Appendix B Positive definiteness

Positive-definite functions play a central role in statistics, approximation theory [Mic86,Wen05], and machine learning [HSS08]. They allow for a convenient Fourier-domain specification of characteristic functions, autocorrelation functions, and interpolation/approximation kernels (e.g., radial basis functions) with the guarantee that the underlying approximation problems are well-posed, irrespective of the location of the data points. In this appendix, we provide the basic definitions of positive definiteness and conditional positive definiteness in the multidimensional setting, together with a review of corresponding mathematical results. We distinguish between continuous functions on the one hand, and generalized functions on the other. We also give a self-contained derivation of Gelfand and Vilenkin's characterization of conditionally-positive definite generalized functions in one dimension and discuss its connection with the celebrated Lévy-Khintchine formula of statisticians. For a historical account of the rich topic of positive definiteness, we refer to [Ste76].

B.1 Positive definiteness and Bochner's Theorem

DEFINITION B.1 A continuous, complex-valued function f of the vector variable $\boldsymbol{\omega} \in \mathbb{R}^d$ is said to be *positive semi-definite* iff.

$$\sum_{m=1}^{N}\sum_{n=1}^{N} \xi_m \overline{\xi}_n f(\boldsymbol{\omega}_m - \boldsymbol{\omega}_n) \geq 0$$

for every choice of $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_N \in \mathbb{R}^d$, $\xi_1, \ldots, \xi_N \in \mathbb{C}$ and $N \in \mathbb{N}$. Such a function is called *positive definite in the strict sense* if the quadratic form is greater than 0 for all $\xi_1, \ldots, \xi_N \in \mathbb{C} \backslash \{0\}$.

In the sequel, we shall abbreviate "positive semi-definite" by *positive-definite*. This property is equivalent to the requirement that the $N \times N$ matrix $\mathbf{F}$ whose elements are given by $[\mathbf{F}]_{m,n} = f(\boldsymbol{\omega}_m - \boldsymbol{\omega}_n)$ is positive semi-definite (or, equivalently, nonnegative definite), for all N, no matter how the $\boldsymbol{\omega}_n$ are chosen.

The prototypical example of a positive-definite function is the Gaussian kernel $\mathrm{e}^{-\omega^2/2}$. To establish the property, we express this Gaussian as the Fourier transform of $g(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}$:

$$\begin{aligned}
\sum_{m=1}^{N}\sum_{n=1}^{N} \xi_m \overline{\xi}_n \mathrm{e}^{-\frac{(\omega_m-\omega_n)^2}{2}} &= \sum_{m=1}^{N}\sum_{n=1}^{N} \xi_m \overline{\xi}_n \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{j}(\omega_m-\omega_n)x} g(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}} \sum_{m=1}^{N}\sum_{n=1}^{N} \xi_m \overline{\xi}_n \mathrm{e}^{-\mathrm{j}(\omega_m-\omega_n)x} g(x)\,\mathrm{d}x \\
&= \int_{\mathbb{R}} \underbrace{\left|\sum_{m=1}^{N} \xi_m e^{-\mathrm{j}\omega_m x}\right|^2}_{\geq 0} \underbrace{g(x)}_{>0}\,\mathrm{d}x \geq 0
\end{aligned}$$

where we made use of the fact that $g(x)$, the (inverse) Fourier transform of $\mathrm{e}^{-\omega^2/2}$, is positive. It is not hard to see that above the argument remains valid for any (multidimensional) function $f(\boldsymbol{\omega})$ that is the Fourier transform of some nonnegative kernel $g(\boldsymbol{x}) \geq 0$. The more impressive result is that the converse implication is also true.

THEOREM B.1 (Bochner's Theorem) *Let f be a bounded continuous function on $\mathbb{R}^d$. Then, f is positive definite if and only if it is the (conjugate) Fourier transform of a nonnegative and finite Borel measure μ*

$$f(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x}\rangle} \mu(d\boldsymbol{x}).$$

In particular, Bochner's theorem implies that f is a valid characteristic function—that is, $f(\boldsymbol{\omega}) = \mathbb{E}\{\mathrm{e}^{\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x}\rangle}\} = \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x}\rangle} \mathscr{P}_X(\,\mathrm{d}\boldsymbol{x})$ where $\mathscr{P}_X$ is some probability measure on $\mathbb{R}^d$—if and only if f is continuous, positive definite with $f(\boldsymbol{0}) = 1$ (cf. Section 3.4.3 and Theorem 3.7).

