Lévy laws in free probability

Ole E. Barndorff-Nielsen[†] and Steen Thorbjørnsen[‡§]

[†]Department of Mathematical Sciences, University of Aarhus, DK-8000 Aarhus C, Denmark; and [‡]Department of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense M, Denmark

Edited by Richard V. Kadison, University of Pennsylvania, Philadelphia, PA, and approved October 7, 2002 (received for review January 29, 2002)

This article and its sequel outline recent developments in the theory of infinite divisibility and Lévy processes in free probability, a subject area belonging to noncommutative (or quantum) probability. The present paper discusses the classes of infinitely divisible probability measures in classical and free probability, respectively, via a study of the Bercovici–Pata bijection between these classes.

In this article and its follow-up (1) we outline certain aspects of the theory of infinite divisibility and Lévy processes in the framework of Voiculescu's free probability. Many researchers in the area have felt that an analogue of the highly developed theory of Lévy processes might be constructed in the framework of free probability, but it has not been clear how to do this. (Recall that one major difficulty in noncommutative probability is the lack of a notion of the joint distribution of two noncommuting operators (as a probability measure.) This fact means that many arguments from classical probability cannot be carried over directly to the noncommutative case). This article and its sequel address some central issues of this research program.

While the follow-up article focuses on the study of Lévy processes in free probability, our study in this paper is concentrated around the bijection, introduced by Bercovici and Pata in ref. 2, between the class of classically infinitely divisible probability measures and the class of freely infinitely divisible probability measures. We derive in *The Bercovici–Pata Bijection* certain algebraic and topological properties of this bijection, in the present article denoted Λ, and explain how these properties imply that Λ maps certain canonical subclasses of classically infinitely divisible probability measures onto their natural free counterparts. *Noncommutative Probability* and *Free Independence* provide background material on noncommutative probability in general and free probability in particular. *Classical and Free Convolution* and *Infinite Divisibility, Self-Decomposability, and Stability* review briefly the theory of convolution and infinite divisibility in free (and classical) probability. The theory, outlined in *The Bercovici–Pata Bijection* and ref. 1, is developed in detail in ref. 3 and in forthcoming articles (O.E.B.-N. and S.T., unpublished work).

1. Noncommutative Probability

In classical probability, the basic objects of study are random variables, i.e. measurable functions from a probability space $(\Omega, \mathcal{F}, P)$ into the real numbers $\mathbb{R}$ equipped with the Borel σ-algebra $\mathcal{B}$. To any such random variable $X : \Omega \to \mathbb{R}$ there is associated a probability measure μ_X on $(\mathbb{R}, \mathcal{B})$ defined by $\mu_X(B) = P(X \in B) = P[X^{-1}(B)]$ for any Borel set B. The measure μ_X is called the distribution of X [with respect to (w.r.t.) P], and it satisfies the property that

$$\int_{\mathbb{R}} f(t)\,\mu_X(dt) = \mathbb{E}[f(X)]\,,$$

for any bounded Borel function $f : \mathbb{R} \to \mathbb{R}$, and where $\mathbb{E}$ denotes expectation (or integration) w.r.t. P. We also shall use the notation $L\{X\}$ for μ_X (where L stands for "law").

In noncommutative probability, one replaces the random variables by (self-adjoint) operators on a Hilbert space $\mathcal{H}$. These operators are then referred to as "noncommutative random variables." The term noncommutative refers to the fact that, in this setting, the multiplication of "random variables" (i.e. composition of operators) is no longer commutative as opposed to the usual multiplication of classical random variables. The noncommutative situation is often remarkably different from the classical one and most often more complicated. By $\mathcal{B}(\mathcal{H})$ we denote the algebra of all bounded operators on $\mathcal{H}$. Recall that $\mathcal{B}(\mathcal{H})$ is equipped with an involution (the adjoint operation) $a \mapsto a^* : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$, which is given by

$$\langle a\xi, \eta\rangle = \langle \xi, a^*\eta\rangle, \qquad [a \in \mathcal{B}(\mathcal{H}),\ \xi, \eta \in \mathcal{H}]\,.$$

Instead of working with the whole algebra $\mathcal{B}(\mathcal{H})$ as the set of "random variables" under consideration, it is, for most purposes, natural to restrict attention to certain subalgebras of $\mathcal{B}(\mathcal{H})$. In this article we shall only consider the nicest cases of such subalgebras, the *von Neumann algebras,* although much of what follows is also valid for more general classes of "noncommutative probability spaces." A von Neumann algebra, acting on a Hilbert space $\mathcal{H}$, is a subalgebra of $\mathcal{B}(\mathcal{H})$ that contains the multiplicative unit $\mathbf{1}$ of $\mathcal{B}(\mathcal{H})$ (i.e. $\mathbf{1}$ is the identity mapping on $\mathcal{H}$) and is closed under the adjoint operation and in the weak operator topology on $\mathcal{B}(\mathcal{H})$ [i.e. the weak topology on $\mathcal{B}(\mathcal{H})$ induced by the linear functionals: $a \mapsto \langle a\xi, \eta\rangle$, $\xi, \eta \in \mathcal{H}$]. A tracial state on a von Neumann algebra $\mathcal{A}$ is a positive linear functional $\tau : \mathcal{A} \to \mathbb{C}$, taking the value 1 at the identity operator $\mathbf{1}$ on $\mathcal{H}$, and satisfying the trace property[¶]

$$\tau(ab) = \tau(ba), \qquad (a, b \in \mathcal{A}).$$

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: w.r.t., with respect to.

[§]To whom correspondence should be addressed. E-mail: steenth@imada.sdu.dk.

