Fourier-based distances and Berry-Esseen like inequalities for smooth densities

Thierry GOUDON[1], Stéphane JUNCA[1,2] and Giuseppe TOSCANI[3]

[1]Labo. J.A. Dieudonné, UMR 6621
Université Nice-Sophia Antipolis, Parc Valrose
F-06108 Nice cedex 02

[2]IUFM Nice
Avenue Georges V, F-06100 Nice

[3]Dipartimento di Matematica
Universitá Pavia, Via Ferrata, 1
I-27100 Pavia

Abstract

This paper is devoted to the rate of convergence problem in the central limit theorem for sums of independent identically distributed random variables with regular probability density function. The method we use depends strictly on Fourier based metrics, and yields Berry–Esseen bounds for the convergence towards both a normal and a stable law in various Sobolev norms.

1 Introduction

In this paper, we investigate bounds for the rate of convergence towards a stable law in the central limit theorem with respect to some Fourier based metrics, for random variables with regular probability density function. In the first part, we shall consider the rate of convergence in the classical central limit theorem for sums of independent identically distributed random variables with finite variance, postponing to the second part extensions of the results to stable laws. For random variables with a smooth density, our method yields Berry-Esseen type convergence rates in any Sobolev space. This research is motivated by a recent study of Lions and Toscani [9], who developed a new method to control the convergence modes towards the Gaussian density in the central limit theorem for independent and equidistributed random variables with a smooth density. While their approach was mainly based on ideas both from information theory and classical kinetic theory of rarefied gases, in the spirit of the information-theoretic proof of the central limit theorem introduced by Linnik [8], in the final part of the paper they addressed the problem in a more classical way, by passing to Fourier transforms.

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent identically distributed random variables with values in the Euclidean space $\mathbb{R}^N$, $N \geq 1$. Assume that the common probability density f satisfies

$$\int_{\mathbb{R}^N} f\,dx = 1, \qquad \int_{\mathbb{R}^N} xf\,dx = 0, \qquad \int_{\mathbb{R}^N} |x|^2 f\,dx < \infty. \tag{1}$$

Let ω_K be the normal density with zero mean and covariance operator K, whose coefficients are

$$K_{ij} = \int_{\mathbb{R}^N} x_i x_j f(x)\,dx \qquad \text{for } 1 \leq i,j \leq N.$$

The classical central limit problem for densities relies in proving convergence of the probability density function $f_n(x)$ of the standardized sum

$$S_n = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n),$$

to the Gaussian density

$$\omega_K(x) = \frac{1}{(2\pi)^{N/2}\sqrt{\det(K)}}\exp\left(-\frac{K^{-1}x\cdot x}{2}\right)$$

as $n \to \infty$ (see [11] for instance). In their paper [9], Lions and Toscani proved convergence of f_n towards ω_K in various norms, showing that the more regular the density f is, the stronger the convergence of f_n is. In particular, under additional regularity assumptions on f (see condition (2) below), the convergence $f_n \to \omega_K$ holds in any Sobolev space

$$H^k(\mathbb{R}^N) = \left\{\phi \in \mathcal{S}'(\mathbb{R}^N),\ \int_{\mathbb{R}^N} (1+|\xi|^2)^k\,|\widehat{\phi}(\xi)|^2\,d\xi < \infty\right\},$$

where

$$\widehat{\phi}(\xi) = \int_{\mathbb{R}^N} e^{-ix\cdot\xi}\phi(x)\,dx$$

stands for the Fourier transform of ϕ. However, the problem of the rate of convergence in these spaces was not discussed. It is well–known, in view of the usual Berry-Esseen-type bounds (cf. [11, 12] and the references therein), that such kind of information may be expected from extra moment assumptions on f. In the first part of the paper we shall prove this quite intuitive result, showing that the rate of convergence in the central limit theorem for smooth densities follows under the same conditions of Theorem 5.1 of [9], if additional moments exist. Our main result is stated as follows.

Theorem 1 *Let f satisfy (1). Moreover, assume there exists $M > 0$ and $k \geq 0$ such that*

$$\int_{\mathbb{R}^N} (1 + |\xi|^2)^k \, |\widehat{f}(\xi)|^M \, d\xi = C_{M,k} < \infty \tag{2}$$

and

$$\int_{\mathbb{R}^N} |x|^s f(x) \, dx = \rho_s < \infty \text{ for some } 2 < s \leq 3. \tag{3}$$

Then

$$\|f_n - \omega_K\|_{H^k} \leq C \frac{1}{n^{(s-2)/2}},$$

the constant C depending on K, ρ_s, M and k.