Bochner's theorem is also fundamental to the theory of scattered data interpolation, although it requires a very slight restriction on the Fourier transform of f to ensure positive definiteness in the strict sense [Wen05].

THEOREM B.2 *A function $f : \mathbb{R}^d \to \mathbb{C}$ that is the (inverse) Fourier transform of a nonnegative, finite Borel measure μ is positive definite in the strict sense if there exists an open set $E \subseteq \mathbb{R}^d$ such that $\mu(E) \neq 0$.*

Proof Let $g(\boldsymbol{x}) \geq 0$ be the (generalized) density associated with μ such that $\mu(E) = \int_E g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ for any Borel set E. We then write $f(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x}\rangle} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and perform the same manipulation as for the Gaussian example above, which yields

$$\sum_{m=1}^{N}\sum_{n=1}^{N} \xi_m \overline{\xi}_n f(\boldsymbol{\omega}_m - \boldsymbol{\omega}_n)) = \int_{\mathbb{R}^d} \underbrace{\left|\sum_{m=1}^{N} \xi_m e^{-\mathrm{j}\langle \boldsymbol{\omega}_m, \boldsymbol{x}\rangle}\right|^2}_{\geq 0} \underbrace{g(\boldsymbol{x})}_{\geq 0}\,\mathrm{d}\boldsymbol{x} > 0$$

The key observation is that the zero set of the sum of exponentials $\sum_{m=1}^{N} \xi_m e^{-\mathrm{j}\langle \boldsymbol{\omega}_m, \boldsymbol{x}\rangle}$ (which is an entire function) has measure zero. Since the above integral involves positive terms only, the only possibility for it to be vanishing is that g be identically zero outside this zero set, which contradicts the assumption on the existence of E. $\square$

In particular, the latter constraint is verified whenever $f(\boldsymbol{\omega}) = \mathscr{F}\{g\}(\boldsymbol{\omega})$ where g is a continuous, nonnegative function with a bounded Lebesgue integral; i.e., $0 < \int_{\mathbb{R}^d} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty$. This kind of result is highly relevant to approximation and learning theory: Indeed, the choice of a strictly positive-definite interpolation kernel (or radial basis function) ensures that the solution of the generic scattered data interpolation problem is well defined and unique, no matter how the data centers are distributed [Mic86]. Here too, the prototypical example of a valid kernel is the Gaussian, which is (strictly) positive definite.

There is also an extension of Bochner's theorem for generalized functions that is due to Laurent Schwartz. In a nutshell, the idea is to replace each finite sum $\sum_{n=1}^{N} \xi_n f(\boldsymbol{\omega} - \boldsymbol{\omega}_n)$ by an infinite one (integral) $\int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega}')f(\boldsymbol{\omega} - \boldsymbol{\omega}')\,\mathrm{d}\boldsymbol{\omega}' = \int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega} - \boldsymbol{\omega}')f(\boldsymbol{\omega}')\,\mathrm{d}\boldsymbol{\omega}' = \langle f, \varphi(\cdot - \boldsymbol{\omega})\rangle$, which amounts to considering appropriate linear functionals of f over Schwartz' class of test functions $\mathscr{S}(\mathbb{R}^d)$. In doing so, the double sum in Definition B.1 collapses into a scalar product between f and the autocorrelation function of the test function $\varphi \in \mathscr{S}(\mathbb{R}^d)$, which leads to

$$(\varphi * \overline{\varphi}^{\vee})(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega}')\overline{\varphi(\boldsymbol{\omega}' - \boldsymbol{\omega})}\,\mathrm{d}\boldsymbol{\omega}'.$$

DEFINITION B.2 A generalized function $f \in \mathscr{S}'(\mathbb{R}^d)$ is said to be *positive-definite* if and only if, for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$,

$$\langle f, (\varphi * \overline{\varphi}^{\vee})\rangle \geq 0.$$

It can be shown that this is equivalent to Definition B.1 in the case where $f(\boldsymbol{\omega})$ is continuous.

THEOREM B.3 (Schwartz-Bochner Theorem) *A generalized function $f \in \mathscr{S}'(\mathbb{R}^d)$ is* positive-definite *if and only if it is the generalized Fourier transform of a nonnegative tempered measure μ; that is,*

$$\langle f, \hat{\varphi}\rangle = \langle \hat{f}, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(\boldsymbol{x})\mu(d\boldsymbol{x}).$$

The term "tempered measure" refers to a generic type of mildly-singular generalized function that can be defined by the Lebesgue integral $\int_{\mathbb{R}^d} \varphi(\boldsymbol{x})\mu(d\boldsymbol{x}) < \infty$ for all $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Such measures are allowed to exhibit polynomial growth at infinity subject to the restriction that they remain finite on any compact set.