[¶]In quantum physics, τ is of the form $\tau(a) = \mathrm{tr}(\rho a)$, where ρ is a trace class self-adjoint operator on $\mathcal{H}$ with trace $\mathbf{1}$, that expresses the state of a quantum system, and a would be an observable, i.e., a self-adjoint operator on $\mathcal{H}$, the mean value of the outcome of observing a being $\tau(a) = \mathrm{tr}(\rho a)$.

1.1. Definition: A W^*-probability space is a pair $(\mathcal{A}, \tau)$, where $\mathcal{A}$ is a von Neumann algebra on a Hilbert space $\mathcal{H}$, and τ is a faithful tracial state on $\mathcal{A}$.

The assumed faithfulness of τ in *Definition 1.1* means that τ does not annihilate any nonzero positive operator. It implies that $\mathcal{A}$ is finite in the sense of Murray and von Neumann.

Suppose now that $(\mathcal{A}, \tau)$ is a W^*-probability space and that a is a self-adjoint operator (i.e. $a^* = a$) in $\mathcal{A}$. Then, as in the classical case, we can associate a (spectral) distribution to a in a natural way: Indeed, by the Riesz representation theorem, there exists a unique probability measure μ_a on $(\mathbb{R}, \mathcal{B})$, satisfying that

$$\int_{\mathbb{R}} f(t)\mu_a(dt) = \tau[f(a)], \qquad \textbf{[1.1]}$$

for any bounded Borel function $f : \mathbb{R} \to \mathbb{R}$. In Eq. **1.1**, $f(a)$ has the obvious meaning if f is a polynomial. For general Borel functions f, $f(a)$ is defined in terms of spectral theory (e.g. see ref. 4).

The (spectral) distribution μ_a of a self-adjoint operator a in $\mathcal{A}$ is automatically concentrated on the spectrum $\mathrm{sp}(a)$, and thus is, in particular, compactly supported. If one wants to be able to consider any probability measure μ on $\mathbb{R}$ as the spectral distribution of some self-adjoint operator, then it is necessary to take unbounded (i.e. noncontinuous) operators into account. Such an operator a is generally not defined on all of $\mathcal{H}$ but only on a subspace $\mathcal{D}(a)$ of $\mathcal{H}$, called the domain of a. We say then that a is an operator *in* $\mathcal{H}$ rather than *on* $\mathcal{H}$. For most of the interesting examples, $\mathcal{D}(a)$ is a dense subspace of $\mathcal{H}$, in which case a is said to be densely defined.

If $(\mathcal{A}, \tau)$ is a W^*-probability space acting on $\mathcal{H}$ and a is an unbounded operator in $\mathcal{H}$, a cannot be an element of $\mathcal{A}$. The closest a can get to $\mathcal{A}$ is to be *affiliated* with $\mathcal{A}$, which means that a commutes with any unitary operator u that commutes with all elements of $\mathcal{A}$. If a is self-adjoint, a is affiliated with $\mathcal{A}$ if and only if $f(a) \in \mathcal{A}$ for any bounded Borel function $f : \mathbb{R} \to \mathbb{R}$. In this case, Eq. **1.1** determines, again, a unique probability measure μ_a on $\mathbb{R}$, which we also refer to as the (spectral) distribution of a and generally has unbounded support. Furthermore, any probability measure on $\mathbb{R}$ can be realized as the (spectral) distribution of some self-adjoint operator affiliated with some W^*-probability space. In the following we also shall use the notation $L\{a\}$ for the distribution of a (possibly unbounded) operator a affiliated with $(\mathcal{A}, \tau)$.

2. Free Independence

The key concept on relations between classical random variables X and Y is *independence*. One way of defining that X and Y [defined on the same probability space $(\Omega, \mathcal{F}, P)$] are independent is to ask that all compositions of X and Y with bounded Borel functions be uncorrelated,

$$\mathbb{E}\{[f(X) - \mathbb{E}\{f(X)\}]\cdot[g(Y) - \mathbb{E}\{g(Y)\}]\} = 0,$$

for any bounded Borel functions f, g: $\mathbb{R} \to \mathbb{R}$.

In the early 1980s, Voiculescu introduced the notion of *free independence* among noncommutative random variables.

2.1. Definition: Let $a_1, a_2, \ldots, a_r$ be self-adjoint operators affiliated with a W^*-probability space $(\mathcal{A}, \tau)$. We say then that $a_1, a_2, \ldots, a_r$ are *freely independent* w.r.t. τ, if

$$\tau\{[f_1(a_{i_1}) - \tau(f_1(a_{i_1}))][f_2(a_{i_2}) - \tau(f_2(a_{i_2}))] \cdots [f_p(a_{i_p}) - \tau(f_p(a_{i_p}))]\} = 0,$$

for any p in $\mathbb{N}$, any bounded Borel functions $f_1, f_2, \ldots, f_p : \mathbb{R} \to \mathbb{R}$, and any indices $i_1, i_2, \ldots, i_p$ in $\{1, 2, \ldots, r\}$ satisfying that $i_1 \neq i_2, i_2 \neq i_3, \ldots, i_{p-1} \neq i_p$.

At first glance, the definition of free independence looks quite similar perhaps to the definition of classical independence given above, and indeed in many respects free independence is conceptually similar to classical independence. For example, if $a_1, a_2, \ldots, a_r$ are freely independent self-adjoint operators affiliated with $(\mathcal{A}, \tau)$, then all numbers of the form $\tau\{f_1(a_{i_1})f_2(a_{i_2}) \cdots f_p(a_{i_p})\}$ (where $i_1, i_2, \ldots, i_p \in \{1, 2, \ldots, r\}$ and $f_1, f_2, \ldots, f_p : \mathbb{R} \to \mathbb{R}$ are bounded Borel functions) are uniquely determined by the distributions $L\{a_i\}$, $i = 1, 2, \ldots, r$. On the other hand, free independence is a truly noncommutative notion, which can be seen, for instance, from the easily checked fact that two classical random variables are never freely independent unless one of them is trivial, i.e. constant with probability 1 (e.g. see ref. 5).

