t}\right| \mu(dt) \\
&\le C\mu(\{t : |t| \ge T\}) + \frac{T}{y}\mu((-T,T)).
\end{aligned}$$

Hence,

$$\limsup_{|z|\to\infty,\ z\in\Gamma_\alpha} |zG_\mu(z) - 1| \le C\mu(\{t : |t| \ge T\}.$$

The results follows since $\lim_{T\to\infty} \mu(\{t : |t| \geq T\}) = 0$. ■

The relation between weak convergence of probability measures and the Cauchy transform is given by the following Lévy continuity type theorem.

Proposition 6 *Let μ_1 and μ_2 be two probability measures on $\mathbb{R}$ and*

$$d(\mu_1, \mu_2) = \sup\left\{\left|G_{\mu_1}(z) - G_{\mu_2}(z)\right|; \quad \mathrm{Im}(z) \geq 1\right\}.$$

Then d is a distance which defines the weak topology of probability measures. In other words, a sequence of probability measures $\{\mu_n\}_{n\geq 1}$ on $\mathbb{R}$ converges weakly to a probability measure μ on $\mathbb{R}$ if and only if

$$\lim_{n\to\infty} G_{\mu_n}(z) = G_\mu(z) \text{ for each } z, \quad \mathrm{Im}(z) \geq 1.$$

There are useful properties of the Cauchy transform when μ is a symmetric probability measure, i.e., $\mu(A) = \mu(-A)$ $\forall A \in B(\mathbb{R})$. Given a symmetric probability measure μ on $\mathbb{R}$, let $\mu^{(2)}$ be the probability measure in $\mathbb{R}_+$ induced by the map $t \to t^2$. More generally, for a probability measure μ on $\mathbb{R}$, the p–th push–forward measure of $\mu^{(p)}$ of μ is defined as

$$\mu^{(p)}(B) = \int_{\mathbb{R}} 1_B(|x|^p)\mu(\mathrm{d}x), \quad B \in \mathcal{B}((0,\infty)).$$

For a symmetric measure we have the following properties of its Cauchy transform. The first is an application of the general inversion formula (2.14).

Proposition 7 *A Borel measure is symmetric if and only if its Cauchy transform is an even function*

Proof. Assume μ is symmetric. Then

$$\begin{aligned} G_\mu(-z) &= \int_{\mathbb{R}} \frac{1}{-z-t}\mu(\mathrm{d}t) = -\int_{\mathbb{R}} \frac{1}{z+t}\mu(\mathrm{d}t) \\ &= -\int_{\mathbb{R}} \frac{1}{z-y}\mu(-\mathrm{d}y) = -\int_{\mathbb{R}} \frac{1}{z-y}\mu(\mathrm{d}y) = -G_\mu(z). \end{aligned}$$

Conversely, if $G_\mu(z) = -G_\mu(-z) = -\overline{G_\mu(-\overline{z})}$, using (2.14), for $t_0 < t_1$

$$\begin{aligned} \mu((t_0,t_1]) &= -\frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{y\to 0^+}\int_{t_0+\delta}^{t_1+\delta} \mathrm{Im}(G_\mu(x+iy))\mathrm{d}x \\ &= -\frac{1}{\pi}\lim_{\delta\to 0^+}\lim_{y\to 0^+}\int_{t_0+\delta}^{t_1+\delta} \mathrm{Im}(-G_\mu(-x-iy))\mathrm{d}x \\ &= -\frac{1}{\pi}\lim_{\delta_0\to 0^-}\lim_{y\to 0^+}\int_{t_0-\delta_0}^{t_1-\delta_0} \mathrm{Im}(G_\mu(-x+iy))\mathrm{d}x \\ &= -\frac{1}{\pi}\lim_{\delta_0\to 0^-}\lim_{y\to 0^+}\int_{-t_1+\delta_0}^{-t_0+\delta_0} \mathrm{Im}(G_\mu(x+iy))\mathrm{d}x \\ &= \mu([-t_1, -t_0)), \end{aligned}$$

showing the symmetry of μ. ■

Proposition 8 *Let μ be a symmetric probability measure μ on $\mathbb{R}$. Then*

$$G_\mu(z) = zG_{\mu^{(2)}}(z^2), z \in \mathbb{C} \setminus \mathbb{R}_+.$$

Proof. Use the symmetry of μ twice to obtain

$$\begin{aligned} G_\mu(z) &= \int_{\mathbb{R}} \frac{1}{z-t}\mu(\mathrm{d}t) = \int_{\mathbb{R}_+} \frac{1}{z-t}\mu(\mathrm{d}t) + \int_{\mathbb{R}_+} \frac{1}{z+t}\mu(\mathrm{d}t) \\ &= 2z \int_{\mathbb{R}_+} \frac{1}{z^2-t^2}\mu(\mathrm{d}t) = z\int_{\mathbb{R}} \frac{1}{z^2-t^2}\mu(\mathrm{d}t) \\ &= z\int_{\mathbb{R}_+} \frac{1}{z^2-t}\mu^{(2)}(\mathrm{d}t) = zG_{\mu^{(2)}}(z^2). \end{aligned}$$

■

Example 9 (Cauchy distribution) *For $\lambda > 0$, the Cauchy distribution c_λ on $\mathbb{R}$ has density*

$$f(x) = \frac{1}{\pi}\frac{\lambda}{\lambda^2 + x^2}, \quad -\infty < x < \infty.$$

Its Fourier transform is easily computed as

$$\widehat{\mathrm{c}_\lambda}(x) = \exp(-\lambda\,|x|).$$

This distribution does not have moments of any order. It is important in non-commutative probability as it will be seen throughout this work. The relation (2.11) between Fourier and Cauchy transforms gives

$$\begin{aligned} G_{\mathrm{c}_\lambda}(z) &= -i\int_{-\infty}^{0} e^{-itz}\widehat{\mathrm{c}_\lambda}(t)dt = -i\int_{-\infty}^{0} e^{-itz}e^{-\lambda|t|}dt \\ &= -i\int_{-\infty}^{0} e^{(\lambda-iz)t}dt = \frac{1}{z+\lambda i},\ z \in \mathbb{C}^+. \end{aligned}$$

2.2.2 Measures with compact support

There are several important distributions with compact support which arise naturally in non-commutative probability, as explained in Chapter 1 and Appendix A. In this section we include some general results for compactly supported probability measures.

If μ has compact support, say $[a, b]$, for some $-\infty < a < b < \infty$, the Cauchy transform has the power series expansion

$$G_\mu(z) = z^{-1} + \sum_{k=1}^{\infty} m_k(\mu)z^{-k-1}, \quad |z| > r_\mu, \tag{2.17}$$

where $m_k(\mu) := \int_{\mathbb{R}} t^k\mu(dt)$ is the k-moment of μ, $k \geq 0$ and $r_\mu := \sup\{|t| : t \in supp(\mu)\}$. For this reason G_μ is thought as a moment generation function.

Likewise, it is possible to recognize from the behavior of its moments when a distribution with all moments has bounded support. Specifically, a Borel measure μ on $\mathbb{R}$ with all moments $m_k(\mu), k \geq 1$, has bounded support if and only if there is a $M > 0$ such that

$$\limsup_{k\to\infty} \sqrt[2k]{m_{2k}} \leq M. \tag{2.18}$$

In other words, the moments do not grow faster than exponentially.

For a compactly supported probability measure there is a simple exponential representation for its Cauchy transform.

Proposition 10 *Let μ be a probability measure with bounded support $[a,b]$. Then,*

$$-G_\mu(z) = \exp\left(c + \int_{-\infty}^{\infty}\left(\frac{1}{t-z} - \frac{t}{1+t^2}\right)\xi(t)\mathrm{d}t\right), \tag{2.19}$$

where the behavior of ξ outside the support is given by

$$\xi(t) = \begin{cases} 0, & \textit{if } t < a \\ 1, & \textit{if } t > b \end{cases}$$

Proof. Let $t \notin [a,b]$. Then

$$\lim_{\varepsilon\downarrow 0}\left(-G_\mu\left(t+\mathrm{i}\varepsilon\right)\right) = -\int_{[a,b]}\frac{1}{\lambda - s}\mu(\mathrm{d}s) \in \begin{cases} \mathbb{R}^+, & \text{if } t < a \\ \mathbb{R}^-, & \text{if } t > b \end{cases}$$

If we apply equations (2.8) and (2.9) to the Herglotz function $-G_\mu$, we obtain

$$\xi(t) = \frac{1}{\pi}\lim_{\varepsilon\downarrow 0}\arg\left(-G_\mu\left(t+\mathrm{i}\varepsilon\right)\right) = \frac{1}{\pi}\arg\left(-\int_{[a,b]}\frac{1}{t-s}\mu(ds)\right) = \begin{cases} 0, & \text{if } t < a \\ 1, & \text{if } t > b \end{cases}.$$

■

Example 11 (Discrete distribution)

If $\mu = \sum_{i=1}^{n} p_i\delta_{x_i}$ is a discrete measure on $A = \{x_1 < x_2 < \ldots < x_n\}$, for some $n \geq 1$. Then its negative Cauchy transform

$$-G_\mu(z) = \sum_{i=1}^{n}\frac{-p_i}{z-x_i} = \frac{\sum_{i=1}^{n} p_i \prod_{j\neq i}^{n}(x_j - z)}{\prod_{i=1}^{n}(z-x_i)} = \frac{P(z)}{Q(z)}.$$

is actually holomorphic in $\mathbb{C}\backslash A$. Clearly, for all $x \in \mathbb{R}\backslash A$, we have that

$$\lim_{\varepsilon\downarrow 0}\left(-G_\mu\left(x+\mathrm{i}\varepsilon\right)\right) = \sum_{i=1}^{n}\frac{-p_i}{x-x_i} \in \mathbb{R}.$$

Therefore the argument (and hence $\xi(x)$) will depend only on the sign of the last expression. If we look at the restriction of $-G_\mu$ to $\mathbb{R}\backslash A$, we can see that

$$\lim_{x\to x_i^\pm}\sum_{i=1}^n \frac{-p_i}{x-x_i} = \mp\infty.$$

By continuity, there exists $y_i \in (x_i, x_{i+1})$, such that $-G_\mu(y_i) = 0$, $1 \le i \le n-1$. If we write

$$\sum_{i=1}^n \frac{-p_i}{x-x_i} = \frac{-\sum_{i=1}^n p_i \prod_{j\neq i}^n (x-x_j)}{\prod_{i=1}^n (x-x_i)} = \frac{P(x)}{Q(x)}.$$

and since $\deg P = n-1$, then $B = \{y_1, \dots, y_{n-1}\}$ are all the roots of P. Therefore, using the knowledge of the roots of P and Q we may compute the sign of our expression. We obtain that for $x \in \mathbb{R}\backslash(A\cup B)$,

$$sgn(-G_\mu(x)) = \begin{cases} 1, & \text{if } x < x_1 \text{ or } x \in (y_i, x_{i+1}), \\ -1, & \text{if } x > x_n \text{ or } x \in (x_i, y_i), \end{cases}$$

and hence

$$\xi(x) = \begin{cases} 1, & \text{if } x > x_n \text{ or } x \in (x_i, y_i), \\ 0, & \text{if } x < x_1 \text{ or } x \in (y_i, x_{i+1}). \end{cases}$$

Finally, when a probability measure μ on $\mathbb{R}$ has all moments, we can expand its Cauchy transform of μ as a continued fraction as follows

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-t}\mu(dt) = \cfrac{1}{z-\beta_0-\cfrac{\gamma_0}{z-\beta_1-\cfrac{\gamma_1}{z-\beta_2-\cfrac{\gamma_2}{\ddots}}}}$$

Here the families $\gamma_m = \gamma_m(\mu) \ge 0$ and $\beta_m = \beta_m(\mu) \in \mathbb{R}$ are called Jacobi parameters and are defined by the recursion

$$xP_m(x) = P_{m+1}(x) + \beta_m P_m(x) + \gamma_{m-1}P_{m-1}(x),$$

where the polynomials $P_{-1}(x) = 0$, $P_0(x) = 1$ and $(P_m)_{m\ge 0}$ is a sequence of orthogonal polynomials with respect to μ, that is,

$$\int_{\mathbb{R}} P_m(x)P_n(x)\mu(dx) = 0 \qquad \text{if } m \neq n.$$

2.2.3 Examples

The Central Limit Theorems 4-6 in Chapter 1 found the moments of the Gaussian distributions with respect to Boolean, free and monotone independencies. In this section we compute the Cauchy transforms of the distributions with the corresponding moments and we identify, via the Stieltjes inversion formula (2.15), the associated distributions: symmetric Bernoulli, semicircle and arcsine distribution. We also compute the Cauchy transform of the Marchenko-Pastur distribution already mentioned at the end of Section 2.3. For all these distributions it is possible to find a closed form for their Cauchy transform. For the classical Gaussian distribution there does not exists a closed form of its Cauchy transform.