The fact that the above form implies positive definiteness can be verified by direct substitution and application of Parseval's relation, by which we obtain

$$\langle f, (\varphi * \overline{\varphi}^{\vee})\rangle = \langle \hat{f}, |\hat{\varphi}|^2\rangle = \int_{\mathbb{R}^d} |\hat{\varphi}(x)|^2 \mu(d\boldsymbol{x}) \geq 0,$$

where the measurability property against $\mathscr{S}(\mathbb{R}^d)$ ensures that the integral is convergent (since $|\hat{\varphi}(x)|^2$ is rapidly decreasing).

The improvement over Theorem B.1 is that $\mu(\mathbb{R}^d)$ is no longer constrained to be finite. While this extension is of no direct help for the specification of characteristic functions, it happens to be quite useful for the definition of spline-like interpolation kernels that result in well-posed data fitting/approximation problems. We also note that the above definitions and results generalize to the infinite-dimensional setting (e.g., the Minlos-Bochner theorem which involves measures over topological vector spaces).

B.2 Conditionally positive-definite functions

DEFINITION B.3 A continuous, complex-valued function f of the vector variable $\boldsymbol{\omega} \in \mathbb{R}^d$ is said to be *conditionally positive-definite* of (integer) order $k \geq 0$ iff.

$$\sum_{m=1}^{N} \sum_{n=1}^{N} \xi_m \overline{\xi}_n f(\boldsymbol{\omega}_m - \boldsymbol{\omega}_n) \geq 0$$

under the condition

$$\sum_{n=1}^{N} \xi_n p(\boldsymbol{\omega}_n) = 0, \quad \text{for all } p \in \Pi_{k-1}(\mathbb{R}^d)$$

for all possible choices of $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_N \in \mathbb{R}^d$, $\xi_1, \dots, \xi_N \in \mathbb{C}$, and $N \in \mathbb{N}$, where $\Pi_{k-1}(\mathbb{R}^d)$ denotes the space of multidimensional polynomials of degree $(k-1)$.

This definition is also extendable for generalized functions using the line of thought that leads to Definition B.2. To keep the presentation reasonably simple and to make the link with the definition of the Lévy exponents in Section 4.2, we now focus on the one-dimensional case ($d = 1$). Specifically, we consider the polynomial constraint $\sum_{n=1}^{N} \xi_n \omega_n^m = 0, m \in \{0, \dots, k-1\}$ and derive the generic form of conditionally positive-definite generalized functions of order k, including the continuous ones which are of greatest interest to us.

The distributional counterpart of the kth-order constraint for $d = 1$ is the orthogonality condition $\int_{\mathbb{R}} \varphi(\omega)\omega^m \, \mathrm{d}\omega = 0$ for $m \in \{0, \dots, k-1\}$. It is enforced by restricting the analysis to the class of test functions whose moments up to order $(k-1)$ are vanishing. Without loss of generality, this is equivalent to considering some alternative test function $\mathrm{D}^k \varphi = \varphi^{(k)}$ where D^k is the kth derivative operator.

DEFINITION B.4 A generalized function $f \in \mathscr{S}'(\mathbb{R})$ is said to be *conditionally positive-definite* of order k iff. for all $\varphi \in \mathscr{S}(\mathbb{R})$

$$\left\langle f, (\varphi^{(k)} * \overline{\varphi^{(k)}}^{\vee}) \right\rangle = \left\langle f, (-1)^k \mathrm{D}^{2k} (\varphi * \overline{\varphi}^{\vee}) \right\rangle \geq 0.$$

This extended definition allows for the derivation of the corresponding version of Bochner's theorem which provides an explicit characterization of the family of conditionally positive-definite generalized functions, together with their generalized Fourier transform.