Voiculescu originally introduced free independence in connection with his deep studies of the von Neumann algebras associated to the free groups (see refs. 6–8). We prefer in this article, however, to indicate the significance of free independence by explaining its connection with random matrices. In the 1950s, the phycisist E. P. Wigner showed that the spectral distribution of large self-adjoint random matrices with independent complex Gaussian entries is, approximately, the semicircle distribution, i.e. the distribution on $\mathbb{R}$ with density $s \mapsto \sqrt{4 - s^2}\cdot 1_{[-2,2]}(s)$ w.r.t. Lebesgue measure. More precisely, for each n in $\mathbb{N}$, let $X^{(n)}$ be a self-adjoint complex Gaussian random matrix of the kind considered by Wigner (and suitably normalized), and let tr_n denote the (usual) tracial state on the $n \times n$ matrices $M_n(\mathbb{C})$. Then for any positive integer p, Wigner showed that

$$\mathbb{E}\{\mathrm{tr}_n[(X^{(n)})^p]\} \xrightarrow[n \to \infty]{} \int_{-2}^{2} s^p \sqrt{4 - s^2}\, ds.$$

In the late 1980s, Voiculescu generalized Wigner's result to families of independent self-adjoint Gaussian random matrices (cf. ref. 7): For each n in $\mathbb{N}$, let $X_1^{(n)}, X_2^{(n)}, \ldots, X_r^{(n)}$ be independent[‖] random matrices of the kind considered by Wigner. Then for any indices $i_1, i_2, \ldots, i_p$ in $\{1, 2, \ldots, r\}$,

‖In the classical sense, at the level of the entries.

$$\mathbb{E}\{\mathrm{tr}_n[X_{i_1}^{(n)}X_{i_2}^{(n)}\cdots X_{i_p}^{(n)}]\}\xrightarrow[n\to\infty]{}\tau\{x_{i_1}x_{i_2}\cdots x_{i_p}\},$$

where $x_1, x_2, \ldots, x_r$ are freely independent self-adjoint operators in a W^*-probability space $(\mathcal{A}, \tau)$ and such that $L\{x_i\}$ is the semicircle distribution for each i.

By Voiculescu's result, free independence describes what the assumed classical independence between the random matrices is turned into, as $n \to \infty$. Also, from a classical probabilistic point of view, free-probability theory may be considered as (an aspect of) the probability theory of large random matrices.

Voiculescu's result reveals another general fact in free probability, namely that the role of the Gaussian distribution in classical probability is taken over by the semicircle distribution in free probability. In particular, as also proved by Voiculescu, the limit distribution appearing in the free version of the central-limit theorem is the semicircle distribution (e.g. see ref. 9).

3. Classical and Free Convolution

In classical probability, the convolution $\mu_1 * \mu_2$ of two probability measures μ_1 and μ_2 on $\mathbb{R}$ is defined as the distribution of the sum $X_1 + X_2$ of two independent random variables X_1 and X_2 with distributions μ_1 and μ_2, respectively. The existence of two independent random variables X_1 and X_2 defined on the same probability space and with prescribed distributions μ_1 and μ_2 follows from a tensor-product construction. In free probability, the corresponding existence result follows from a similar construction, where the tensor product is replaced by the so-called free product (we refer to ref. 9 for details). Furthermore, as previously indicated, if x_1 and x_2 are freely independent self-adjoint operators with spectral distributions μ_1 and μ_2, the distribution $L\{x_1 + x_2\}$ depends only on μ_1 and μ_2. Hence, it makes sense to define the free convolution $\mu_1 \boxplus \mu_2$ of μ_1 and μ_2 by setting $\mu_1 \boxplus \mu_2 = L\ \{x_1 + x_2\}$. Once the free-convolution $\boxplus$ has thus been defined, one could from a probabilistic point of view forget about the underlying operator construction and merely consider $\boxplus$ as a new type of convolution on the set of probability measures on $\mathbb{R}$. To a large extent, this approach can in fact be followed through by virtue of the analytical function tools that we describe next.

The main tool for dealing with classical convolution is the Fourier transform. The Fourier transform (or characteristic function) of a probability measure μ on $\mathbb{R}$ is the function $f_\mu : \mathbb{R} \to \mathbb{C}$ given by

$$f_\mu(u) = \int_{\mathbb{R}} e^{isu}\mu(ds), \quad (u \in \mathbb{R}).$$

We denote by C_μ the *cumulant transform* of μ, i.e. the logarithm of the Fourier transform of μ,

$$C_\mu(u) = \log f_\mu(u), \quad (u \in \mathbb{R}).$$

The key property of the Fourier transform in this connection is that

$$f_{\mu_1 * \mu_2}(u) = f_{\mu_1}(u)\cdot f_{\mu_2}(u), \quad (u \in \mathbb{R}),$$

for any probability measures μ_1, μ_2 on $\mathbb{R}$. Thus, the cumulant transform linearizes classical convolution.

In ref. 10, Voiculescu found a transformation that linearizes free convolution; the so-called R transform. Since then, several modifications of Voiculescu's R transform have appeared in the literature, in particular the so-called *Voiculescu transform* and what we shall refer to as the *free cumulant transform.* These transforms are defined as follows.