2.2.4.1 Symmetric Bernoulli Distribution

Let β_σ be the **Symmetric Bernoulli distribution on** $\{-\sigma, \sigma\}$ given by

$$\beta_\sigma = \frac{1}{2}(\delta_{-\sigma} + \delta_\sigma).$$

Theorem 4 in Chapter 1 establishes that β_σ is the Gaussian distribution with respect to Boolean independence. Simple calculations give

$$G_{\beta_\sigma}(z) = \int_{\mathbb{R}} \frac{1}{z-t}\beta_\sigma(dt) = \frac{1}{2}\left(\frac{1}{z+\sigma} + \frac{1}{z-\sigma}\right) = \frac{z}{z^2-\sigma^2}$$

Moreover, by Proposition 10

$$-G_{\beta_\sigma}(z) = \exp\left(c + \int_{-\infty}^{\infty}\left(\frac{1}{t-z} - \frac{t}{1+t^2}\right)\xi_{\beta_\sigma}(t)\mathrm{d}t\right)$$

where ξ_{β_σ} is given by

$$\xi_{\beta_\sigma}(t) = \mathbf{1}_{(\sigma,\infty)}(t).$$

2.2.4.2 Semicircle distribution

It was seen in Chapter 1 that the semicircle distribution is the limiting distribution in the free central limit theorem, playing the role the Gaussian distribution does in classical probability. For this reason the semicircle distribution is sometimes called the free Gaussian.

This distribution appears in several problems related to free probability and random matrices. The most important point is the fact that it is the asymptotic spectral distribution of Gaussian and more general ensembles of random matrices, when the dimension goes to infinite - as mentioned also in Chapter 1 and detailed in Chapter 6; a result that goes back to the pioneering work of Wigner. For that reason, the semicircle distribution is also called the Wigner distribution.

To compute the Cauchy transform and its density we start from the sequence of moments found in the Central Limit Theorem with respect to free independence (Theorem 5 in Chapter 1). That is, let μ be a probability measures on

$$m_k(\mu) = \begin{cases} 0 & \text{if } k \text{ is odd} \\ C_n & \text{if } k \text{ is even, } k = 2n \end{cases} \quad (k \geq 1) \tag{2.20}$$

where $C_n = \binom{2n}{n}/(n+1)$ are the Catalan numbers. These moments satisfied the condition (2.18) and hence μ has bounded support. Then the moments generating expansion (2.17) exists and is given by

$$G_\mu(z) = \frac{1}{z} + \sum_{k=1}^{\infty} \frac{C_k}{z^{2k+1}}. \tag{2.21}$$

Use now the useful recurrence equation for the Catalan numbers $C_k = \sum_{j=1}^{k} C_{j-1}C_{k-j}$ to obtain

$$G_\mu(z) = \frac{1}{z} + \frac{1}{z} G_\mu(z) G_\mu(z).$$

Thus, $G_\mu(z)$ satisfies the quadratic equation

$$G_\mu(z)^2 - zG_\mu(z) + 1 = 0, \qquad z \in \mathbb{C}^+ \tag{2.22}$$

whose solution for $G_\mu(z) : \mathbb{C}^+ \to \mathbb{C}^-$ is

$$G_\mu(z) = \frac{z - \sqrt{z^2 - 4}}{2}, \tag{2.23}$$

where the negative sign of the square root is taken, since (iii) in Proposition 5 has to be satisfied.

Now, using the Stieltjes inversion formula (2.15), the semicircle distribution $\mathrm{w}_{0,2}$ on $(-2,2)$ is found to have density

$$w_{0,2}(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \cdot 1_{[-2,2]}(x). \tag{2.24}$$

More generally, for m real and $\sigma^2 > 0$, the density of the semicircle distribution $\mathrm{w}_{m,\sigma}$ is given by

$$w_{m,\sigma}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - (x-m)} \cdot 1_{[m-2\sigma, m+2\sigma]}(x). \tag{2.25}$$

This distribution has mean m and variance σ^2. Its Cauchy transform is

$$G_{\mathrm{w}_{m,\sigma}}(z) = \frac{2}{r^2}(z - \sqrt{(z-m)^2 - r^2}),$$

where $r = 2\sigma$ is the radius of the support.

Finally, for the semicircle distribution $\mathrm{w}_{0,\sigma}$ on $(-2\sigma, 2\sigma)$, the negative of its Cauchy transform has the exponential representation (2.19) with $\xi_{\mathrm{w}_{0,L}}(\lambda)$ computed from (2.8) as follows:

$$\begin{aligned} \xi_{\mathrm{w}_{0,L}}(\lambda) &= \lim_{\varepsilon \downarrow 0} \arg\left(\frac{-2}{\sigma^2}\left(\lambda + \mathrm{i}\varepsilon - \sqrt{(\lambda + \mathrm{i}\varepsilon)^2 - \sigma^2}\right)\right) \\ &= \mathbf{1}_{(-2\sigma,2\sigma)}(\lambda)\sin^{-1}\left(\frac{\lambda}{2\sigma}\right) + \mathbf{1}_{(2\sigma,\infty)}(\lambda). \end{aligned}$$

2.2.4.3. Arcsine distribution

It was also seen in Chapter 1 that the arcsine distribution is the limiting distribution in the central limit theorem with respect to monotone and anti-monotone independence. More specifically, Theorem 6 in Chapter 1 says that the limiting distribution in the Central Limit Theorem with respect monotone convolution has moments given by

$$m_k(\mu) = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \binom{2n}{n} & \text{if } k \text{ is even, } k = 2n. \end{cases}.$$

This moment sequence satisfies condition (2.18) and hence its distribution has bounded support. Let G be the Cauchy transform (2.23) and consider the function $H(z) = G(z)/z$. Then

$$\frac{\partial}{\partial z} H(z) = \frac{-2}{z^2\sqrt{z^2-4}}.$$

On the other hand, taking derivatives in (2.21)

$$\begin{aligned} \frac{\partial}{\partial z} H(z) &= \sum_{k=0}^{\infty} \frac{\partial}{\partial z}\left(\frac{C_k}{z^{2k+2}}\right) = \sum_{k=0}^{\infty}\left(\frac{-(2k+2)C_k}{z^{2k+3}}\right) \\ &= -2\sum_{k=0}^{\infty} \frac{\binom{2k}{k}}{z^{2k+3}} = \frac{-2G_\mu(z)}{z^2}, \end{aligned}$$

gives

$$G_\mu(z) = \frac{1}{\sqrt{z^2-4}}. \tag{2.26}$$

Using the Stieltjes inversion formula (2.15) we obtain the arcsine distribution $\mathrm{a}_{0,2}$ with density

$$a_{0,2}(x) = \frac{1}{\pi}\frac{1}{\sqrt{4-x^2}} \cdot 1_{(-2,2)}(x). \tag{2.27}$$

In general, for any $\sigma > 0$, let $\mathrm{a}_{0,\sigma}$ denote the arcsine distribution on $(-\sigma, \sigma)$ with density

$$a_{0,\sigma}(x) = \frac{1}{\pi}\frac{1}{\sqrt{\sigma^2-x^2}} \cdot 1_{(-\sigma,\sigma)}(x).$$

In this case,

$$G_{\mathrm{a}_{0,\sigma}}(z) = \frac{1}{\sqrt{z^2-\sigma^2}}$$

and, by Proposition 10, $-G_{\mathrm{a}_{0,\sigma}}$ has the exponential representation (2.19) with

$$\begin{aligned} \xi_{\mathrm{a}_{0,\sigma}}(\lambda) &= \lim_{\varepsilon \downarrow 0} \arg\left(\frac{-1}{\sqrt{(\lambda + \mathrm{i}\varepsilon)^2 - \sigma^2}}\right) \\ &= \frac{1}{2}\mathbf{1}_{(-\sigma,\sigma)}(\lambda) + \mathbf{1}_{(\sigma,\infty)}(\lambda). \end{aligned}$$

2.2.4.4 Marchenko-Pastur distribution

This distribution was also already mentioned in Chapter 1. It plays a key role in several areas. One of the most important comes from the fact that it is the asymptotic spectral distribution of Wishart and more general ensembles of sample covariance random matrices, when the dimension goes to infinite - as mentioned also in Chapter 1 and also explained in Chapter 5; a result that goes back to the pioneering work of Marchenko and Pastur. It also plays in free probability a role analogous to that of the Poisson distribution in classical probability. For that reason, this distribution is also called the free Poisson distribution.

This distribution is considered in more generality in Section 2.4.4. Here we introduce a particular case arising as the square of a random variable with a symmetric semicircle distribution. More specifically, if s has the semicircle distribution $\mathrm{w}_{0,2}$ on $(-2,2)$, then s^2 has the Marchenko-Pastur distribution with density

$$f(x) = \frac{\sqrt{x(4-x)}}{2\pi x}.$$

From Proposition 8 we then have that the Cauchy transform of $\mu^{(2)}$ can be computed from

$$zG_{\mathrm{w}^{(2)}}(z^2) = \frac{z-\sqrt{z^2-4}}{2},$$

that is

$$G_{\mathrm{w}^{(2)}}(z) = \frac{1}{2} - \frac{\sqrt{z(z-4)}}{2z}.$$

Moreover,

$$\arg(-G_{\mathrm{w}^{(2)}}(z)) = \arg(-2G_{\mathrm{w}^{(2)}}(z)) = \arg\left(\frac{\sqrt{z(z-4)}}{z} - 1\right)$$

and

$$\begin{aligned}\lim_{\varepsilon\downarrow 0}\arg\left(\frac{\sqrt{(x+\mathrm{i}\varepsilon)\,(x+\mathrm{i}\varepsilon-4)}}{x+\mathrm{i}\varepsilon} - 1\right) &= \arg\left(\frac{\sqrt{x(x-4)}}{x} - 1\right)\\ &= \arg\left(\frac{\mathrm{i}\sqrt{x(4-x)}}{x} - 1\right).\end{aligned}$$

Observing that

$$\left\|\frac{\mathrm{i}\sqrt{x(4-x)}}{x} - 1\right\| = \sqrt{\left(\frac{\sqrt{x(4-x)}}{x}\right)^2 + 1^2} = \sqrt{\frac{4}{x}},$$

we find

$$\arg\left(\frac{\mathrm{i}\sqrt{x(4-x)}}{x} - 1\right) = \frac{\pi}{2} + \arcsin\sqrt{\frac{x}{4}}.$$

Hence $-G_{\mathrm{w}^{(2)}}$ has the exponential representation (2.19) with

$$\xi(x) = \left(\frac{1}{2} + \frac{1}{\pi}\arcsin\sqrt{\frac{x}{4}}\right)\mathbf{1}_{(0,4)}(x) + \mathbf{1}_{(4,\infty)}(x).$$

2.3 Transforms in non-classical independencies

2.3.1 Reciprocal of a Cauchy transform

The reciprocal of the Cauchy transform is an important tool in the study of convolutions with respect to the non-classical independencies. In particular, the composition of two reciprocal Cauchy transforms gives the convolution with respect to monotone and anti-monotone independencies. It also plays a key role in the study of additive convolutions with respect to free and Boolean independencies. In this section we present its main properties.

Definition 12 *Let μ be a finite measure on $\mathbb{R}$ with Cauchy transform G_μ. We define the application $F_\mu : \mathbb{C}^+ \to \mathbb{C}^+$ by*

$$F_\mu(z) = \frac{1}{G_\mu(z)}, \quad z \in \mathbb{C}^+ .$$

We observe that F_μ is a Herglotz function and that there is a one-to-one correspondence between finite measures on $\mathbb{R}$ and the reciprocal of a Cauchy transform. Moreover, a characterization of reciprocal Cauchy transforms is possible, similar to Proposition 5. Recall that Γ_α denotes the Herglotz region (2.16).

Proposition 13 *Let $F : \mathbb{C}^+ \to \mathbb{C}^+$ be a Herglotz function. The following four statements are equivalently:*

i) There exists a finite Borel measure μ on $\mathbb{R}$ such that $F_\mu = F$ in $\mathbb{C}^+$.

ii) For each $\alpha > 0$

$$\lim_{|z|\to\infty, z\in\Gamma_\alpha} \frac{F(z)}{z} = \mu(\mathbb{R}).$$

iii) $\lim_{y\to\infty} F(iy)/iy = \mu(\mathbb{R})$. .

iv) There exist $\gamma \in \mathbb{R}$ and a finite Borel measure σ on $\mathbb{R}$ such that

$$F(z) = \gamma + \mu(\mathbb{R})z + \int_{\mathbb{R}} (\frac{1+tz}{t-z})\sigma(dt),\ z \in \mathbb{C}^+.$$

Proof. The equivalences among (i), (ii) and (iii) follow from Proposition 5. Since $F : \mathbb{C}^+ \to \mathbb{C}^+$ is a Herglotz function the representation (2.6) holds with $\psi = \lim_{y\to\infty} F(iy)/iy = \mu(\mathbb{R})$. This proves the equivalence between (iii) and (iv). ■

In particular, when μ is a probability measure we obtain a very useful property of F_μ.

Corollary 14 *For each probability measure μ on $\mathbb{R}$ it holds that*

$$\operatorname{Im}(F_\mu(z)) \geq \operatorname{Im}(z)$$

with equality for all $z \in \mathbb{C}^+$ if and only if μ is a Dirac measure.

The fact that, inside the region Γ_α, $F_\mu(z)$ behaves as z when z is large allows to show that F_μ is univalent in a certain domain. For α and β positive we consider the regions

$$\Gamma_{\alpha,\beta} = \{z = x + iy : y > \beta, |x| < \alpha y\}. \tag{2.28}$$

We say that F_μ has a right inverse F_μ^{-1} in a domain Γ if $F_\mu(F_\mu^{-1}(z)) = z$ for each z in Γ. This right inverse plays a key role in free probability and its existence is guaranteed by the following result.