As one can easily argue, condition (2) is fulfilled with $M = 2$ whenever f belongs to H^k. Furthermore, according to [9], it is worth pointing out that Theorem 1 holds in any H^k if, instead of satisfying condition (2), the probability density f is such that, for some $\varepsilon > 0$

$$|\xi|^\varepsilon \, |\widehat{f}(\xi)| \leq C \text{ for } |\xi| \geq 1. \tag{4}$$

Indeed, (4) yields (2) with $M > (2k + N)/\varepsilon$. Notice also that $\sqrt{f} \in H^1$ ensures (4) with $\varepsilon = 1$.

In the second part of the paper we extend our analysis to cover the case of stable limit theorems. To this aim, we need to determine the corresponding of condition (3). A simple criterion relies on finiteness of some relative pseudo-moment. Let $f_p, p \leq 2$, denote a p-stable law [6, 7]. In the normal case, corresponding to $p = 2$, $f_2 = \omega_K$, all moments are bounded, and the integrability of $x^\alpha(f - f_2)$ is equivalent to the integrability of $x^\alpha f$. This does not apply when $p < 2$. Hence, we will substitute (3) with the following

$$\int |x|^{p+\delta} \, |f(x) - f_p(x)| \, dx < \infty; \tag{5}$$

where $\delta \in]0, 1 - p]$ if $0 \leq p < 1$, or $\delta \in]0, 2 - p]$ if $1 \leq p < 2$. Actually, we will see that rates of convergence to the stable law f_p can be obtained for densities in a larger set, containing the f's satisfying (5).

Berry-Esseen type estimates are of vital interest in statistics, where many results about the asymptotic distribution of estimators depend on the central limit theorem, and the validity of the statistical method depends on how good the asymptotic distribution approximates the real distribution (see [6] for further details). For these reasons, the study of bounds for the rate of convergence in the central limit theorem has been the object of an intense activity of research in the last thirty years. We address the interested reader to the recent book by Rachev and Rüschendorf [12] for an almost complete list of references to this topic. Chapter 8 of this book is concerned with the rate of convergence in the central limit theorem with respect to the Kantorowich metric. As in (almost all) previous estimates of Berry–Esseen bounds, the proof is based on smoothing inequalities for this metric, in the spirit of the Bergström convolution method (cf. Senatov [13] for a similar proof of the multi-dimensional central limit theorem in the uniform metric). All assumptions concerning the domain of attraction and the order of convergence are described in [12] in terms of finiteness conditions for certain convolution-type metrics. For this reason, it is not easy to compare the domain of attraction in the Kantorowich metric with the normal domain of attraction of the stable law.

On the contrary, our analysis is simple and direct, taking advantage of the introduction of a suitable Fourier-based metric for probability distributions (see definition (6)). We shall introduce this metric and its properties in Section 2. The interest of this metric both in connection with the study of the asymptotic behavior of linear diffusions and with interpolation inequalities in Sobolev spaces is illustrated in Section 2 by means of two simple examples. The first relies in a amazing application to the heat equation, while the second deals with the behavior of f_n, where a slow rate of decay in Sobolev spaces can be found by means of the method developed by Carrillo and Toscani [2] for the study of the asymptotic behaviour of the Fokker-Planck equation. Combined to the results of [9], this method yields in a line convergence in any Sobolev space with an explicit rate (which is not sharp). Section 3 is devoted to the classical central limit theorem, and contains basically the proof of Theorem 1. The proof is essentially based on the splitting (in the Fourier space) of the H^k–norm on high and low frequencies. The regularity assumption (2) allows to control the decay of high frequencies, while the extra moment condition (3), combined with properties of the Fourier-based metric, gives estimates of the decay of low frequencies. In Section 4, we extend the discussion to stable limit theorems. Proceeding as in the proof of Theorem 1, we look for conditions that give estimates of the decay of low frequencies. This requires the understanding of various properties of the metric, as well as a detailed evaluation of subsets of the normal domain of attraction of a stable law. Finally, we show in the Appendix how the Fourier based metric can be used to provide a simple proof of the central limit theorem for stable laws, with an explicitly computable size of the distance in terms of the common distribution function.

2 Fourier-based metrics

2.1 Definition

Denote by $P_s(\mathbb{R}^N)$, $s > 0$, the class of all probability distributions F on $\mathbb{R}^N$, $N \geq 1$, such that

$$\int_{\mathbb{R}^N} |x|^s \, dF(x) < \infty.$$

We introduce a metric on $P_s(\mathbb{R}^N)$ by

$$d_s(F, G) = \sup_{\xi \in \mathbb{R}^N \setminus \{0\}} \left\{ \frac{|\widehat{f}(\xi) - \widehat{g}(\xi)|}{|\xi|^s} \right\}, \tag{6}$$

where f, g are the densities associated to F, G respectively. In the following, if X and Y have distributions F and G, or f and g are their probability densities we shall write $d_s(X, Y)$ or $d_s(f, g)$ instead of $d_s(F, G)$.