By $\mathbb{C}^+$ (respectively $\mathbb{C}^-$) we denote the strictly upper (respectively strictly lower) complex half-plane. For a probability measure μ on $\mathbb{R}$, the *Cauchy transform* $G_\mu : \mathbb{C}^+ \to \mathbb{C}^-$ is defined by

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z - t}\,\mu(dt), \quad (z \in \mathbb{C}^+).$$

It was proved in ref. 11 that the mapping $F_\mu := 1/G_\mu : \mathbb{C}^+ \to \mathbb{C}^+$ always has a right inverse, F_μ^{-1}, defined on a region of the form: $\Gamma(\eta, M) = \{x + iy \in \mathbb{C}^+ \mid x^2 + y^2 > M^2, |x| < \eta y\}$, where η and M are positive numbers. The Voiculescu transform ϕ_μ is then defined by

$$\phi_\mu(z) = F_\mu^{-1}(z) - z, \quad [z \in \Gamma(\eta, M)],$$

and the free cumulant transform $\mathcal{C}_\mu$ is defined by

$$\mathcal{C}_\mu(z) = z\phi_\mu\left(\frac{1}{z}\right) = zF_\mu^{-1}\left(\frac{1}{z}\right) - 1,$$

for $1/z$ in $\Gamma(\eta, M)$, i.e. for z in the region $\{x - iy \in \mathbb{C}^- \mid x^2 + y^2 < M^{-2}, |x| < \eta y\}$. As indicated above, the key property of the Voiculescu transform, proved in ref. 11, is that $\phi_{\mu_1 \boxplus \mu_2}(z) = \phi_{\mu_1}(z) + \phi_{\mu_2}(z)$, for any probability measures μ_1, μ_2 on $\mathbb{R}$. A similar property holds, of course, for the free cumulant transform, i.e. we have

$$\mathcal{C}_{\mu_1 \boxplus \mu_2}(z) = \mathcal{C}_{\mu_1}(z) + \mathcal{C}_{\mu_2}(z).$$

We prefer in this article to work with the free cumulant transform rather than the Voiculescu transform or other modifications of it. The reason is that this particular modification is especially close in nature to the classical cumulant transform. In particular, it behaves exactly like the classical cumulant transform w.r.t. scalar multiplication, which is important for the discussion of *free* *self-decomposability* introduced in ref. 3 (see *Infinite Divisibility, Self-Decomposability, and Stability*). Indeed, if μ is the distribution of a random variable X and $c \geq 0$, then denoting by $D_c\mu$ the distribution of cX we have the relation

$$C_{D_c\mu}(z) = C_\mu(cz). \quad \textbf{[3.1]}$$

Furthermore, in terms of C_μ the free Lévy–Khintchine representation of freely infinitely divisible probability measures resembles more closely the classical Lévy–Khintchine representation, as we shall see in *The Bercovici–Pata Bijection.*

4. Infinite Divisibility, Self-Decomposability, and Stability

In classical probability theory one has the following hierarchy of classes of probability measures on $\mathbb{R}$,

$$\mathcal{G}(*) \subset \mathcal{S}(*) \subset \mathcal{L}(*) \subset \mathcal{ID}(*) \subset \mathcal{P},$$

where

(*i*) $\mathcal{P}$ is the class of all probability measures on $\mathbb{R}$,
(*ii*) $\mathcal{ID}(*)$ is the class of infinitely divisible probability measures on $\mathbb{R}$, i.e.

$$\mu \in \mathcal{ID}(*) \Leftrightarrow \forall\, n \in \mathbb{N}\ \exists\, \mu_n \in \mathcal{P}\colon\ \mu = \underbrace{\mu_n * \mu_n * \cdots * \mu_n}_{n \text{ terms}},$$

(*iii*) $\mathcal{L}(*)$ is the class of self-decomposable probability measures on $\mathbb{R}$, i.e.

$$\mu \in \mathcal{L}(*) \Leftrightarrow \forall\, c \in]0, 1[\ \exists \mu_c \in \mathcal{P}\colon\ \mu = D_c\mu * \mu_c,$$

(*iv*) $\mathcal{S}(*)$ is the class of stable probability measures on $\mathbb{R}$, i.e.

$$\mu \in \mathcal{S}(*) \Leftrightarrow \{\psi(\mu) \mid \psi\colon \mathbb{R} \to \mathbb{R}, \text{ increasing affine transformation}\} \text{ is closed under convolution } *, \text{ and}$$

(*v*) $\mathcal{G}(*)$ is the class of Gaussian (or normal) distributions on $\mathbb{R}$.

The classes of probability measures, defined above, are all of great importance in classical probability. This is partly explained by their characterizations as limit distributions of certain types of sums of independent random variables (e.g. see ref. 12 or 13).

In free probability, we denote by $\mathcal{ID}(\boxplus)$, $\mathcal{L}(\boxplus)$, and $\mathcal{S}(\boxplus)$ the classes of, respectively, freely infinitely divisible, freely self-decomposable, and freely stable probability measures on $\mathbb{R}$. These classes are defined exactly as the corresponding classical classes except that one replaces classical convolution $*$ by free convolution $\boxplus$ throughout in *ii–iv* above. Furthermore, we shall denote by $\mathcal{G}(\boxplus)$ the class of free Gaussian distributions, i.e. that of semicircle distributions. It turns out, then, that in free probability, we also have the hierarchy

$$\mathcal{G}(\boxplus) \subset \mathcal{S}(\boxplus) \subset \mathcal{L}(\boxplus) \subset \mathcal{ID}(\boxplus) \subset \mathcal{P}.$$