Lemma 15 *Let μ be a probability measure on $\mathbb{R}$ and let $0 < \varepsilon < \alpha$. There exists $\beta > 0$ such that*

i) The function F_μ is univalent in $\Gamma_{\alpha,\beta}$.

ii) $\Gamma_{\alpha-\varepsilon,\beta(1+\varepsilon)} \subset F_\mu(\Gamma_{\alpha,\beta})$

Proof. By (ii) in Proposition 13 we can chose β large enough such that $|F_\mu(z) - z| < \varepsilon |z|$ for $z \in \Gamma_{\alpha+\varepsilon,\beta(1-\varepsilon)}$. By continuity, the inequality holds also for $z \in \partial\Gamma_{\alpha+\varepsilon,\beta(1-\varepsilon)}$. Let $z_0 \in \Gamma_{\alpha,\beta}$, Then for some $\beta' > \beta$ large enough, the image of the boundary of $\{z \in \Gamma_{\alpha+\varepsilon,\beta(1-\varepsilon)} : |z| < (2+\varepsilon)\beta'\}$ is a curve with winding number 1 around z_0. By analyticity of $F_\mu - z_0$ and the argument principle, $F(z) = z_0$ has exactly one solution in the connected component of $\mathbb{C}\backslash F(\gamma)$ which contains $\{z \in \Gamma_{\alpha,\beta} : |z| < \beta'\}$. Since β' can be chosen arbitrarily large, F is univalent in $\Gamma_{\alpha,\beta}$. A similar argument shows that $\Gamma_{\alpha-\varepsilon,\beta(1+\varepsilon)} \subset F_\mu(\Gamma_{\alpha,\beta})$. ■

In summary we have the following useful result for the reciprocal of a Cauchy transform.

Proposition 16 *Let μ be a probability measure on $\mathbb{R}$. There exists a domain Γ of the form $\Gamma = \cup_{\alpha>0}\Gamma_{\alpha,\beta_\alpha}$ such that F_μ has a right inverse F_μ^{-1} defined in Γ. Moreover*

$$\operatorname{Im}(F_\mu^{-1}(z)) \leq \operatorname{Im}(z), \qquad z \in \Gamma \tag{2.29}$$

and for each $\alpha > 0$

$$\lim_{|z|\to\infty, z\in\Gamma_\alpha} \frac{F_\mu^{-1}(z)}{z} = 1. \tag{2.30}$$

Proof. The existence of the domain $\Gamma = \cup_{\alpha>0}\Gamma_{\alpha,\beta_\alpha}$ follows from the last lemma. Inequality (2.29) is a consequence of Corollary 14, and (2.30) follows from Proposition 13 (ii). ■

2.3.2 Voiculescu transforms

Using F_μ^{-1} we now introduce the Voiculescu transform in the domain $\Gamma = \cup_{\alpha>0}\Gamma_{\alpha,\beta_\alpha}$ of Proposition 16. The key property of the Voiculescu transform is that it linearizes the free additive convolution of Chapter 1 for measures with compact support. We will come back to this point in the next section.

Definition 17 *Let μ be a probability measure on $\mathbb{R}$ with reciprocal Cauchy transform F_μ. The Voiculescu transform $\phi_\mu : \Gamma \to \mathbb{C}^-$ is defined as*

$$\phi_\mu(z) = F_\mu^{-1}(z) - z \qquad z \in \Gamma.$$

A probability measure on $\mathbb{R}$ is uniquely determined by its Voiculescu transform. To see this, suppose μ and μ' are probability measures on $\mathbb{R}$, such that $\phi_\mu = \phi_{\mu'}$, on a region $\Gamma_{\eta,\beta}$. It follows then that also $F_\mu = F_{\mu'}$ on some open subset of $\mathbb{C}^+$, and hence (by analytic continuation), $F_\mu = F_{\mu'}$ on all of $\mathbb{C}^+$. Consequently μ and μ' have the same Cauchy transform, and by the Stieltjes inversion formula (2.14), this means that $\mu = \mu'$.

A Bochner type characterization of the Voiculescu transform is as follows.

Proposition 18 *Let ϕ be an analytic function defined on a region $\Gamma_{\alpha,\beta}$, for some positive numbers α and β. Then the following two assertions are equivalent:*

(i) There exists a probability measure μ on $\mathbb{R}$, such that $\phi(z) = \phi_\mu(z)$ for all z in a domain $\Gamma_{\alpha,\beta'}$, where $\beta' \geq \beta$.

(ii) There exists a number β' greater than or equal to β, such that

(a) $\mathrm{Im}(\phi(z)) \leq 0$ *for all z in $\Gamma_{\alpha,\beta'}$.*

(b) $\phi(z)/z \to 0$, as $|z| \to \infty$, $z \in \Gamma_{\alpha,\beta'}$.

(c) For any positive integer n and any points $z_1, \ldots, z_n$ in $\Gamma_{\alpha,\beta'}$, the $n \times n$ matrix

$$\left[\frac{z_j - \overline{z_k}}{z_j + \phi(z_j) - \overline{z_k} - \overline{\phi(z_k)}}\right]_{1\leq j,k\leq n},$$

is positive semi-definite.

Proof. (i)$\Rightarrow$(ii). Assume $\phi_\mu(z) = \phi(z)$ for all z in a domain $\Gamma_{\alpha,\beta'}$ with $\beta' \geq \beta$. Proposition 16 $\phi_\mu(z)$ implies (a) and (b). By (a) and the definition of ϕ_μ, there is a Herglotz function F_μ such that $F_\mu^{-1}(z) = \phi_\mu(z) + z \in \mathbb{C}^+$, for all $z \in \Gamma_{\alpha,\beta'}$. So let $z_1, \ldots, z_n$ in $\Gamma_{\alpha,\beta'}$. Then the matrix

$$\left[\frac{z_j - \overline{z_k}}{z_j + \phi(z_j) - \overline{z_k} - \phi(\overline{z_k})}\right]_{j,k} = \left[\frac{F_\mu(F_\mu^{-1}(z_j)) - F_\mu(\overline{F_\mu^{-1}(z_k)})}{F_\mu^{-1}(z_j) - \overline{F_\mu^{-1}(z_k)}}\right]_{j,k} \tag{2.31}$$

is positive semi-definite by Proposition 3.

It remains to prove that (ii) implies (i). We first observe that $g(z) = \phi(z) + z$ is analytic and satisfies $g(z)/z \to 1$, as $|z| \to \infty$, $z \in \Gamma_{\alpha,\beta'}$. Therefore we can proceed as in Lemma 15 to show that $g(z)$ is univalent in $\Gamma_{\alpha,\beta''}$, and $g^{-1}(z) \in \mathbb{C}^+$

for some $\beta'' > \beta'$. Then the equation $F(\varsigma) = F(z+\phi(z)) = z$ defines an analytic function F, which, by (c), satisfies that the matrices

$$\left[\frac{z_j - \overline{z_k}}{z_j + \phi(z_j) - \overline{z_k} - \phi(\overline{z_k})}\right]_{j,k} = \left[\frac{F(\varsigma_j) - F(\overline{\varsigma_k})}{\varsigma_j - \overline{\varsigma_k}}\right]_{j,k}$$

are positive semi-definite for all $\varsigma_1, \ldots, \varsigma_n \in g^{-1}(\Gamma_{\alpha,\beta''}) \subset \mathbb{C}^+$. Theorem 3 allows F to be extended to a Herglotz function. The fact that $\phi(z)/z \to 0$ is equivalent to $F(z)/z \to 1$ and from Proposition 13, $F = F_\mu$ for a probability measure μ in $\mathbb{R}$. Finally, F_μ is univalued for some $\Gamma_{\alpha,\beta'''}$, and $F_\mu(z + \phi(z)) = z = F_\mu(z + \phi_\mu(z))$, so $\phi_\mu = \phi$ in $\Gamma_{\alpha,\beta'''}$. ■

Similar to the Fourier transform, we have a Lévy type continuity theorem for the Voiculescu transform.

Proposition 19 *Let (μ_n) be a sequence of probability measures on $\mathbb{R}$. Then the following assertions are equivalent:*

(a) The sequence (μ_n) converges weakly to a probability measure μ on $\mathbb{R}$.

(b) There exist positive numbers α and β, and a function ϕ, such that all the functions ϕ, ϕ_{μ_n} are defined on $\Gamma_{\alpha,\beta}$, and such that

(b1) $\phi_{\mu_n}(z) \to \phi(z)$, as $n \to \infty$, uniformly on compact subsets of $\Gamma_{\alpha,\beta}$,

(b2) $\sup_{n\in\mathbb{N}} \left|\frac{\phi_{\mu_n}(z)}{z}\right| \to 0$, as $|z| \to \infty$, $z \in \Gamma_{\alpha,\beta}$.

(c) There exist positive numbers α and β, such that all the functions ϕ_{μ_n} are defined on $\Gamma_{\alpha,\beta}$, and such that

(c1) $\lim_{n\to\infty} \phi_{\mu_n}(\mathrm{i}y)$ exists for all y in $[\beta, \infty[$.

(c2) $\sup_{n\in\mathbb{N}} \left|\frac{\phi_{\mu_n}(\mathrm{i}y)}{y}\right| \to 0$, as $y \to \infty$.

If the conditions (a),(b) and (c) are satisfied, then $\phi = \phi_\mu$ on $\Gamma_{\alpha,\beta}$.

Proof. Assume that $\{\mu_n\}_{n=1}^\infty$ converges in distribution to μ. Then, it is not difficult to see from Proposition 13 that $\lim_{y\to\infty} F_{\mu_n}(z) = F_\mu(z)$ uniformly on compact subsets of $\mathbb{C}^+$ and

$$F_{\mu_n}(z) = z(1 + o(1)),\ F_\mu(z) = z(1 + o(1)) \text{ as } |z| \to \infty,\ z \in \Gamma_\alpha$$

uniformly on n. Then, it follows that there exist constants $\alpha, \beta > 0$ such that the Voiculescu transforms ϕ_μ and ϕ_{μ_n} are well defined and have negative imaginary part in $\Gamma_{\alpha,\beta}$. Moreover, $\phi_{\mu_n} = o(z)$ uniformly in n as $|z| \to \infty$, $z \in \Gamma_{\alpha,\beta}$. The family $\{\phi_{\mu_n}\}_{n=1}^\infty$ is normal in $\Gamma_{\alpha,\beta}$. Thus, it is enough to prove that the limit of a convergent subsequence $\{\phi_{\mu_{n_j}}\}_{j=1}^{\infty}$ is in fact ϕ_μ. Indeed, if $z \in \Gamma_{\alpha,\beta}$ one has $z + \phi(z) \in \mathbb{C}^+$ and

$$\begin{aligned}|F_\mu(z+\phi(z)) - z| &= \left|F_\mu(z+\phi(z)) - F_{\mu_{n_j}}(z+\phi_{\mu_{n_j}}(z))\right| \\ &\leq \left|F_\mu(z+\phi(z)) - F_\mu(z+\phi_{\mu_{n_j}}(z))\right| \\ &+ \left|F_\mu(z+\phi_{\mu_{n_j}}(z)) - F_{\mu_{n_j}}(z+\phi_{\mu_{n_j}}(z))\right|.\end{aligned}$$

Since $F_{\mu_{n_j}}$ converges to F_μ uniformly in a neighborhood of $z + \phi(z)$, we have

$$\lim_{j\to\infty} \left|F_\mu(z+\phi_{\mu_{n_j}}(z)) - F_{\mu_{n_j}}(z+\phi_{\mu_{n_j}}(z))\right| = 0.$$

Then $z + \phi(z) = F^{-1}(z)$ in $\Gamma_{\alpha,\beta}$ and therefore $\phi = \phi_\mu$.

Assume now that (ii) holds. From the first part of the proof it is enough to prove the convergence of $\{\mu_n\}_{n=1}^{\infty}$. We have $F_{\mu_n}^{-1}(z) = z + \phi_{\mu_n}(z) = z(1 + o(1))$ uniformly in n as $|z| \to \infty$, $z \in \Gamma_{\alpha,\beta}$. Then, in particular, $iyG_{\mu_n}(iy) - 1 = o(1)$ uniformly in n as $y \to \infty$. The inequality

$$-\operatorname{Re}(iyG\mu_n(iy) - 1) = \int_{\mathbb{R}} \frac{t^2}{y^2+t^2}\mu_n(dt) \geq \frac{1}{2}\mu_n(\{t : |t| \geq y\})$$

gives the weak convergence of $\{\mu_n\}_{n=1}^{\infty}$. ■

Now we can show that the sum of two Voiculescu transforms is again a Voiculescu transform. We rely on the compactly supported case which was mentioned in Chapter 1.