THEOREM B.4 (Gelfand-Villenkin) *A generalized function $f \in \mathscr{S}'(\mathbb{R})$ is conditionally* *positive-definite of order k if and only if it admits the following representation over $\mathscr{S}(\mathbb{R})$:*

$$\langle f, \hat{\varphi}\rangle = \langle \hat{f}, \varphi\rangle = \int_{\mathbb{R}\setminus\{0\}} \left(\varphi(x) - r(x)\sum_{n=0}^{2k-1} \frac{\varphi^{(n)}(0)}{n!} x^n\right)\mu(\mathrm{d}x) + \sum_{n=0}^{2k} a_n \frac{\varphi^{(n)}(0)}{n!}, \tag{B.1}$$

where μ is a positive tempered Borel measure on $\mathbb{R}\setminus\{0\}$ satisfying

$$\int_{|x|<1} |x|^{2k}\mu(\mathrm{d}x) < \infty.$$

Here, $r(x)$ is a function in $\mathscr{S}(\mathbb{R})$ such that $(r(x)-1)$ has a zero of order $(2k+1)$ at $x=0$, while the a_n are appropriate real-valued constants with the constraint that $a_{2k} \geq 0$.

Below, we provide a slightly adapted version of Gelfand and Vilenkin's proof which is remarkably concise and quite illuminating [GV64, Theorem 1, pp. 178], at least if one compares it with the standard derivation of the Lévy-Khintchine formula, which has a much more technical flavor (cf. [Sat94]), and is ultimately less general.

Proof Since $\left\langle f, (-1)^k \mathrm{D}^{2k}(\varphi * \overline{\varphi}^{\vee})\right\rangle = \left\langle (-1)^k\mathrm{D}^{2k} f, (\varphi * \overline{\varphi}^{\vee})\right\rangle$, we interpret Definition B.4 as the property that $(-1)^k\mathrm{D}^{2k} f$ is positive-definite. By the Schwartz-Bochner theorem, this is equivalent to the existence of a tempered measure ν such that

$$\left\langle (-1)^k \mathrm{D}^{2k} f, \hat{\varphi}\right\rangle = \left\langle f, (-1)^k \mathrm{D}^{2k}\hat{\varphi}\right\rangle = \langle \hat{f}, x^{2k}\varphi\rangle = \int_{\mathbb{R}} \varphi(x)\nu(\mathrm{d}x).$$

By defining $\phi(x) = x^{2k}\varphi(x)$, this can be rewritten as

$$\langle \hat{f}, \phi\rangle = \int_{\mathbb{R}} \frac{\phi(x)}{x^{2k}}\nu(\mathrm{d}x) = \langle f, \hat{\phi}\rangle,$$

where ϕ is a test function that has a zero of order $2k$ at the origin. In particular, this implies that $\lim_{\epsilon\downarrow 0}\int_{|x|<\epsilon} \frac{\phi(x)}{x^{2k}}\nu(\mathrm{d}x) = \frac{\phi^{(2k)}(0)}{(2k)!} a_{2k}$ where $a_{2k} \geq 0$ is the ν-measure at point $x = 0$. Introducing the new measure $\mu(\mathrm{d}x) = \nu(\mathrm{d}x)/x^{2k}$, we then decompose the Lebesgue integral as

$$\langle \hat{f}, \phi\rangle = \int_{\mathbb{R}\setminus\{0\}} \phi(x)\mu(\mathrm{d}x) + a_{2k}\frac{\phi^{(2k)}(0)}{(2k)!}, \tag{B.2}$$

which specifies f on the subset of test functions that have a $2k$th-order zero at the origin. To extend the representation to the whole space $\mathscr{S}(\mathbb{R})$, we associate to every $\varphi \in \mathscr{S}(\mathbb{R})$ the corrected function

$$\phi_{\mathrm{c}}(x) = \varphi(x) - r(x)\sum_{n=0}^{2k-1} \frac{\varphi^{(n)}(0)}{n!} x^n \tag{B.3}$$

with $r(x)$ as specified in the statement of the theorem. By construction, $\phi_{\mathrm{c}} \in \mathscr{S}(\mathbb{R})$ and has the $2k$th-order zero that is required for (B.2) to be applicable. By combining (B.2) and (B.3), we find that

$$\langle \hat{f}, \varphi\rangle = \int_{\mathbb{R}\setminus\{0\}} \phi_{\mathrm{c}}(x)\mu(\mathrm{d}x) + a_{2k}\frac{\phi_{\mathrm{c}}^{(2k)}(0)}{(2k)!} + \sum_{n=0}^{2k-1}\frac{\varphi^{(n)}(0)}{n!}\langle \hat{f}, r(x)x^n\rangle.$$

Next, we identify the constants $a_n = \langle \hat{f}, r(x)x^n \rangle$ and note that $\phi_c^{(2k)}(0) = \varphi^{(2k)}(0)$. The final step is to substitute these together with the expression (B.3) of ϕ_c in the above formula, which yields the desired result.