The first inclusion is well known and easily verified, and the second one is not hard to prove by application of the free-cumulant transform. The third inclusion is of a deeper nature. As in the classical case, it is a consequence of the fact that the infinitely divisible distributions may be characterized as the possible limit distributions, as $n \to \infty$, of sums $S_n = X_{n,1} + \cdots + X_{n,k_n}$ of (freely) independent random variables such that the terms $X_{n,1}, \ldots, X_{n,k_n}$ are uniformly negligible (in probability) as $n \to \infty$. The latter result was proved in 1937 by Khintchine (14) in the classical case and recently by Bercovici and Pata in the free case (15). Based on this, it remains to remark that by successive applications of *iii* above, a measure μ in $\mathcal{L}(\boxplus)$ can be considered as the (fixed) distribution of sums S_n of the kind described above (see ref. 3 for details).

We shall focus here mostly on the class $\mathcal{L}(*)$ of self-decomposable measures (and its free counterpart), which until fairly recently seemed to be more or less forgotten except by a few experts, although it did receive considerable attention in the early studies of infinite divisibility. It was introduced first as a class of limit distributions by Lévy and is now playing a substantial role in mathematical finance (see the contribution by O.E.B.-N. and Shephard in ref. 16).

A random variable Y has distribution in $\mathcal{L}(*)$ if and only if Y has, for any c in $]0, 1[$, a representation in law of the form**

$$Y \overset{\mathrm{d}}{=} cY + Y_c,$$

for some random variable Y_c, which is independent of Y. This latter formulation makes the idea of self-decomposability of immediate appeal from the viewpoint of mathematical modeling. In the follow-up article (1), we consider yet another characterisation of self-decomposability, which describes the self-decomposable probability measures exactly as the laws of certain stochastic integrals w.r.t. certain Lévy processes. One of the main results of the follow-up article is a free analogue of this particular characterization.

5. The Bercovici–Pata Bijection

We present next a bijection between the classes $\mathcal{ID}(*)$ and $\mathcal{ID}(\boxplus)$, which was introduced by Bercovici and Pata in ref. 2. The bijection is defined in terms of the Lévy–Khintchine representations of classical and free infinitely divisible probability measures.

In the classical case, a famous result, due to Lévy and Khintchine (who build on initial work by Kolmogorov) states that a probability measure μ on $\mathbb{R}$ is in $\mathcal{ID}(*)$ if and only if its cumulant transform C_μ has a representation in the form

**The symbol "$\overset{\mathrm{d}}{=}$" means "has the same distribution as."

$$C_\mu(u) = i\gamma u + \int_{\mathbb{R}} \left(e^{iut} - 1 - \frac{iut}{1+t^2}\right)\frac{1+t^2}{t^2}\,\sigma(dt), \quad (u \in \mathbb{R}), \qquad \textbf{[5.1]}$$

where γ is a real constant and σ is a finite measure on $\mathbb{R}$. The pair (γ, σ) is uniquely determined, and it is termed the *generating pair* for μ.

The free version of the Lévy-Khintchine representation was proved, in the general case, by Bercovici and Voiculescu in ref. 11. It asserts that a probability measure ν on $\mathbb{R}$ is in $\mathcal{ID}(\boxplus)$ if and only if its Voiculescu transform ϕ_ν has a representation in the form

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

where γ is a real constant and σ is a finite measure on $\mathbb{R}$. Again, the pair (γ, σ) is uniquely determined, and it is called the *free generating pair* for ν.

5.1. Definition: The Bercovici–Pata bijection is the mapping Λ: $\mathcal{ID}(*) \to \mathcal{ID}(\boxplus)$ defined in the following way: Suppose μ is in $\mathcal{ID}(*)$ and has generating pair (γ, σ); then $\Lambda(\mu)$ is the measure in $\mathcal{ID}(\boxplus)$ with free generating pair (γ, σ).

From the characterizations of $\mathcal{ID}(*)$ and $\mathcal{ID}(\boxplus)$ in terms of the Lévy–Khintchine representations, it is immediate that Λ is in fact a bijection. At first glance, Λ may seem like a very formal correspondence, but as we shall see, Λ preserves a lot of structure between $\mathcal{ID}(*)$ and $\mathcal{ID}(\boxplus)$.

In modern literature on (classical) infinite divisibility (cf. ref. 17), the Lévy–Khintchine representation (Eq. **5.1**) is often rewritten to the equivalent form

$$C_\mu(u) = i\eta u - \frac{1}{2}au^2 + \int_{\mathbb{R}} (e^{iut} - 1 - iut1_{[-1,1]}(t))\rho(dt), \quad (u \in \mathbb{R}), \qquad \textbf{[5.2]}$$

where η is a real constant, a is a nonnegative constant, and ρ is a measure on $\mathbb{R}$ satisfying the conditions

$$\rho(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}} \min\{1, t^2\}\rho(dt) < \infty,$$

i.e. ρ is a *Lévy measure.* The triplet (a, ρ, η) is uniquely determined and is called the generating triplet for μ. The relationship between the two representations Eqs. **5.1** and **5.2** is as follows:

$$\begin{aligned} a &= \sigma(\{0\}), \\ \rho(dt) &= \frac{1+t^2}{t^2}\cdot 1_{\mathbb{R}\setminus\{0\}}(t)\sigma(dt), \qquad \textbf{[5.3]} \\ \eta &= \gamma + \int_{\mathbb{R}} t\left(1_{[-1,1]}(t) - \frac{1}{1+t^2}\right)\rho(dt). \end{aligned}$$

It turns out that the resemblance between the classical and free Lévy–Khintchine representations becomes stronger if one uses the free cumulant transform C_μ rather than the Voiculescu transform ϕ_μ as well as generating triplets rather than pairs.