Theorem 20 *Let ϕ_{μ_1} and ϕ_{μ_2} be Voiculescu transforms of two probability measures μ_1 and μ_2 on $\mathbb{R}$. Then, $\phi = \phi_{\mu_1} + \phi_{\mu_2}$ is the Voiculescu transform of a probability measure μ in $\mathbb{R}$.*

Proof. We consider sequences of compactly supported probability measures $(\mu_n^{(i)})_{n\in\mathbb{N}}$ converging in distribution to μ_i, $i = 1, 2$. From the compactly supported case, there are probability measures $(\mu_n)_{n\in\mathbb{N}}$ such that $\phi_{\mu_n} = \phi_{\mu_n^{(1)}} + \phi_{\mu_n^{(2)}}$. By Proposition 19 (b) there is a domain $\Gamma_{\eta,M}$ where each $\phi_{\mu_n^{(i)}} \to \phi_{\mu_i}$ uniformly on compact subsets of $\Gamma_{\eta,M}$. Hence $(\phi_{\mu_n^{(1)}} + \phi_{\mu_n^{(2)}}) \to \phi$ uniformly on compact subsets $\Gamma_{\eta,M}$. Thus, by Proposition 19 (a) there exist a probability measure μ such that $\phi = \phi_\mu$. ■

R-transform and the free cumulant transform

There are some variants of the Voiculescu transform that are useful when dealing with free additive convolutions. One of them it the R_μ-transform defined by

$$R_\mu(z) = \phi_\mu(\frac{1}{z}) = F_\mu^{-1}\left(z^{-1}\right) - \frac{1}{z} \qquad z^{-1} \in \Gamma. \tag{2.32}$$

We then have the following useful relation between the Cauchy transform and the $R_\mu-$transform

$$G_\mu(R_\mu(z) + \frac{1}{z}) = z. \tag{2.33}$$

Another variant is the free cumulant transform $\mathcal{C}_\mu^{\boxplus}$ given by

$$\mathcal{C}_\mu^{\boxplus}(z) = z\phi_\mu(\frac{1}{z}) = zF_\mu^{-1}\left(z^{-1}\right) - 1, \qquad z^{-1} \in \Gamma. \tag{2.34}$$

This will be useful in Chapter 3 to describe the Lévy-Khintchine representation in terms of triples, in analogy to the classical case.

2.3.3 The self-energy

For a probability measure μ on $\mathbb{R}$ we define its *self-energy* transform K_μ by

$$K_\mu(z) = z - \frac{1}{G_\mu(z)} = z - F_\mu(z), \qquad z \in \mathbb{C}^+. \tag{2.35}$$

By Corollary 14 $\mathrm{Im}(F_\mu(z)) \leq \mathsf{Im}(z)$, and hence K_μ maps $\mathbb{C}^+$ into $\mathbb{C}^- \cup \mathbb{R}$. The following characterization follows from the theory of Herglotz functions in Section 2.1

Theorem 21 *For any function $K\colon \mathbb{C}^+ \to \mathbb{C}^- \cup \mathbb{R}$ the following two statements are equivalent:*

(a) $K = K_\mu$ for some probability measure μ on $\mathbb{R}$.

(b) There exists a real constant γ and a finite measure σ on $\mathbb{R}$, such that

$$K(z) = \gamma + \int_{\mathbb{R}} \frac{1+tz}{z-t}\,\sigma(\mathrm{d}t), \qquad z \in \mathbb{C}^+. \tag{2.36}$$

As for the Voiculescu transform ϕ_μ, there is a variant of the self-energy K_μ that we will use often, called the *Boolean cumulant transform* defined as

$$\mathcal{C}_\mu^{\uplus}(z) = zK_\mu(1/z). \tag{2.37}$$

2.3.4 Examples

In this subsection we compute several transforms for some of the distributions in Examples 2.2.3.

2.3.4.1 Symmetric Bernoulli Distribution (continuation)

We have seen in Example 2.2.3 that the symmetric Bernoulli distribution $\beta_\sigma = \frac{1}{2}(\delta_{-\sigma} + \delta_\sigma)$ has Cauchy transform $G_{\beta_\sigma}(z) = z/(z^2 - \sigma^2)$. Then

$$F_{\beta_\sigma}(z) = \frac{z^2 - \sigma^2}{z}, \tag{2.38}$$

$$K_\mu(z) = z - \frac{z^2 - \sigma^2}{z} = \frac{\sigma^2}{z} \tag{2.39}$$

and

$$\mathcal{C}_\mu^{\uplus}(z) = \sigma^2. \tag{2.40}$$

2.3.4.2 Semicircle distribution (continuation)

We have seen in Example 2.2.3 that the semicircle distribution $\mathrm{w}_{0,2}$ with density $w_{0,2}(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \cdot 1_{[-2,2]}(x)$ has Cauchy transform $G_\mu(z) = (z - \sqrt{z^2 - 4})/2$. Then

$$F_{\mathrm{w}_{0,2}}(z) = \frac{1}{G_{\mathrm{w}_{0,2}}(z)} = \frac{2}{z - \sqrt{z^2 - 4}}, \quad z \in \mathbb{C}^+, \tag{2.41}$$

and

$$F^{-1}_{\mathrm{w}_{0,2}}(z) = z + 1/z, \quad z \in \mathbb{C}^+.$$

Moreover, from the definition of the Voiculescu transform and (2.34) we easily see that the Voiculescu and free cumulant transforms are given, respectively, by

$$\phi_{\mathrm{w}_{0,2}}(z) = 1/z, \quad z \in \mathbb{C}^+ \tag{2.42}$$

and

$$\mathcal{C}^{\boxplus}_{\mathrm{w}_{0,2}}(z) = z^2, \quad z \in \mathbb{C}^+. \tag{2.43}$$

It is worthy to observe that $\phi_{\mathrm{w}_{0,2}}(z)$ is an analytic function on $\mathbb{C}^+$.

More generally, for $m \in \mathbb{R}$ and $\sigma^2 > 0$, the Cauchy transform of $\mathrm{w}_{m,\sigma}$ is given by

$$G_{\mathrm{w}_{m,\sigma}}(z) = \frac{2}{r^2}(z - \sqrt{(z-m)^2 - r^2}), \quad z \in \mathbb{C}^+,$$

while its free cumulants transform is given by

$$\mathcal{C}^{\boxplus}_{\mathrm{w}_{m,\sigma}}(z) = mz + \frac{r^2 z^2}{4} = mz + \sigma^2 z^2, \quad z \in \mathbb{C}^+ \tag{2.44}$$

where $r = 2\sigma$ is the radius of the support.

2.3.4.3 Arcsine distribution (continuation)

From Example 2.2.3 the arcsine distribution $\mathrm{a}_{0,2}$ has Cauchy transform $G_\mu(z) = 1/\sqrt{z^2-4}$. Then its reciprocal Cauchy transform is

$$F_{\mathrm{a}_{0,2}}(z) = \frac{1}{G_{\mathrm{a}_{0,2}}(z)} = \sqrt{z^2-4}, \quad z \in \mathbb{C}^+ \tag{2.45}$$

and

$$F^{-1}_{\mathrm{a}_{0,2}}(z) = \sqrt{z^2+4}.$$

Then, similar to the case of the semicircle distribution, one can compute the Voiculescu and the free cumulant transforms of $\mathrm{a}_{0,2}$ and they are given by

$$\phi_{\mathrm{a}_{0,2}}(z) = \sqrt{z^2+4} - z \tag{2.46}$$

and

$$\mathcal{C}^{\boxplus}_{\mathrm{a}_{0,2}}(z) = \sqrt{4z^2+1} - 1. \tag{2.47}$$

We observe that $\phi_{\mathrm{a}_{0,2}}(z)$ is not an analytic function on all $\mathbb{C}$.

More generally, for $s > 0$, the Cauchy transform of the arcsine distribution a_s on $(-\sqrt{s}, \sqrt{s})$ is given by

$$G_{\mathrm{a}_s}(z) = (\sqrt{z^2-s})^{-1}. \tag{2.48}$$

2.3.4.4 Marchenko-Pastur distribution (continuation)

The Marchenko-Pastur distribution m_c of general parameter $c > 0$ is a compactly supported distribution given by

$$\mathrm{m}_c(dx) = \begin{cases} (1-c)\delta_0 + \frac{1}{2\pi x}\sqrt{(x-a)(b-x)}1_{[a,b]}(x)dx, & \text{if } 0 \le c \le 1 \\ \frac{1}{2\pi x}\sqrt{(x-a)(b-x)}1_{[a,b]}(x)dx & \text{if } c > 1 \end{cases} \tag{2.49}$$

where $a = (1-\sqrt{c})^2$ and $b = (1+\sqrt{c})^2$. Its Cauchy transform is given by

$$G_{\mathrm{m}_c}(z) = \frac{1}{2} - \frac{\sqrt{(z-a)(z-b)}}{2z} + \frac{1-c}{2z} \tag{2.50}$$

while its free cumulant transform is

$$\mathcal{C}^{\boxplus}_{\mathrm{m}_c}(z) = \frac{cz}{1-z}. \tag{2.51}$$

Indeed, let us compute the distribution m_c starting from the free cumulant transform (2.51). From (2.34) we have

$$1 + \frac{cG_{\mathrm{m}_c}(z)}{1-G_{\mathrm{m}_c}(z)} = zG_{\mathrm{m}_c}(z)$$

and therefore

$$zG_{\mathrm{m}_c}(z)^2 - (z-c+1)G_{\mathrm{m}_c}(z) + 1 = 0.$$

Solving for $G_{\mathrm{m}_c}(z)$

$$\begin{aligned} G_{\mathrm{m}_c}(z) &= \frac{1}{2z}\left(z-c+\sqrt{(z-(1-\sqrt{c})^2)(z-(1+\sqrt{c})^2)}+1\right) \\ &= \frac{1}{2}+\frac{\sqrt{(z-a)(z-b)}}{2z}+\frac{1-c}{2z},\ z\neq 0. \end{aligned}$$

For $c \leq 1$, $G_{(1-c)\delta_0}(z) = \frac{1-c}{z}$, and for the absolutely continuos part $f_{\mathrm{m}_c}(x)\mathrm{dx}$ we have

$$f_{\mathrm{m}_c}(x) = -\frac{1}{\pi}\lim_{y\to+0}\mathrm{Im}\left(\frac{1}{2}+\frac{\sqrt{(z-a)(z-b)}}{2z}\right) = \frac{1}{\pi}\frac{\sqrt{x\,(4c-8-x)}}{2x} \quad x\in[a,b]. \tag{2.52}$$

Then, from the inversion formula (2.14),

$$G_{\mathrm{m}_c}(dx) = \begin{cases} (1-c)\delta_0 + \frac{1}{2\pi x}\sqrt{(x-a)(b-x)}1_{[a,b]}(x)dx, & 0\leq c\leq 1 \\ \frac{1}{2\pi x}\sqrt{(x-a)(b-x)}\cdot 1_{[a,b]}(x)dx & c>1. \end{cases}$$

2.4 Additive Convolutions: Analytic approach

In this section we define through analytic methods the non-classical additive convolutions corresponding to each of the non-classical independencies. These convolutions coincide with the additive convolutions defined in the algebraic framework of Section 1.2.3 in Chapter 1 for the case of compactly supported probability measures and in Appendix B for the general case of non necessarily bounded probability measures. However, we will not provide proofs of these correspondences. Throughout this monograph we will study non-classical convolutions by combining tools from both the analytical and algebraic point of views. References for the proofs of this equivalence are presented in the Section of Bibliographic Notes at the end of this chapter.

2.4.1 Free convolution

Let $\mathcal{C}^{\boxplus}_{\mu_1}$ and $\mathcal{C}^{\boxplus}_{\mu_2}$ be the free cumulant transforms of μ_1 and μ_2 defined in the domains $\Gamma_{\alpha_1,\beta_1}$ and $\Gamma_{\alpha_2,\beta_2}$, respectively. By (2.34) and Theorem 20 we know that there exists a probability μ on $\mathbb{R}$ such that $\mathcal{C}^{\boxplus}_{\mu} = \mathcal{C}^{\boxplus}_{\mu_1} + \mathcal{C}^{\boxplus}_{\mu_2}$ in $\Gamma_{\alpha_1,\beta_1}\cap\Gamma_{\alpha_2,\beta_2}$. This measure is unique by the uniqueness of the Voiculescu transform. Then the following definition makes sense.

Definition 22 *Let μ_1 and μ_2 be probability measures on $\mathbb{R}$. The free convolution of μ_1 and μ_2 is defined as the unique probability measure $\mu_1 \boxplus \mu_2$ on $\mathbb{R}$ such that*

$$\mathcal{C}^{\boxplus}_{\mu_1\boxplus\mu_2}(z) = \mathcal{C}^{\boxplus}_{\mu_1}(z) + \mathcal{C}^{\boxplus}_{\mu_2}(z), \qquad z^{-1}\in\Gamma_{\alpha_1,\beta_1}\cap\Gamma_{\alpha_2,\beta_2}. \tag{2.53}$$

In terms of the Voiculescu transform, condition (2.53) is equivalent to

$$\phi_{\mu_1\boxplus\mu_2}(z) = \phi_{\mu_1}(z) + \phi_{\mu_2}(z) \qquad z\in\Gamma_{\alpha_1,\beta_1}\cap\Gamma_{\alpha_2,\beta_2} \tag{2.54}$$

and also to

$$R_{\mu_1 \boxplus \mu_2}(z) = R_{\mu_1}(z) + R_{\mu_2}(z) \qquad z^{-1} \in \Gamma_{\alpha_1,\beta_1} \cap \Gamma_{\alpha_2,\beta_2}. \tag{2.55}$$

Free convolution corresponds to the distribution of sums of free independent random variables. More precisely, if X and Y are free independent random variables in some non-commutative probability space $(\tau, \mathcal{A})$, then $\mu_X \boxplus \mu_Y = \mu_{X+Y}$.