To prove the sufficiency of the representation, we apply (B.1) to evaluate the functional

$$\langle f, (\hat{\varphi}^{(k)} * \overline{\hat{\varphi}^{(k)}}^{\vee}) \rangle = \langle \hat{f}, x^{2k}|\varphi(x)|^2 \rangle = \int_{\mathbb{R}} x^{2k}|\varphi(x)|^2 \mu(\mathrm{d}x) + a_{2k}|\varphi(0)|^2 \geq 0,$$

where we have used the property that the derivatives of $x^{2k}|\varphi(x)|^2$ are all vanishing at the origin, except the one of order $2k$, which equals $(2k)!\,|\hat{\varphi}(0)|^2$ for $x = 0$. □

It is important to note that the choice of the function r is arbitrary as long as it fulfills the boundary condition $r(x) = 1 + O(|x|^{2k+1})$ as $x \to 0$, so as to regularize the potential kth-order singularity of μ at the origin, and that it decays sufficiently fast to temper the Taylor-series correction in (B.3) at infinity. If we compare the effect of using two different tempering functions r_1 and r_2, the modification is only in the value of the constants a_n, with $a_{n,2} - a_{n,1} = \langle \hat{f}, \big(r_2(x) - r_1(x)\big)x^n \rangle$. Another way of putting it is that the corresponding distributions $\hat{f}_1$ and $\hat{f}_2$ specified by the leading integral in (B.1) will only differ by a $(2k-1)$th-order point distribution that is entirely localized at $x = 0$; that is, $\hat{f}_2(x) - \hat{f}_1(x) = \sum_{n=0}^{2k-1} \frac{a_{n,2}-a_{n,1}}{n!}\delta^{(n)}(x)$, owing to the property that a_{2k} is common to both scenarios, or, equivalently, that the difference of their inverse Fourier transforms f_1 and f_2 is a polynomial of degree $(2k-1)$.

Thanks to Theorem B.4, it is also possible to derive an integral representation that is the kth-order generalization of the Lévy-Khintchine formula. For a detailed treatment of the multidimensional version of the problem, we refer to the works of Madych, Nelson, and Sun [MN90a, Sun93].

COROLLARY B.5 *Let $f(\omega)$ be a continuous function of $\omega \in \mathbb{R}$. Then, f is conditionally positive-definite of order k if and only if it can be represented as*

$$f(\omega) = \frac{1}{2\pi}\int_{\mathbb{R}\backslash\{0\}} \left(\mathrm{e}^{\mathrm{j}\omega x} - r(x)\sum_{n=0}^{2k-1} \frac{(\mathrm{j}\omega x)^n}{n!} \right) \mu(\mathrm{d}x) + \sum_{n=0}^{2k} a_n \frac{(\mathrm{j}\omega)^n}{n!}$$

where μ is a positive Borel measure on $\mathbb{R}\backslash\{0\}$ satisfying

$$\int_{\mathbb{R}} \min(|x|^{2k}, 1)\mu(\mathrm{d}x) < \infty,$$

where $r(x)$ and a_n are as in Theorem B.4.

The result is obtained by plugging $\varphi(x) = \frac{1}{2\pi} e^{\mathrm{j}\omega x} \longleftrightarrow \hat{\varphi}(\cdot) = \delta(\cdot - \omega)$ into (B.1), which is justifiable using a continuity argument. The key is that the corresponding integral is bounded when μ satisfies the admissibility condition, which ensures the continuity of $f(\omega)$ (by Lebesgue's dominated-convergence theorem), and vice versa.

B.3 Lévy-Khintchine formula from the point of view of generalized functions

We now make the link with the Lévy-Khintchine theorem of statisticians (cf. Section 4.2.1) which is equivalent to characterizing the functions that are conditionally positive-definite of order one. To that end, we rewrite the formula in Corollary B.5 for $k=1$ under the additional constraint that $f_1(0)=0$ (which fixes the value of a_0) as

$$\begin{aligned} f_1(\omega) &= a_0 + a_1 \mathrm{j}\omega - \frac{a_2}{2}\omega^2 + \frac{1}{2\pi}\int_{\mathbb{R}\setminus\{0\}} \left(\mathrm{e}^{\mathrm{j}\omega x} - r(x) - r(x)\mathrm{j}\omega x\right)\mu(\mathrm{d}x) \\ &= a_1 \mathrm{j}\omega - \frac{a_2}{2}\omega^2 + \int_{\mathbb{R}\setminus\{0\}} \left(\mathrm{e}^{\mathrm{j}\omega x} - 1 - r(x)\mathrm{j}\omega x\right) v(x)\,\mathrm{d}x \end{aligned}$$