5.2. Proposition.

(*i*) *A probability measure ν on $\mathbb{R}$ is $\boxplus$-infinitely divisible if and only if there exists a nonnegative number a, a real number η, and a Lévy measure ρ such that the free-cumulant transform C_ν has the representation*

$$C_\nu(z) = \eta z + az^2 + \int_{\mathbb{R}} \left[\frac{1}{1-tz} - 1 - tz1_{[-1,1]}(t)\right]\rho(dt), \quad (z \in \mathbb{C}^-). \qquad \textbf{[5.4]}$$

In that case, the triplet (a, ρ, η) is uniquely determined and is called the free generating triplet for ν.

(*ii*) *If μ is a measure in $\mathcal{ID}(*)$ with (classical) generating triplet (a, ρ, η), then $\Lambda(\mu)$ has free generating triplet (a, ρ, η).*

Proof:

(*i*) Let ν be a measure in $\mathcal{ID}(\boxplus)$ with free generating pair (γ, σ), and consider its free Lévy–Khintchine representation (in terms of the Voiculescu transform):

$$\phi_\nu(z) = \gamma + \int_{\mathbb{R}} \frac{1+tz}{z-t}\,\sigma(dt), \quad (z \in \mathbb{C}^+). \qquad \textbf{[5.5]}$$

Then define the triplet (a, ρ, η) by Eq. **5.3**, and note that

$$\sigma(dt) = a\delta_0(dt) + \frac{t^2}{1+t^2}\rho(dt),$$

$$\gamma = \eta - \int_{\mathbb{R}} t\left(1_{[-1,1]}(t) - \frac{1}{1+t^2}\right)\rho(dt).$$

Now, for z in $\mathbb{C}^-$, the corresponding free cumulant transform $\mathcal{C}_\nu$ is given by

$$\begin{aligned}
\mathcal{C}_\nu(z) = z\phi_\nu(1/z) &= z\left[\gamma + \int_{\mathbb{R}} \frac{1+t(1/z)}{(1/z)-t}\sigma(dt)\right] \\
&= \gamma z + z\int_{\mathbb{R}} \frac{z+t}{1-tz}\sigma(dt) = \gamma z + \int_{\mathbb{R}} \frac{z^2+tz}{1-tz}\sigma(dt) \\
&= \eta z - \left[\int_{\mathbb{R}} t\left[1_{[-1,1]}(t) - \frac{1}{1+t^2}\right]\rho(dt)\right]z + az^2 + \int_{\mathbb{R}} \frac{z^2+tz}{1-tz}\frac{t^2}{1+t^2}\rho(dt).
\end{aligned}$$

Note here that

$$1_{[-1,1]}(t) - \frac{1}{1+t^2} = 1 - \frac{1}{1+t^2} - 1_{\mathbb{R}\setminus[-1,1]}(t) = \frac{t^2}{1+t^2} - 1_{\mathbb{R}\setminus[-1,1]}(t),$$

such that

$$\int_{\mathbb{R}} t\left[1_{[-1,1]}(t) - \frac{1}{1+t^2}\right]\rho(dt) = \int_{\mathbb{R}}\left[\frac{t}{1+t^2} - t^{-1}1_{\mathbb{R}\setminus[-1,1]}(t)\right]t^2\rho(dt).$$

Note also that

$$\frac{z^2+tz}{(1-tz)(1+t^2)} = \frac{z^2}{1-tz} + \frac{tz}{1+t^2}.$$

Therefore,

$$\begin{aligned}
\mathcal{C}_\nu(z) &= \eta z - \left\{\int_{\mathbb{R}}\left[\frac{t}{1+t^2} - t^{-1}1_{\mathbb{R}\setminus[-1,1]}(t)\right]t^2\rho(dt)\right\}z + az^2 + \int_{\mathbb{R}}\left(\frac{z^2}{1-tz} + \frac{tz}{1+t^2}\right)t^2\rho(dt) \\
&= \eta z + az^2 + \int_{\mathbb{R}}\left[\frac{z^2}{1-tz} + t^{-1}z1_{\mathbb{R}\setminus[-1,1]}(t)\right]t^2\rho(dt) \\
&= \eta z + az^2 + \int_{\mathbb{R}}\left[\frac{(tz)^2}{1-tz} + tz1_{\mathbb{R}\setminus[-1,1]}(t)\right]\rho(dt).
\end{aligned}$$

Further,

$$\begin{aligned}
\frac{(tz)^2}{1-tz} + tz1_{\mathbb{R}\setminus[-1,1]}(t) = \left[\frac{(tz)^2}{1-tz} + tz\right] - tz1_{[-1,1]}(t) &= \frac{tz}{1-tz} - tz1_{[-1,1]}(t) \\
&= \frac{1}{1-tz} - 1 - tz1_{[-1,1]}(t).
\end{aligned}$$

We conclude that

$$\mathcal{C}_\nu(z) = \eta z + az^2 + \int_{\mathbb{R}}\left[\frac{1}{1-tz} - 1 - tz1_{[-1,1]}(t)\right]\rho(dt). \qquad \textbf{[5.6]}$$

Clearly the above calculations may be reversed such that Eqs. **5.5** and **5.6** are equivalent, which proves *i*.

(*ii*) Suppose $\mu \in \mathcal{ID}(*)$ with generating pair (γ, σ) and generating triplet (a, ρ, η), the relationship between which is given by Eq. **5.3**. Then by definition of Λ, $\Lambda(\mu)$ has free generating pair (γ, σ), and the calculations in the proof of *i* [with ν replaced by $\Lambda(\mu)$] show that $\Lambda(\mu)$ has free generating triplet (a, ρ, η), as desired.