As in the case of the classical convolution $*$ it is easy to see that the free convolution operation $\boxplus$ is commutative and associative, since addition of functions is. Also, if μ is a probability measure on $\mathbb{R}$ and δ_x is a Dirac probability measure, $\mu \boxplus \delta_x$ is the translation of μ by x, that is $\mu \boxplus \delta_x(A) = \mu(A - x)$. Therefore, $\mu \boxplus \delta_x = \mu * \delta_x$.

Furthermore, for a constant $c \neq 0$ the dilation of a probability measure μ on $\mathbb{R}$ by c is the probability measure $D_c\mu$ on $\mathbb{R}$ such that $D_c\mu(A) = \mu(c^{-1}A)$. We observe that the free cumulant transform behaves, with respect to the dilation D_c, as the classical cumulant transform, that is $\mathcal{C}^{\boxplus}_{D_c\mu}(z) = \mathcal{C}^{\boxplus}_{\mu}(cz)$, for any probability measure μ on $\mathbb{R}$ and any constant $c \neq 0$.

On the other hand, some aspects of free convolution are different to what is found in classical convolution. For example, if μ_1 and μ_2 are probability measures on $\mathbb{R}$ with compact support, for the classical convolution $supp(\mu_1 * \mu_2) = supp(\mu_1) + supp(\mu_2)$, while if a, b, c are the maxima of the support of μ_1, μ_2 and $\mu_1 \boxplus \mu_2$, respectively, then $c = a + b$ if $\mu_1(\{a\}) + \mu_2(\{b\}) \geq 1$, while $c < a + b$ if $\mu_1(\{a\}) + \mu_2(\{b\}) < 1$.

Also, in contrast with classical convolution, the free convolution of two atomic measures can be absolutely continuous. Example 2.4.4 shows that the free convolution of the symmetric Bernoulli distribution $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$ with itself is the arcsine distribution $\mathrm{a}_{0,2}$. This shows furthermore that, contrary to the classical convolution case, the operation $\boxplus$ is not distributive with respect to the convex sum of probability measures. Indeed, if $\boxplus$ were distributive, with $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$, we would have

$$\begin{aligned} \mu \boxplus \mu = \mu \boxplus \frac{1}{2}(\delta_{-1} + \delta_1) &= \frac{1}{2}(\mu \boxplus \delta_{-1}) + \frac{1}{2}(\mu \boxplus \delta_1) \\ &= \frac{1}{2}(\frac{1}{2}(\delta_{-1} + \delta_1) \boxplus \delta_{-1}) + \frac{1}{2}(\frac{1}{2}(\delta_{-1} + \delta_1) \boxplus \delta_1) \\ &= \frac{1}{4}\delta_{-2} + \frac{1}{2}\delta_0 + \frac{1}{4}\delta_2, \end{aligned}$$

which is a contradiction since

$$\frac{1}{4}\delta_{-2} + \frac{1}{2}\delta_0 + \frac{1}{4}\delta_2 \neq \mathrm{a}_{0,2}$$

2.4.2 Boolean convolution

Recall from Section 2.3.3 that the self-energy is given by $K_\mu(z) = z - F_\mu(z), z \in \mathbb{C}^+$.

Definition 23 *The Boolean convolution of two probability measures μ_1 and μ_2 on $\mathbb{R}$ is defined as the unique probability measure $\mu_1 \uplus \mu_2$ on $\mathbb{R}$ such that*

$$K_{\mu_1 \uplus \mu_2}(z) = K_{\mu_{1}(z)} + K_{\mu_2}(z), \qquad z \in \mathbb{C}^+. \tag{2.56}$$

In terms of the Boolean cumulant transform (2.40), this is equivalent to say that

$$\mathcal{C}^{\uplus}_{\mu_1 \uplus \mu_2}(z) = \mathcal{C}^{\uplus}_{\mu_1}(z) + \mathcal{C}^{\uplus}_{\mu_2}(z), \qquad z^{-1} \in \mathbb{C}^+. \tag{2.57}$$

The Boolean convolution is associative and commutative, i.e. $(\mu_1 \uplus \mu_2) \uplus \mu_3 = \mu_1 \uplus (\mu_2 \uplus \mu_3)$ for all probability measures μ_1, μ_2, μ_3.

Boolean convolution corresponds to the distribution of the sum of Boolean independent random variables. More precisely, let X, Y be Boolean independent random variables in some non-commutative probability space $(\tau, \mathcal{A})$ then

$$\mu_X \uplus \mu_Y = \mu_{X+Y}. \tag{2.58}$$

Different to the free and classical cases, for the Boolean convolution a Dirac mass does not shift the measure, as shown by the following example. This may be explained by the fact that scalars are not well behaved under Boolean independence.

Example 24 *Let $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. Then $\mu \uplus \delta_1 \neq \mu * \delta_1$.*

$$G_\mu(z) = \frac{1}{2}(\frac{1}{z-1} + \frac{1}{z+1}) = \frac{z}{z^2-1}. \tag{2.59}$$

from where

$$F_\mu(z) = \frac{z^2-1}{z}. \tag{2.60}$$

Hence

$$K_\mu(z) = z - \frac{z^2-1}{z} = \frac{1}{z}.$$

Since $F_{\delta_1}(z) = z - 1$ and $K_{\delta_1}(z) = -1$, we have $K_{\mu \uplus \delta_1}(z) = \frac{1}{z} - 1 = \frac{1-z}{z}$.

Since $\mu * \delta_1 = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$,

$$F_{\mu * \delta_1}(z) = \frac{1}{2}(\frac{1}{z} + \frac{1}{z-2}) = \frac{z+1}{(z)(z-1)}$$

2.4.3 Monotone convolution

Definition 25 *The monotone convolution of two probability measures μ_1 and μ_2 on $\mathbb{R}$ is defined as the unique probability measure $\mu_1 \triangleright \mu_2$ on $\mathbb{R}$ such that*

$$F_{\mu_1 \triangleright \mu_2}(z) = F_{\mu_1}(F_{\mu_2}(z)) \qquad z \in \mathbb{C}^+.$$

By the uniqueness of the reciprocal Cauchy transform F_μ, monotone convolution is well defined. Moreover, since composition of functions is associative, then the monotone convolution is associative too. That is, $(\mu_1 \triangleright \mu_2) \triangleright \mu_3 = \mu_1 \triangleright (\mu_2 \triangleright \mu_3)$ for all probability measures μ_1, μ_2, μ_3. It is not, however, commutative, i.e. in general $\mu_1 \triangleright \mu_2 \neq \mu_2 \triangleright \mu_1$ as Example 26 below shows.

Monotone convolution corresponds to the sum of monotone independent random variables. Namely, if X and Y are monotone independent non-commutative random variables in some non-commutative probability space $(\tau, \mathcal{A})$ then $\mu_X \triangleright \mu_Y = \mu_{X+Y}$.

Example 26 *Let $\mu_1 = a_{0,2}$ and $\mu_2 = \delta_2$. Then $\mu_1 \triangleright \mu_2 \neq \mu_2 \triangleright \mu_1$. Indeed, on one hand, from (2.45) the reciprocal of the Cauchy transform of the arcsine law is $F_{\mu_1}(z) = \sqrt{z^2-4}$. On the other hand, it is easily seen that $F_{\mu_2}(z) = z - 2$. Then*

$$F_{\mu_1 \triangleright \mu_2}(z) = \sqrt{(z-2)^2 - 4} = \sqrt{z^2 - 4z}$$

and

$$F_{\mu_2 \triangleright \mu_1}(z) = \sqrt{(z-2)^2 - 4} = \sqrt{z^2 - 4} - 2.$$

Since the reciprocal Cauchy transforms $F_{\mu_1 \triangleright \mu_2}$ and $F_{\mu_2 \triangleright \mu_1}$ do not coincide, $\mu_1 \triangleright \mu_2 \neq \mu_2 \triangleright \mu_1$.

2.4.4 Examples

2.4.4.1 Free convolution of the symmetric Bernoulli distribution: Arcsine distribution

The free convolution of two atomic measure can be absolutely continuous.

Proposition 27 *Let μ be the symmetric Bernoulli distribution $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Then $\mu \boxplus \mu$ is the arcsine distribution $a_{0,2}$ on $(-2, 2)$.*

Proof. We have seen in Example 2.2.3 that $G_\mu(z) = z/(z^2-1)$. From (2.33) and writing $L_\mu(z) = R_\mu(z) + \frac{1}{z}$ we have $G_\mu(L_\mu(z)) = z$ and therefore

$$(L_\mu(z))^2 - \frac{(L_\mu(z))}{z} = 1.$$

Solving for $L_\mu(z)$ we have the two solutions

$$L_\mu(z) = \frac{1 \pm \sqrt{1+4z^2}}{2z}.$$

Then the R-transform of μ is

$$R_\mu(z) = \frac{\sqrt{1+4z^2} - 1}{2z}$$

and computing $R_{\mu \boxplus \mu}$ we have

$$\frac{1}{z} + R_{\mu \boxplus \mu}(z) = \frac{1}{z} + 2R_\mu(z) = \frac{\sqrt{1+4z^2}}{z}$$

and

$$\frac{\sqrt{1+4(G_{\mu\boxplus\mu}(z))^2}}{G_{\mu\boxplus\mu}(z)} = z.$$

Therefore

$$G_{\mu\boxplus\mu}(z) = \frac{1}{\sqrt{z^2-4}}, \quad z \in \mathbb{C}^+, \tag{2.61}$$

which according to Example 2.2.3 is the Cauchy transform of the arcsine distribution on $(-2, 2)$. ■

2.4.4.2 Boolean convolution of the symmetric Bernoulli distribution

Let β_{σ_1} and β_{σ_2} be two symmetric Bernoulli distributions with parameters σ_1 and σ_2. Then $\beta_{\sigma_1} \uplus \beta_{\sigma_2} = \beta_{\sigma_1+\sigma_2}$. Indeed, from (2.39) $K_{\mu_i}(z) = \frac{\sigma_i}{z}$ and then

$$K_{\beta_{\sigma_1}\uplus\beta_{\sigma_2}}(z) = \frac{\sigma_1+\sigma_2}{z}.$$

More generally we have

Proposition 28 $\beta_{\sigma_1} \uplus \cdots \uplus \beta_{\sigma_n} = \beta_{\sigma_1+\cdots+\sigma_n}$.

2.4.4.3 semicircle distribution (continuation)

From Example 2.2.3 the free cumulant transform of the semicircle distribution $\mathrm{w}_{m,\sigma}$ with zero mean and variance σ^2 is given by

$$\mathcal{C}^{\boxplus}_{\mathrm{w}_{m,\sigma}}(z) = mz + \sigma^2 z^2,\ z \in \mathbb{C}. \tag{2.62}$$

Several important consequences follow from this fact, which resembles the fact that the sum of independent classical Gaussian distribution has the Gaussian distribution with the sum of the means and variances.

Proposition 29 *Let* $\mathrm{w}_{m_1,\sigma_1}, ..., \mathrm{w}_{m_n,\sigma_n}$ *be semicircle distributions with means* $m_1,..., m_n$ *and variances* $\sigma_1^2, ..., \sigma_n^2$*, respectively. Then* $\mathrm{w}_{m,\sigma} = \mathrm{w}_{m_1,\sigma_1} \boxplus \cdots \boxplus \mathrm{w}_{m_n,\sigma_n}$ *has the semicircle distribution with mean* $m_1 + \cdots + m_n$ *and variance* $\sigma_1^2 + \cdots + \sigma_n^2$.

Proof. Using (2.62) we have

$$\begin{aligned}
\mathcal{C}^{\boxplus}_{\mathrm{w}_{m,\sigma}}(z) &= \mathcal{C}^{\boxplus}_{\mathrm{w}_{m_1,\sigma_1}}(z) + \cdots + \mathcal{C}^{\boxplus}_{\mathrm{w}_{m_n,\sigma_n}}(z) = \\
&= (\sigma_1^2 z^2 + m_1 z) + \cdots + (\sigma_n^2 z^2 + m_n z) \\
&= (\sigma_1^2 + \cdots + \sigma_n^2) z^2 + (m_1 + \cdots + m_n) z
\end{aligned}$$

from which the proposition follows by the uniqueness of the free cumulant transform. ■

2.4.4.4 Arcsine distribution (continuation)

Let $\mu_1 = a_{0,\sigma_1}$ and $\mu_2 = a_{0,\sigma_2}$. Then $\mu_1 \triangleright \mu_2 = a_{0,\sqrt{\sigma_1^2+\sigma_2^2}}$. Indeed, from (2.45) the reciprocal Cauchy transform of the arcsine law with parameter σ_i is given by $F_{\mu_i}(z) = \sqrt{z^2 - \sigma_i^2}$. Then,

$$F_{\mu_1 \triangleright \mu_2} = \sqrt{(\sqrt{z^2 - \sigma_2^2})^2 - \sigma_1^2} = \sqrt{z^2 - (\sigma_2^2 + \sigma_1^2)}$$

More generally the following results is easily proved.