where $v(x)\,\mathrm{d}x = \frac{1}{2\pi}\mu(\mathrm{d}x)$, $r(x) = 1 + O(|x|^3)$ as $x \to 0$ and $\lim_{x\to\pm\infty} r(x) = 0$. Clearly, the new form is equivalent to the Lévy-Khintchine formula (4.3) with the slight difference that the bias compensation is achieved by using a bell-shaped, infinitely-differentiable function r instead of the rectangular window $\mathbb{1}_{|x|<1}(x)$.

Likewise, we are able to transcribe the generalized Fourier-transform-pair relation (B.1) for the Lévy Khintchine representation (4.3), which yields

$$\begin{aligned} \langle \hat{f}_{\mathrm{L-K}}, \varphi\rangle &= \langle f_{\mathrm{L-K}}, \hat{\varphi}\rangle \\ &= \int_{\mathbb{R}\setminus\{0\}} \left(\varphi(x) - \varphi(0) - x\,\mathbb{1}_{|x|<1}(x)\varphi^{(1)}(0)\right) v(x)\,\mathrm{d}x + b_1'\varphi^{(1)}(0) + \frac{b_2}{2}\varphi^{(2)}(0). \end{aligned} \tag{B.4}$$

The interest of (B.4) is that it uniquely specifies the generalized Fourier transform of a Lévy exponent $f_{\mathrm{L-K}}$ as a linear functional of φ. We can also give a "time-domain" (or pointwise) interpretation of this result by distinguishing between three cases.

1) Lebesgue-integrable Lévy density $v \in L_1(\mathbb{R})$

Here, we are able to split the leading integral in (B.4) into its subparts, which results in

$$\hat{f}_{\mathrm{L-K}}(x) = v(x) - \delta(x)\left(\int_{\mathbb{R}} v(a)da\right) + \delta'(x)\left(b_1' - \int_{|a|<1} a v(a)\,\mathrm{d}a\right) + \delta''(x)\frac{b_2}{2}.$$

The underlying principle is that the so-defined generalized function will result in the same measurements as (B.4) when applied to the test function φ. In particular, the values of $\varphi^{(n)}$ at the origin are sampled using the Dirac distribution and its derivatives.

2) Non-integrable Lévy density ($v \notin L_1(\mathbb{R})$) with finite absolute moment $\int_{\mathbb{R}} |a| v(a)\,\mathrm{d}a < \infty$

To ensure that the integral in (B.4) is convergent, we need to retain the zero-order correction. Yet, we can still pull out the third term which results in the interpretation

$$\hat{f}_{\mathrm{L-K}}(x) = \mathrm{p.f.}\big(v\big) + \delta'(x)\left(b_1' - \int_{|a|<1} a v(a)\,\mathrm{d}a\right) + \delta''(x)\frac{b_2}{2},$$

where p.f. stands for the finite part operator that implicitly implements the Taylor-series adjustment that stabilizes the scalar-product integral $\langle v, \varphi \rangle$ (see Appendix A).

3) Non-integrable Lévy density and unbounded absolute moment $\int_{|\mathbb{R}} |a| v(a)\, \mathrm{d}a = \infty$

Here, we cannot split the integral anymore. In the particular case where $\int_{|a|>1} |a| v(a)\, \mathrm{d}a < \infty$, we can stabilize the integral by applying a full first-order Taylor-series correction. This leads to the finite-part interpretation

$$\hat{f}_{\mathrm{L-K}}(x) = \mathrm{p.f.}\big(v\big) + b_1 \delta'(x) + \frac{b_2}{2} \delta''(x),$$

which is the direct counterpart of (4.5). For $\int_{|a|>1} |a| v(a)\, \mathrm{d}a = \infty$, the proper pointwise interpretation becomes more delicate and it is safer to stick to the distributional definition (B.4).

The relevance of those results is that they properly characterize the impulse response of the infinitesimal semigroup generator G investigated in Section 9.7. Indeed, we have that $g(x) = \mathrm{G}\{\delta\}(x) = \mathscr{F}\{f\}(x)$, which is the generalized Fourier transform of the Lévy exponent f.