Apart from the striking similarity between Eqs. **5.2** and **5.4**, note that these particular representations clearly exhibit how μ (respectively ν) is always the convolution of a Gaussian distribution (respectively a semicircle distribution) and a distribution of generalized Poisson (respectively free Poisson) type (see also *The Lévy–Itô Decomposition* in ref. 1). In particular, the cumulant transform for the Gaussian distribution with mean η and variance a is $u \mapsto i\eta u - (1/2)au^2$, and the free cumulant transform for the semicircle distribution with mean η and variance a is $z \mapsto \eta z + az^2$.

As mentioned above, the bijection Λ is more than just a formal correspondence. Indeed, as just indicated, it follows easily from *Proposition 5.2* that Λ maps the Gaussian distributions onto the semicircle distributions. Furthermore, it was proved by Bercovici and Pata in ref. 2 that Λ actually preserves stability,[††] i.e. $\Lambda[\mathcal{S}(*)] = \mathcal{S}(\boxplus)$. When investigating the corresponding question for self-decomposability we realized that, in fact, Λ has the following algebraic properties.

5.3. Theorem. *The Bercovici–Pata bijection* Λ: $\mathcal{ID}(*) \to \mathcal{ID}(\boxplus)$ *satisfies*

(*i*) *If* $\mu_1, \mu_2 \in \mathcal{ID}(*)$, *then* $\Lambda(\mu_1 * \mu_2) = \Lambda(\mu_1) \boxplus \Lambda(\mu_2)$.
(*ii*) *If* $\mu \in \mathcal{ID}(*)$ *and* $c \in \mathbb{R}$, *then* $\Lambda(D_c\mu) = D_c\Lambda(\mu)$.

Proof:

(*i*) Suppose $\mu_1, \mu_2 \in \mathcal{ID}(*)$ with generating pairs (γ_1, σ_1), respectively (γ_2, σ_2). Then because the classical cumulant transform linearizes classical convolution, it follows immediately that $\mu_1 * \mu_2$ has generating pair $(\gamma_1 + \gamma_2, \sigma_1 + \sigma_2)$. By definition of Λ, $\Lambda(\mu_i)$ has free generating pair (γ_i, σ_i), $i = 1, 2$, and because the Voiculescu transform linearizes free convolution, it follows similarly that $\Lambda(\mu_1) \boxplus \Lambda(\mu_2)$ has free generating pair $(\gamma_1 + \gamma_2, \sigma_1 + \sigma_2)$. By definition of Λ, this means that *i* holds.

(*ii*) Assume that $\mu \in \mathcal{ID}(*)$ with generating triplet (a, ρ, η), and assume for simplicity that $c > 0$. Then for any u in $\mathbb{R}$,

$$C_{D_c\mu}(u) = C_\mu(cu) = i\eta(cu) - \frac{1}{2}a(cu)^2 + \int_{\mathbb{R}} [e^{i(cu)t} - 1 - i(cu)t1_{[-1,1]}(t)]\rho(dt)$$

$$= i(c\eta)u - \frac{1}{2}(c^2a)u^2 + \int_{\mathbb{R}} [e^{ius} - 1 - ius1_{[-c,c]}(s)]D_c\rho(ds)$$

$$= i\eta' u - \frac{1}{2}(c^2a)u^2 + \int_{\mathbb{R}} [e^{ius} - 1 - ius1_{[-1,1]}(s)]D_c\rho(ds),$$

where $\eta' = c\eta + \int_{\mathbb{R}} s[1_{[-1,1]}(s) - 1_{[-c,c]}(s)]D_c\rho(ds)$. Thus, the generating triplet for $D_c\mu$ is $(c^2a, D_c\rho, \eta')$. Note next that, by *Proposition 5.2,* $\Lambda(\mu)$ has free generating triplet (a, ρ, η), so by Eq. **3.1** we find for z in $\mathbb{C}^-$,

$$C_{D_c\Lambda(\mu)}(z) = C_{\Lambda(\mu)}(cz) = \eta(cz) + a(cz)^2 + \int_{\mathbb{R}} \left[\frac{1}{1 - t(cz)} - 1 - t(cz)1_{[-1,1]}(t)\right]\rho(dt)$$

$$= (c\eta)z + (c^2a)z^2 + \int_{\mathbb{R}} \left[\frac{1}{1 - sz} - 1 - sz1_{[-c,c]}(s)\right]D_c\rho(ds)$$

$$= \eta' z + (c^2a)z^2 + \int_{\mathbb{R}} \left[\frac{1}{1 - sz} - 1 - sz1_{[-1,1]}(s)\right]D_c\rho(ds),$$

with η' as above. Thus, the free generating triplet for $D_c\Lambda(\mu)$ is $(c^2a, D_c\rho, \eta')$, and hence by *Proposition 5.2,* $\Lambda(D_c\mu) = D_c\Lambda(\mu)$.

Together with the easily checked property that all Dirac measures are fixed points of Λ, *Theorem 5.3* shows that Λ preserves the affine structure on $\mathcal{ID}(*)$ and $\mathcal{ID}(\boxplus)$, which provides another explanation of the fact that Λ preserves stability and also shows that the same holds for self-decomposability, i.e. that $\Lambda[\mathcal{L}(*)] = \mathcal{L}(\boxplus)$. Indeed, suppose that $\mu \in \mathcal{L}(*)$ and that $c \in]0, 1[$. Then $\mu = D_c\mu * \mu_c$ for some probability measure μ_c. It is a well known fact that μ_c is automatically in $\mathcal{ID}(*)$ (see ref. 12) and hence by *Theorem 5.3,*

$$\Lambda(\mu) = \Lambda(D_c\mu * \mu_c) = D_c\Lambda(\mu) \boxplus \Lambda(\mu_c),$$

which shows that $\Lambda(\mu) \in \mathcal{L}(\boxplus)$. The same argumentation applies to the converse inclusion.