Proposition 30 *Let* $\mathrm{a}_{0,\sigma_i}, i = 1, ...n$ *be symmetric arcsine distributions on* $(-\sigma_i, \sigma_i)$, $i = 1, ..., n$. *Then* $\mathrm{a}_{0,\sigma_1} \triangleright \cdots \triangleright \mathrm{a}_{0,\sigma_n}$ *has the semicircle distribution* $\mathrm{a}_{0,\sqrt{\sigma_1^2+\cdots+\sigma_n^2}}$.

2.4.4.5 Free Poisson distribution (continuation)

From the free cumulant transform (2.51) we easily see that the free convolution of Marchenko-Pastur distributions is again a Marchenko-Pastur distribution. More specifically,

Proposition 31 *Let* $\mathrm{m}_{c_1}, ..., \mathrm{m}_{c_n}$ *be Marchenko-Pastur distributions with parameters* $c_1, ..., c_n$ *respectively. Then* $\mathrm{m}_c = \mathrm{m}_{c_1} \boxplus \cdots \boxplus \mathrm{m}_{c_n}$ *has the Marchenko-Pastur distribution with parameter* $c_1 + \cdots + c_n$.

2.5 Free multiplicative convolutions: Analytic approach

2.5.1 The S-transform

The free product $\boxtimes$ of two random variables was considered in Chapter 1. In this section we present the main elements of the analytic approach to the free product of probability measures. In this context, it is useful to consider another analytic tool called the S-transform. We briefly indicate the main results of the subject without proofs.

For a probability measure μ on $\mathbb{R}$ define de moment generating function $\Psi_\mu : \mathbb{C}\backslash\mathbb{R} \to \mathbb{C}$ as

$$\Psi_\mu(z) = \int_{\mathbb{R}} \frac{zx}{1 - zx} \mu(\mathrm{d}x) = z^{-1} G_\mu(z^{-1}) - 1, \quad z \in \mathbb{C}\backslash\mathbb{R}. \tag{2.63}$$

Using Proposition 8 one can easily see that if μ is a symmetric probability measure $\Psi_\mu(z) = \Psi_{\mu^2}(z^2)$, $z \in \mathbb{C}\backslash\mathbb{R}_+$.

On the other hand, if μ has compact support with moments $m_n(\mu)$, $n \geq 1$, we obtain the power series expansion

$$\Psi_\mu(z) = \sum_{n=1}^{\infty} m_n(\mu) z^n. \tag{2.64}$$

When $m_1(\mu) \neq 0$, the inverse $\chi_\mu(z)$ of $\Psi_\mu(z)$ exists and is unique as a formal power series in z. In this case, the S-transform is defined as

$$S_\mu(z) = \chi_\mu(z)\frac{1+z}{z}. \tag{2.65}$$

Since $S_\mu(z)$ is obtained from the Cauchy transform G_μ and the inverse $\chi_\mu(z)$ is unique, then there is a one-to-one correspondence between non-zero mean probability measures with compact support and the corresponding S-transforms.

When $m_1(\mu) = 0$, the main problem is that $\Psi_\mu(z)$ does not have a unique inverse. Indeed, if $\mu \neq \delta_0$ has zero mean, then $m_2(\mu) > 0$ and the series (2.64) starts with $m_2(\mu)z^2$. This means that $\Psi_\mu(z)$ cannot be invertible by a power series in z but in $\sqrt{z}$. Then there are two inverses corresponding to the branches of $\sqrt{z}$. In other words, there are two power series χ and $\widetilde{\chi}$ in $\sqrt{z}$ that satisfy $\Psi_\mu(\chi_\mu(z)) = z$.

On the other hand, for probability measures with unbounded support on $\mathbb{R}_+$ and such that $\mu(\{0\}) < 1$, the function $\Psi_\mu(z)$ has a unique inverse $\chi_\mu(z)$ in the left-half plane $i\mathbb{C}^+$ and $\Psi_\mu(i\mathbb{C}^+)$ is a region contained in the circle with diameter $(\mu(\{0\}) - 1, 0)$. In this case the S-transform of μ is defined also as in (2.65). It satisfies $z = \mathcal{C}_\mu^{\boxplus}(zS_\mu(z))$ for sufficiently small $z \in \Psi_\mu(i\mathbb{C}^+)$.

The definition of S-transform can be extended to symmetric probability measures μ on $\mathbb{R}$ as follows. Let $H = \{z \in \mathbb{C}^-; \quad |\mathrm{Re}(z)| < |\mathrm{Im}(z)|\}$ and $\widetilde{H} = \{z \in \mathbb{C}^+; \quad |\mathrm{Re}(z)| < \mathrm{Im}(z)\}$. When $\mu(\{0\}) < 1$, the transform Ψ_μ has a unique inverse on H, $\chi_\mu : \Psi_\mu(H) \to H$ and a unique inverse on $\widetilde{H}$, $\widetilde{\chi}_\mu : \Psi_\mu(\widetilde{H}) \to \widetilde{H}$. In this case there are two S-transforms for μ given by

$$S_\mu(z) = \chi_\mu(z)\frac{1+z}{z} \text{ and } \widetilde{S}_\mu(z) = \widetilde{\chi}_\mu(z)\frac{1+z}{z} \tag{2.66}$$

and these are such that

$$S_\mu^2(z) = \frac{1+z}{z}S_{\mu^{(2)}}(z) \text{ and } \widetilde{S}_\mu^2(z) = \frac{1+z}{z}S_{\mu^{(2)}}(z) \tag{2.67}$$

for z in $\Psi_\mu(H)$ and $\Psi_\mu(\widetilde{H})$, respectively. Moreover the following result holds.

Lemma 32 *Assume that μ is a probability measure on $\mathbb{R}_+$ or symmetric on $\mathbb{R}$. For some sufficiently small $\varepsilon > 0$, we have a region D_ε that includes $\{-it; 0 < t < \varepsilon\}$ such that*

$$z = \mathcal{C}_\mu^{\boxplus}(zS_\mu(z)) \tag{2.68}$$

for $z \in D_\varepsilon$.

2.5.2 Free multiplicative convolutions via the S-transform

We now show how to compute the multiplicative convolution of two probability measures on $\mathbb{R}_+$ using the S-transform.

Proposition 33 *Let μ_1 and μ_2 be probability measures on $\mathbb{R}_+$ with $\mu_i \neq \delta_0$, $i = 1, 2$. Then $\mu_1 \boxtimes \mu_2 \neq \delta_0$ and*

$$S_{\mu_1 \boxtimes \mu_2}(z) = S_{\mu_1}(z) S_{\mu_2}(z)$$

in that component of the common domain which contains $(-\varepsilon, 0)$ for small $\varepsilon > 0$. Moreover, $(\mu_1 \boxtimes \mu_2)(\{0\}) = \max\{\mu_1(\{0\}), \mu_2(\{0\})\}$.

Multiplicative convolution is closed under weak convergence. More precisely,

Proposition 34 *Let $\{\mu_n\}_{n=1}^{\infty}$ and $\{\nu_n\}_{n=1}^{\infty}$ be sequences of probability measures on $\mathbb{R}_+$ converging to probability measures μ and ν on $\mathbb{R}_+$, respectively, in the weak* topology and such that $\mu \neq \delta_0 \neq \nu$. Then, the sequence $\{\mu_n \boxtimes \nu_n\}_{n=1}^{\infty}$ converges to $\mu \boxtimes \nu$ in the weak* topology.*

On the other hand, the free multiplicative convolution $\mu_1 \boxtimes \mu_2$ of a probability measure μ_1 supported on $\mathbb{R}_+$ with a symmetric probability measure μ_2 on $\mathbb{R}$ can be computed via the S-transform as follows

$$S_{\mu_1 \boxtimes \mu_2}(z) = S_{\mu_1}(z) S_{\mu_2}(z). \tag{2.69}$$

Theorem 35 *Let μ and ν be probability measures on $\mathbb{R}$ such that μ is symmetric, $\nu \in \mathcal{M}^+$ and $\mu \neq \delta_0 \neq \nu$. Let S_μ and $\widetilde{S}_\mu$ be the two S-transforms of μ. Then*

$$S_{\mu \boxtimes \nu}(z) = S_\mu(z) S_\nu(z) \text{ and } \widetilde{S}_{\mu \boxtimes \nu}(z) = \widetilde{S}_\mu(z) S_\nu(z) \tag{2.70}$$

are the two S-transforms of the symmetric probability measure $\mu \boxtimes \nu$, where the functions in (2.70) are considered in the common domain which contains $(-\varepsilon, 0)$ for small $\varepsilon > 0$.

The key in proving the above theorem is the following lemma, which is a result of independent interest.

Lemma 36 *Let μ and ν be probability measures on $\mathbb{R}$ such that μ is symmetric, $\nu \in \mathcal{M}^+$ and $\mu \neq \delta_0 \neq \nu$. Then*

$$\mu_1 \boxtimes \mu_2^{(2)} \boxtimes \mu_1 = (\mu_1 \boxtimes \mu_2)^{(2)}. \tag{2.71}$$

Free additive powers of a probability measure μ on $\mathbb{R}_+$ may also be described by the S-transform in the following way

$$S_{\mu^{\boxplus t}}(z) = \frac{1}{t} S_\mu(z/t), \tag{2.72}$$

while the S transform of a dilation is given by

$$S_{D_t(\mu)}(z) = \frac{1}{t} S_\mu(z). \tag{2.73}$$

Then we can deduce the following important equality,

$$(\mu \boxtimes \nu)^{\boxplus t} = D_t(\mu^{\boxplus t} \boxtimes \nu^{\boxplus t}) \qquad t > 1. \tag{2.74}$$

The fact that the free powers $\mu^{\boxplus t}$, $t > 1$ exist is proved as follows. We first prove that for μ with support on the positive real line μ we have that

$$\mu \boxtimes \left((1 - \frac{1}{t})\delta_0 + \frac{1}{t}\delta_t \right) = (1 - \frac{1}{t})\delta_0 + \frac{1}{t}\mu^{\boxplus t}.$$

Since the LHS is a well defined probability measure then the RHS is also a probability measure. Moreover, by the second part of Proposition 33 $\mu \boxtimes \left((1-\frac{1}{t})\delta_0 + \frac{1}{t}\delta_t\right)$ has an atom of size at least $1-t$ at 0 and then $\mu^{\boxplus t}$ is a probability measure.

Let $\nu = (1-\frac{1}{t})\delta_0 + \frac{1}{t}\mu^{\boxplus t}$ then $G_\nu(z) = \frac{t-1}{tz} + \frac{1}{t}G_{\mu^{\boxplus t}}(z)$, from where

$$\begin{aligned}
\Psi_\nu(z) &= 1/zG_\nu(1/z) - 1 \\
&= \frac{1}{z}\left(\frac{t-1}{t(1/z)} + \frac{1}{t}G_{\mu^{\boxplus t}}(1/z)\right) - 1 \\
&= \frac{t-1}{t} + \frac{1}{tz}G_{\mu^{\boxplus t}}(1/z) - 1 \\
&= (\frac{1}{t}\left(zG_{\mu^{\boxplus t}}(z) - 1\right) = \frac{1}{t}\Psi_{\mu^{\boxplus t}}(z).
\end{aligned}$$

This implies that $\chi_\nu(z) = \chi_{\mu^{\boxplus t}}(tz)$ and hence

$$\begin{aligned}
S_\nu(z) &= \frac{1+z}{z}\chi_{\mu^{\boxplus t}}(tz) = \frac{1+z}{1+tz}tS_{\mu^{\boxplus t}}(tz) \\
&= \frac{1+z}{1+tz}S_\mu(z) = S_\rho(z)S_\mu(z)
\end{aligned}$$

where $\rho = (1-\frac{1}{t})\delta_0 + \frac{1}{t}\delta_t$, as desired.

2.5.3 Examples

The S-transform of the Gaussian distributions associated to the four non-classical independencies can be straightforwardly computed from previous considerations. We leave the details to the readers. For the Wigner measure $\mathrm{w}_{0,a}$ with zero mean and variance a

$$S_{\mathrm{w_{0,a}}}(z) = \sqrt{\frac{1}{az}} \tag{2.75}$$

and for "its square", the Marchenko-Pastur distribution $\mathrm{m_c}$ with parameter $c > 0$

$$S_{\mathrm{m_c}}(z) = \frac{1}{z+c}. \tag{2.76}$$

For the arcsine distribution $(-\sqrt{s}, \sqrt{s})$

$$S_{\mathrm{a}_s}(z) = \sqrt{\frac{z+2}{sz}}, \tag{2.77}$$

while for the positive arcsine distribution on $(0, s)$

$$\mathrm{a}_s^+(\mathrm{d}x) = \frac{1}{\pi}\frac{1}{\sqrt{x(s-x)}}1_{(0,s)}(x)\mathrm{d}x; \tag{2.78}$$

its S-transform is

$$S_{\mathrm{a}_s^+}(z) = \frac{z+2}{s(z+1)}. \tag{2.79}$$

2.5.4 A symmetric beta distribution

We end the section with an example of the multiplicative convolution of the arcsine distribution with the Marchenko-Pastur distribution. More specifically, we are interested in identifying the distribution $\lambda_s = \mathrm{a}_s \boxtimes \mathrm{m}_1$.

Using (2.77) and (2.76) in (2.70) we have that $S_{\lambda_s}(z) = S_{\mathrm{a}_c}(z)S_{\mathrm{m}_1}(z)$, i.e.

$$S_{\lambda_s}(z) = \frac{1}{z+1}\sqrt{\frac{z+2}{sz}}.$$

Then, from (2.66)

$$\begin{aligned} zS_{\lambda_s}(z) &= \chi_{\mathrm{a}_s}(z), \\ \Psi_{\mathrm{a}_s}(zS_{\lambda_s}(z)) &= \Psi_{\mathrm{a}_s}(\chi_{\mathrm{a}_s}(z)) = z. \end{aligned}$$

This means by equation (2.68) that

$$\mathcal{C}^{\boxplus}_{\lambda_s}(z) = \Psi_{\mathrm{a}_s}(z). \tag{2.80}$$

We will prove that λ_s has the symmetric beta distribution $SB_s(3/2, 1/2)$ on $(-2\sqrt{s}, 2\sqrt{s})$ with density

$$g(x) = \frac{1}{2\pi\sqrt{s}}\,|x|^{-1/2}\,(2\sqrt{s} - |x|)^{1/2}, \quad |x| < 2\sqrt{s}. \tag{2.81}$$

First, the Cauchy transform of λ_s is given by

$$G_{\lambda_s}(z) = \frac{-1}{\sqrt{2s}}\sqrt{1 - \sqrt{z^{-2}(z^2 - 4s)}}. \tag{2.82}$$

Indeed, using (2.80), (2.63) and (2.48), we have that λ_s is such that

$$G_{\lambda_s}\left(\frac{z^2}{\sqrt{z^2 - s}}\right) = z^{-1}.$$

Making the change of variable $r = z^2/\sqrt{z^2 - s}$, we observe that $r \in \mathbb{C}^+$ when $z \in \mathbb{C}^+$ and that r and z satisfy

$$z^4 - z^2r^2 + r^2s = 0.$$

Solving for z^2, we find $z^2 = \left(r^2 \pm \sqrt{r^2(r^2-4s)}\right)/2$ and hence

$$z = \pm\sqrt{\frac{r^2 \pm \sqrt{r^2(r^2-4s)}}{2}}.$$

Then, the potential candidates for G_{λ_s} are

$$\frac{\sqrt{2}}{\pm\sqrt{z^2 \pm \sqrt{z^2(z^2-4s)}}}.$$

Since, by Proposition 4, G_{λ_s} must be such that $G_{\lambda_s} : \mathbb{C}^+ \to \mathbb{C}^-$ and $|z|\, G_{\lambda_s}(z) \to 1$ when $|z| \to \infty$, we deduce that

$$-G_{\lambda_s}(z) = \frac{\sqrt{2}}{\sqrt{z^2 + \sqrt{z^2(z^2-4s)}}}.$$

Then, multiplying and dividing by $\left(\sqrt{z^2 - (z^2(z^2-4s))^{1/2}}\right)$,

$$\begin{aligned}
-G_{\lambda_s}(z) &= \frac{\sqrt{2}\sqrt{z^2 - \sqrt{z^2(z^2-4s)}}}{\sqrt{z^4 - z^2(z^2-4s)}} \\
&= \frac{\sqrt{2}\sqrt{z^2 - \sqrt{z^2(z^2-4s)}}}{\sqrt{4sz^2}} \\
&= \frac{1}{\sqrt{2s}}\sqrt{1 - \sqrt{z^{-2}(z^2-4s)}}.
\end{aligned}$$

Next, we use the inversion formula (2.15) to show that the density of G_{λ_s} is given by (2.81). We notice that there is an imaginary part when $|x| < 2\sqrt{s}$ and $x \neq 0$. Thus, we are looking for $b < 0$ such that

$$\sqrt{1 - \sqrt{x^{-2}(x^2-4s)}} = a + ib.$$

That is,

$$\begin{aligned}
1 - i\sqrt{x^{-2}(4s-x^2)} &= a^2 - b^2 + 2iab, \\
\sqrt{x^{-2}(4s-x^2)} &= -2ab
\end{aligned}$$

if and only if

$$x^{-2}(4s - x^2) = 4a^2b^2 = 4b^4 + 4b^2.$$

Then, solving for b^2 in the equation

$$b^4 + b^2 - \frac{1}{4}x^{-2}(4s - x^2) = 0$$

we obtain

$$b^2 = \frac{-1 \pm \sqrt{1 + x^{-2}(4s - x^2)}}{2}.$$

Since b^2 is real and nonnegative, we have

$$b^2 = \frac{1}{2} |x|^{-1} (\sqrt{4s} - |x|)$$

and therefore

$$b = -\sqrt{\frac{1}{2} |x|^{-1} (\sqrt{4s} - |x|)},$$

from where we obtain the result.

2.6 Divisibility for convolutions in the 5 universal independencies

The analytic approach to infinite divisibility is the subject of Chapter 3. However, in this section we present a brief introduction and some simple criteria for infinite divisibility with respect to the additive convolutions for the five independencies. These conditions are for distributions determined by moments and are based on properties of the cumulants introduced in Chapter 1, as well as a few other parameters of the distributions.

Although the condition of the existence of moments seems restrictive these conditions are useful, since contrary to the case of classical infinite divisibility some of the most important compactly supported distributions are infinitely divisible with respect to non-classical convolutions.

2.6.1 Definition and first examples

Let $\circledast$ denote, generically, any of the five convolutions $*, \boxplus, \uplus, \triangleright$ and $\triangleleft$ associated to classical, free, Boolean, monotone and anti-monotone independencies, respectively. We write $\circledast \in \{*, \boxplus, \uplus, \triangleright, \triangleleft\}$. A probability measure μ on $\mathbb{R}$ is *n-divisible with respect to the convolution* $\circledast$ if there exists a probability measure μ_n on $\mathbb{R}$ such that

$$\mu = \underbrace{\mu_n \circledast \mu_n \circledast \cdots \circledast \mu_n}_{n \text{ times}}.$$

In the case of the commutative convolutions $*, \boxplus$ and $\uplus$, this is equivalent to saying that there exists a probability measure μ_n such that

$$\mathcal{C}_{\mu}^{\circledast}(z) = \mathcal{C}_{\mu_1}^{\circledast}(z) + ... + \mathcal{C}_{\mu_n}^{\circledast}(z)$$

for z in a suitable domain of definition of $\mathcal{C}_{\mu}^{\circledast}$.

Similar to the definition of infinite divisibility with respect to classical convolution, it is said that a probability measure μ on $\mathbb{R}$ is *infinitely divisible with respect to the convolution* $\circledast$ if for every positive integer n, μ is n-divisible.

The Gaussian distribution for each of the five universal independence is infinitely divisible with respect to the associated convolution. Indeed, it is well known that the classical Gaussian distribution is infinitely divisible with respect to the classical convolution. Proposition 29 gives the infinite divisibility of the semicircle distribution with respect to the free convolution $\boxplus$. Likewise, Proposition 30 gives the infinite divisibility of the arcsine distribution with respect to the monotone $\rhd$ and anti-monotone $\lhd$ convolutions. Finally, the infinite divisibility of the symmetric Bernoulli distribution with respect to Boolean convolution $\uplus$ is given by Proposition 28, but also from the fact that any probability measure on $\mathbb{R}$ is infinitely divisible with respect to the Boolean convolution, as easily seen as follows.

Proposition 37 *Let μ be a probability measure on $\mathbb{R}$. Then μ is infinitely divisible with respect to the Boolean convolution $\uplus$.*

Proof. Given μ, from Theorem 21 there is a finite measure σ on $\mathbb{R}$ and a real constant γ, such that

$$K_\mu(z) = \gamma + \int_{\mathbb{R}} \frac{1+tz}{z-t}\,\sigma(\mathrm{d}t), \qquad z \in \mathbb{C}^+.$$

For n in $\mathbb{N}$, let μ_n be the probability measure on $\mathbb{R}$ such that

$$K_{\mu_n}(z) = n^{-1}\gamma + n^{-1}\int_{\mathbb{R}} \frac{1+tz}{z-t}\,\sigma(\mathrm{d}t), \qquad z \in \mathbb{C}^+.$$

Then for any z in $\mathbb{C}^+$

$$K_{\underbrace{\mu_n \uplus \cdots \uplus_n}_{n \text{ times}}}(z) = \sum_{j=1}^{n} K_{\mu_n}(z) = nK_{\mu_n}(z) = K_\mu(z).$$

By uniqueness of the self-energy and the Cauchy transforms, this means that $\mu_n \uplus \cdots \uplus \mu_n = \mu$, which completes the proof. ∎

2.6.2 Kurtosis and n-divisibility

The kurtosis of a probability distribution is a widely used quantity in statistics and gives information about the shape of a given distribution. Here we derive a simple necessary conditions for n-divisibility with respect to the classical, free, monotone and anti-monotone additive convolutions in terms of the kurtosis. We use the first fourth cumulants with respect to these convolutions, as defined in Chapter 1.

The *classical kurtosis* of a probability measure μ with finite fourth moment is defined as

$$Kurt(\mu) = \frac{c_4(\mu)}{(c_2(\mu))^2} = \frac{\widetilde{m}_4(\mu)}{(\widetilde{m}_2(\mu))^2} - 3,$$

where $c_2(\mu)$ and $c_4(\mu)$ are the second and fourth classical cumulants, and $\widetilde{m}_2(\mu)$ and $\widetilde{m}_4(\mu)$ the second and fourth moments around the mean. It is always true that $Kurt(\mu) \geq -2$.

The kurtosis of the Gaussian distributions with respect to the five universal independencies are easily computed. For the classical Gaussian the kurtosis is zero. For the semicircle distribution $\mathrm{w}_{0,\sigma}$, $Kurt(\mathrm{w}_{0,\sigma}) = -1$ while for the arcsine distribution $\mathrm{a}_{0,\sigma}$, $Kurt(\mathrm{a}_{0,\sigma}) = -1.5$. Finally, for the symmetric Bernoulli distribution its kurtosis is -2.

Proposition 38 *Let μ be a probability measure on $\mathbb{R}$ with finite fourth moment. If μ is n-divisible in the classical sense then $Kurt(\mu) \geq -\frac{2}{n}$.*

Proof. Suppose μ is n-divisible. Let μ_n be such that $\underbrace{\mu_n * \cdots * \mu_n}_{n \text{ times}} = \mu$, by linearity of the cumulants we can see that

$$Kurt(\mu_n) = \frac{\frac{1}{n}c_4(\mu)}{(\frac{1}{n}c_2(\mu))^2} = n\frac{c_4(\mu)}{(c_2(\mu))^2} = nKurt(\mu)$$

So $Kurt(\mu) = \frac{1}{n}Kurt(\mu_n) > -\frac{2}{n}$, where we used the fact that $Kurt \geq -2$. ∎

The free kurtosis is defined similarly using the free cumulants instead of the classical cumulants. That is, the *free kurtosis* of a probability measure μ is defined as

$$Kurt^{\boxplus}(\mu) = \frac{\kappa_4(\mu)}{(\kappa_2(\mu))^2} = \frac{\widetilde{m}_4(\mu)}{(\widetilde{m}_2(\mu))^2} - 2 = Kurt(\mu) + 1$$

where $\kappa_2(\mu)$ and $\kappa_4(\mu)$ are the second and fourth free cumulants. Notice that $Kurt^{\boxplus}(\mu) \geq -1$.

Using similar arguments as in Proposition 38, we obtain a sufficient condition for free n-divisibility.

Proposition 39 *Let μ be a probability measure on $\mathbb{R}$ with finite fourth moment. If μ is infinitely divisible in the free sense then $Kurt^{\boxplus}(\mu) \geq -1/n$.*

Proof. Let μ be n-divisible in the free sense and μ_n be such that

$$\underbrace{\mu_n \boxplus \cdots \boxplus \mu_n}_{n \text{ times}} = \mu.$$

Since $Kurt^{\boxplus}(\mu_n) = nKurt^{\boxplus}(\mu)$ and $Kurt^{\boxplus}(\mu_n) \geq -1$, we get the result. ∎

As a consequence of Proposition 27 the arcsine distribution $\mathrm{a}_{0,2}$ is 2-divisible with respect to $\boxplus$. However, Proposition 39 shows that it is not 3-divisible with respect to $\boxplus$, since $Kurt^{\boxplus}(\mathrm{a}_{0,2}) = Kurt(\mathrm{a}_{0,2}) + 1 = -.5$.

The *monotone kurtosis* of a zero-mean distribution μ is defined as

$$Kurt^{\rhd}(\mu) = \frac{2m_4(\mu) - 3m_2(\mu)^2}{2(m_2(\mu))^2} = Kurt(\mu) + 1.5.$$

In general, the monotone kurtosis is then defined as

$$Kurt^{\rhd}(\mu) = \frac{r_4(\mu)}{(r_2(\mu))^2}.$$

where $r_2(\mu)$ and $r_4(\mu)$ are the second and fourth monotone cumulants. The following result gives a necessary condition in terms of kurtosis for a measure with zero-mean to be n-divisible with respect to monotone convolution. It is proved similarly to the classical and free cases above.