In ref. 3 we also studied the topological properties of Λ. Recall that a sequence (σ_n) of finite measures on $\mathbb{R}$ is said to converge weakly to a finite measure σ on $\mathbb{R}$ if $\int_{\mathbb{R}} f(s)\sigma_n(ds) \to \int_{\mathbb{R}} f(s)\sigma(ds)$ for any continuous bounded function $f : \mathbb{R} \to \mathbb{R}$. In that case we write $\sigma_n \xrightarrow{w} \sigma$.

[††] Bercovici and Pata actually proved an even stronger result, namely that Λ preserves the so-called partial domain of attraction.

5.4. Theorem. *The Bercovici–Pata bijection* $\Lambda : \mathcal{ID}(*) \to \mathcal{ID}(\boxplus)$ *is a homeomorphism w.r.t. weak convergence. In other words, if* (μ_n) *is a sequence of measures in* $\mathcal{ID}(*)$ *and* μ *is another measure in* $\mathcal{ID}(*)$*, then* $\mu_n \xrightarrow{w} \mu$ *if and only if* $\Lambda(\mu_n) \xrightarrow{w} \Lambda(\mu)$.

Proof (Sketch): For this particular result, the proof is simpler when expressed in terms of generating pairs rather than generating triplets. Thus for each n, let (γ_n, σ_n) be the generating pair for μ_n, and let (γ, σ) be the generating pair for μ. Then by a result of Gnedenko (ref. 12), we have that

$$\mu_n \xrightarrow{w} \mu \Leftrightarrow \gamma_n \to \gamma \text{ and } \sigma_n \xrightarrow{w} \sigma.$$

Note that (γ_n, σ_n) and (γ, σ) are also the free generating pairs for $\Lambda(\mu_n)$ and $\Lambda(\mu)$, respectively. Hence, in order to prove *Theorem 5.4,* we need a free version of the result of Gnedenko. This can be obtained by virtue of the description of weak convergence in terms of the Voiculescu transform, which was established by Bercovici and Voiculescu in ref. 11. Based on ref. 11, the proof only involves relatively standard measure-theoretic techniques, and we refer to ref. 3 for the details.

6. Concluding Remarks

In the sequel to this article (1), the results established above form the basis for a discussion of some aspects of free Lévy processes. The theory of such processes is in many ways analogous to, although at present far from as extensively developed as, the theory of Lévy processes in classical probability theory. We shall focus on (*i*) an integral representation of any freely self-decomposable operator as an integral with respect to a free Levy process, (*ii*) a free version of the Lévy–Itô representation of Levy processes in the classical sense, and (*iii*) a stochastic version of the Bercovici–Pata bijection.

We are grateful to the referee for helpful suggestions for an improved exposition of the material in the article. This work was supported by the Danish National Research Foundation (to MaPhySto, Centre for Mathematical Physics and Stochastics, and O.E.B.-N.) and the Danish Natural Science Research Council (to S.T.).

1. Barndorff-Nielsen, O. E. & Thorbjørnsen, S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 16576–16580.
2. Bercovici, H. & Pata, V. (1999) *Ann. Math.* **149,** 1023–1060.
3. Barndorff-Nielsen, O. E. & Thorbjørnsen, S. (2002) *Bernoulli* **8,** 323–366.
4. Rudin, W. (1991) *Functional Analysis* (McGraw–Hill, Singapore), 2nd Ed.
5. Voiculescu, D. V. (2000) in *Lectures on Probability Theory and Statistics* (Springer, Berlin), pp. 279–349.
6. Voiculescu, D. V. (1985) in *Operator Algebras and Their Connections with Topology and Ergodic Theory* (Springer, Berlin), pp. 556–588.
7. Voiculescu, D. V. (1991) *Invent. Math.* **104,** 201–220.
8. Voiculescu, D. V. (1990) in *Operator Algebras, Unitary Representations, Enveloping Algebras and Invariant Theory* (Birkhäuser, Boston), Vol. 92, pp. 45–60.
9. Voiculescu, D. V., Dykema, K. J. & Nica, A. (1992) *Free Random Variables* (Am. Math. Soc., Providence, RI), Vol. 1.
10. Voiculescu, D. V. (1986) *J. Funct. Anal.* **66,** 323–346.
11. Bercovici, H. & Voiculescu, D. V. (1993) *Indiana Univ. Math. J.* **42,** 733–773.
12. Gnedenko, B. V. & Kolmogorov, A. N. (1968) *Limit Distributions for Sums of Independent Random Variables* (Addison–Wesley, Reading, MA).
13. Feller, W. (1971) *An Introduction to Probability Theory and Its Applications* (Wiley, New York), Vol. 2.
14. Khintchine, A. Y. (1937) *Rec. Math. [Mat. Sbornik]* N.S. **2,** 79–120.
15. Bercovici, H. & Pata, V. (2000) *Proc. Am. Math. Soc.* **128,** 1011–1015.
16. Barndorff-Nielsen, O. E., Mikosch, T. & Resnick, S., eds. (2001) *Lévy Processes—Theory and Applications* (Birkhäuser, Boston).
17. Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions* (Cambridge Univ. Press, Cambridge, U.K.), Vol. 68.