Proposition 40 *Let μ be a probability measure on $\mathbb{R}$ with zero-mean and finite fourth moment. If μ is n-divisible with respect to monotone convolution then $Kurt^{\rhd}(\mu) \geq -\frac{1}{2n}$..*

Using similar ideas we define, for a probability measure μ on $\mathbb{R}$ with zero-mean and fourth moment, the *Boolean kurtosis* as

$$Kurt^{\uplus}(\mu) = \frac{h_4(\mu)}{(h_2(\mu))^2} = Kurt^{\boxplus}(\mu) + 1 = Kurt(\mu) + 2. \tag{2.83}$$

We might expect to obtain a similar criterion as above for Boolean infinite divisibility, but since any measure is infinitely divisible with respect to Boolean convolution, this would only lead to the fact that kurtosis is greater than -2.

2.6.3 Infinite divisibility and cumulants

For distributions having compact support or being determined by moments, working with cumulants turns out to be very useful to rule out measures which are not infinitely divisible.

The main criteria is the conditionally positive definiteness of cumulants for infinitely divisible measures, similar to the classical case. Let $\{\mu_N\}_{N>0}$ be a sequence of measures and, as in Chapter 1, denote by c_n, κ_n, r_n, h_n the corresponding sequence of classical, free, Boolean and monotone cumulants and by m_n the moments.

Lemma 41 *The following statements are equivalent.*

(1) For each $n > 1$ the following limit exists:

$$\lim_{N\to\infty} N m_n(\mu_N).$$

(2) For each $n > 1$ the following limit exists:

$$\lim_{N\to\infty} N c_n(\mu_N).$$

(3) For each $n > 1$ the following limit exists:

$$\lim_{N\to\infty} N k_n(\mu_N).$$

(4) For each $n > 1$ the following limit exists:

$$\lim_{N\to\infty} N r_n(\mu_N).$$

(5) For each $n > 1$ the following limit exists:

$$\lim_{N\to\infty} N h_n(\mu_N).$$

Moreover, if the limits exists they are all equal.

Proof. Let us prove that (2) implies (1). The other implications are similar from the corresponding formulas relating cumulants with moments in Chapter 1. By the moment cumulant formula we have

$$\begin{aligned}\lim_{N\to\infty} N m_n(\mu_N) &= \lim_{N\to\infty} N \sum_{\pi\in\mathcal{P}(n)} c_\pi(\mu_N) && (2.84)\\ &= \lim_{N\to\infty} N c_n(\mu_N) && (2.85)\end{aligned}$$

where the last equality follows since by assumption (2), all partitions π with more than one block tend to zero. ∎

Recall that a sequence $\{a_n\}_{n\geq}$ is conditionally positive definite if for every $n \geq 1$ and $\alpha_i \in \mathbb{C}, i = 1, ..., n$

$$\sum_{i,j=1}^{n} \alpha_i \overline{\alpha_j} a_{i+j} \geq 0.$$

Theorem 42 *Let $\circledast \in \{*, \boxplus, \uplus, \rhd, \lhd\}$ be any of the additive convolutions. Let μ be a probability measure on $\mathbb{R}$ determined by moments and $\circledast$ infinitely divisible. Then the sequence of cumulants with respect to the convolution $\circledast$ is conditionally positive definite.*

Proof. Let μ_N be the N-component of M, i.e. $\mu_N = \mu^{\circledast 1/N}$. Then, for every N, we have that $c_n(\mu) = N c_n(\mu_N)$. By the last lemma we have

$$c_n(\mu_N) = \lim_{N\to\infty} N m_n(\mu_N)$$

and then

$$\begin{aligned}\sum_{n,m=1}^{k} \alpha_n \overline{\alpha}_m \kappa_{n+m} &= \lim_{N\to\infty} N \sum_{n,m=1}^{k} \alpha_n \overline{\alpha}_m m_{m+n}(\mu_n)\\ &= \lim_{N\to\infty} N \sum_{n,m=1}^{k} \alpha_n \overline{\alpha}_m \int_{\mathbb{R}} x^{m+n}(\mu_n)\\ &= \lim_{N\to\infty} N \int_{\mathbb{R}} (\sum_{n=1}^{k} \alpha_n x^n) \cdot (\sum_{m=1}^{k} \overline{\alpha}_m x^m)(\mu_n)\\ &\geq 0.\end{aligned}$$

∎

2.7 Additional Comments

The Herglotz exponential representation (2.8) links free probability to the seemingly unrelated area of representation theory of symmetric groups. More specifically, the irreducible representations of symmetric groups are in one-to-one correspondence with Young diagrams, which in turn, admit a unique Herglotz exponential representation, and hence correspond to a probability measure. Then the limiting shape of a typical diagram obtained by adding boxes according to the Plancherell measure on the symmetric group corresponds to the semicircle law. Some other quite standard shapes of Young diagrams such as the square or the horizontal line, are mapped, respectively, to the Gaussian distributions for Boolean and monotone independencies, namely, the symmetric Bernoulli distribution and the arcsine law. Furthermore, the statistics of operations on representations, such as outer products and restrictions, can be explained in terms of free operations on probability measures, such as free convolutions and free compressions. The interested reader is referred to the following works: The limiting shape of the Young diagrams was established by Kerov and Vershik [KV77]. Its relation to the semicircle law was later found by Kerov in [Ke93]; see also [Ke00], [KV81]. Biane [Bi98] proved that the structure behind this correspondence is much richer, by showing the relations between operations on representation theory and operations in free probability.

2.8 Bibliographic Notes

The material in Herglotz and Pick-Nevanlinna functions in Section 2.1 is mainly a collection of results taken from the books by Akhiezer [Ak65] and Donoghue [Do74]. The canonical representation of a Pick function is essentially a consequence of Cauchy´s integral formula. Its proof and importance in classical infinite divisibility are given in the monograph by Bondesson [Bo92]. For the exponential representation (2.8) we refer to the book by Teschl [Te02].

Section 2.2 on the Cauchy transform is a collection of several results available in the literature and most of them follows from the fact that a Cauchy transform is a Pick function; see also the books by Hiai and Petz [HP00] and Hora and Obata [HO07]. The negative of a Cauchy transform is sometimes called the Borel transform in Theory of Operators (e.g.[Te02]) or the Stieltjes transform. The latter is extensively used in the study of the asymptotic analysis of the spectrum of random matrices; see for example the book by Bai and Silverstein [BS10].

The classical convolution of measures and its role in classical infinite divisibility is systematically presented in the book by Sato [Sa99]. The analytic approach to free convolution was first initiated by Voiculescu [Vo86] in the case of probability measures with compact support and by Maassen [Ma92] for distributions with finite variance. The material in Sections 2.3 and 2.4 on analytic transforms for possibly unbounded distributions, are mainly based in the seminal paper by Bercovici and Voiculescu [BV93] in the case of free convolutions, Speicher and Woroudi [SW97] in the Boolean case and Franz and Muraki [FM05] in the monotone case. In some cases we have presented additional details for the convenience of the reader. See also [BNT06], [HP00].

The combinatorial approach to non-classical convolutions is not considered in this work, apart from the material already presented in Chapter 1. For a comprehensive combinatorial treatment of free convolutions of measures with compact support we refer to the book by Nica and Speicher [NS06] where the free cumulates are extensively studied.

The S-transform for probability measures on $\mathbb{R}_+$ was also introduced by Voiculescu [Vo87] and systematically studied by Bercovici and Voiculescu [BV93]. Proposition 33 collects Corollaries 6.6 and 6.7 and Lemma 6.9 in [BV93]. The S-transform for probability measures with bounded support and zero mean was considered by Raj Rao and Speicher [SRR07] and the case of unbounded support by Arizmendi and Pérez-Abreu [APA09].

The symmetric beta distribution obtained as the multiplicative convolution $\mathrm{a}_s \boxtimes \mathrm{m}_1$ in Example 2.5.4 is an uncommon distribution with an explicit Cauchy transform. This example is due to Arizmendi et al. [ABNPA10]. A generalization was obtained by Arizmendi and Hasebe [AH12] who found a whole class of distributions with explicit Cauchy transforms, which includes the symmetric beta distribution.

The book by Steutel and Van Harn [SV03] contains some sufficient conditions for classical infinite divisibility based on the first four cumulants. The relevance of kurtosis in infinite divisibility is considered in Arizmendi and Pérez-Abreu [APA10]. The conditionally positive definiteness of classical cumulants is part of the folklore in the literature. The free case is considered in Nica and Speicher [NS06] and the monotone case is due to Hasebe and Saigo [HS11].

Bibliography

[Ak65] Akhiezer, N.I. (1965). *The Classical Moment Problem.* Oliver & Boyd, Edinburgh-London.

[ABNPA10] Arizmendi, O., O. E. Banrdorff-Nielsen & V. Pérez-Abreu (2010). On free and classical type G distributions. *Braz. J. Probab. Statist.* **24**, 106-127.

[AH12] Arizmendi, O. & T. Hasebe (2012). On a class of explicit Cauchy-Stieltjes transforms related to monotone stable and free Poisson laws. *Bernoulli.* To appear.

[APA09] Arizmendi, O. & V. Pérez-Abreu (2009). The S-transform of symmetric probability measures with unbounded support. *Proc. Amer. Math. Soc.* **137**, 3057-3066.

[APA10] Arizmendi, O. & V. Pérez-Abreu (2010). On the non-classical infinite divisibility of power semicircle distributions. *Comm. Stoch. Anal.* **4**, 161-178.

[BNT06] Barndorff-Nielsen, O. E. & S. Thorbjørnsen (2006). Classical and free infinite divisibility and Lévy processes. In: *Quantum Independent Increment Processes II.* M. Schürmann and U. Franz (eds). Lecture Notes in Mathematics **1866**, Springer, Berlin.

[BS10] Bai, Z. and J.W. Silverstein (2010). *Spectral Analysis of Large Dimensional Random Matrices.* Springer.

[BV93] Bercovici, H. & D. Voiculescu (1993). Free convolution of measures with unbounded supports. *Indiana Journal of Mathematics* **42**, 733-773.

[Bi98] Biane, P. (1998). Representations of symmetric groups and free probability. *Advances in Mathematics* **138,** 126-181.

[Bo92] Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities.* Lecture Notes in Statist., **76**, Springer, New York.

[BOO10] Borodin, A. Okounkov, A. & Olshanski, G. (2000). Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* **13**, 481–515.

[Do74] Donoghue, W (1974). *Monotone Matrix Functions and Analytic Continuation.* Springer, New York.

[FM05] Franz, U. & N. Muraki (2005). Markov property of monotone Lévy processes. In *Infinite Dimensional Harmonic Analysis III.* H. Heyer, T. Hirai, T. Kawazoe, K. Saito (eds). World Scientific, Singapore, pp 37-57.

[Ga76] Gardner, M. (1976). Catalan numbers: An integer sequence that materializes in unexpected places. *Scientific American* **234**, 120-125.

[HS11] Hasebe, T. & Saigo, H. (2011). The monotone cumulants. *Ann. Inst. Henri Poincaré*, **47**, 1160-1170-

[HP00] Hiai, F. & D. Petz (2000). *The Semicircle Law, Free Random Variables and Entropy.* Mathemathical Surveys and Monographs **77,** American Mathematical Society, Providence.

[HO07] Hora, A. & N. Obata (2007). *Quantum Probability and Spectral Analysis of Graphs.* Series Theoretical and Mathematical Physics, XVIII, 371 Springer, New York.

[Ke93] Kerov, S. V. (1993). Transition probabilities for continual Young diagrams and the Markov moment problem. *Functional Analysis and Applications* **27**, 104–117.

[Ke00] Kerov, S. V. (2000). Anisotropic Young diagrams and Jack symmetric functions. *Functional Analysis and Applications* **31**, 41–51.

[KV77] Kerov, S. V. & Vershik, A. M. (1977). Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables. *Soviet Mathematics Doklady* **18**, 527–531.

[KV81] Kerov, S. V. & Vershik, A. M. (1981). The asymptotics character theory of the symmetric group. *Functional Analysis and Applications* **15**, 246–255.

[Ma92] Maaseen, H. (1992). Addition of freely independent random variables. *Journal of Functional Analysis* **106** 409-438.

[Mo09] Mlotkowski, W. (2009). Combinatorial relation between free cumulants and Jacobi parameters. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **12**, 291-306.

[NS06] Nica, A. & R. Speicher (2006). *Lectures on the Combinatorics of Free Probability.* London Mathematical Society Lecture Notes Series **335,** Cambridge University Press, Cambridge.

[Sa99] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

[SW97] Speicher, R. & R. Woroudi (1997). Boolean convolution, *Fields Institute Communications*, **12** D. Voiculescu, (ed). American Mathematical Society, pp. 267-279

[SRR07] Speicher, R. & N. Raj Rao (2007). Multiplication of free random variables and the S-transform: The case of vanishing mean. *Electronic Communications in Probability* **12,** 248-258.

[SV03] Steutel, F. W. & K. Van Harn (2003). *Infinite Divisibility of Probability Distributions on the Real Line.* Marcel-Dekker, New York.

[Te02] Teschl, G. (2000). *Jacobi Operators and Completely Integrable Nonlinear Lattices.* Mathematical Surveys and Monographs **72,** American Mathematical Society, Providence.

[Vo86] Voiculescu, D. (1986). Addition of certain non-commuting random variables. *Journal of Functional Analysis* **66**, 323-346.

[Vo87] Voiculescu, D. (1987). Multiplication of certain non-commuting random variables. *Journal of Operator Theory* **18**, 223-235.