Connection between boundedness of these metrics and moment properties are easily understood. Let us write $s = m + \delta$, where m is an integer and $0 \leq \delta < 1$. Then, an easy computation shows that in order that $d_s(F, G)$ be finite, it suffices that F and G have the same moments up to order m.

The metric (6) has been introduced in [4] to investigate the trend to equilibrium of the solutions to the Boltzmann equation for Maxwell molecules. There, the case $s = 2 + \delta$, $\delta > 0$, was considered. Further applications of d_s, with $s = 4$, were studied in [1], while the case $s = 2$ has been considered in [14] in connection with the uniqueness of the solutions to the Boltzmann equation for Maxwell molecules. There, the main properties and the relationships with other classical metrics have been investigated. Here, we remark that, for all $s > 0$, d_s is a regular metric [15], i.e. for all random variables X, Y, Z such that Z is independent of (X, Y)

$$d_s(X + Z, Y + Z) \leq d_s(X, Y). \tag{7}$$

Moreover for any constant $a > 0$ and all random variables X and Y,

$$d_s(aX, aY) \leq a^s d_s(X, Y). \tag{8}$$

Properties (7) and (8) characterize ideal metrics [15] of order s.

This type of metrics proved to be a very useful instrument in the estimate of the convergence rate in the central limit theorem. It seems to us worth illustrating how this tool can be used with the following very simple examples. The former deals with the heat equation: by means of the metric (6) we shall find the rate of decay of the solution of the heat equation to the fundamental one, showing that this rate is related to the number of moments which are initially finite. The latter explains how we can use this metric to prove Berry-Esseen like inequalities with an easy proof of convergence in Sobolev spaces at a suboptimal rate.

2.2 A remark on the heat equation

It is well known that a solution of the heat equation

$$\begin{cases} \partial_t u = k\Delta_x u \quad x \in \mathbb{R}^N \\ u(0,x) = f(x), \end{cases} \tag{9}$$

where $f(x)$ is a probability density function satisfying (1), behaves asymptotically in time as the fundamental solution of (9), which is given by ω_{2kt}, and the rate of convergence in L^2 norm is governed by $t^{-N/2}$. By means of the Fourier based metric (6) it is immediate to show that this rate of convergence can be improved as soon as we have more information on the moments of the initial data.

Theorem 2 *Let u_1 and u_2 be solutions of (9) with initial data f_1 and f_2 respectively, and let s be such that $d_s(f_1, f_2)$ is finite . Then, we have*

$$\|(u_1 - u_2)(t,\cdot)\|^2_{L^2(\mathbb{R}^N)} \le d_s^2(f_1, f_2)\, C_{s,N}\, t^{-(s+N/2)},$$

where

$$C_{s,N} = \frac{\Gamma(s + N/2)}{\Gamma(N/2)} (4\pi)^{-N/2} (2k)^{-(s+N/2)}.$$

Proof. The proof is entirely explicit. Indeed, by Fourier transform, we get

$$\partial_t \widehat{u_i}(t,\xi) = -k|\xi|^2 \widehat{u_i}(t,\xi).$$

Solving the ordinary differential equation leads to

$$|\widehat{u_1}(t,\xi) - \widehat{u_2}(t,\xi)|^2 = \exp(-2kt|\xi|^2)|\widehat{f_1}(\xi) - \widehat{f_2}(\xi)|^2.$$

It follows that the L^2 norm is estimated by

$$\begin{aligned} \int_{\mathbb{R}^N} |u_1(t,x) - u_2(t,x)|^2\, dx &= (2\pi)^{-N} \int_{\mathbb{R}^N} |\widehat{u_1}(t,\xi) - \widehat{u_2}(t,\xi)|^2\, d\xi \\ &= (2\pi)^{-N} \int_{\mathbb{R}^N} \exp(-2kt|\xi|^2)|\widehat{f_1}(\xi) - \widehat{f_2}(\xi)|^2\, d\xi \\ &\le d_s^2(f_1, f_2)\, F_s(t) \end{aligned}$$

by using the Parseval equality. The function of time $F_s(t)$ is computed as follows

$$\begin{aligned}
F_s(t) &= (2\pi)^{-N} \int_{\mathbb{R}^N} \exp(-2kt|\xi|^2)|\xi|^{2s}\, d\xi \\
&= (2\pi)^{-N} \int_{S^{N-1}} \int_0^\infty \exp(-2kt\rho^2)\rho^{2s}\rho^{N-1}\, d\rho\, d\sigma \\
&= (2\pi)^{-N}\, 2\pi^{N/2}/\Gamma(N/2) \int_0^\infty \exp(-z) z^{(2s+N-2)/2} \frac{dz}{2.2kt\,(2kt)^{(2s+N-2)/2}} \\
&= \Gamma(s+N/2)/\left(\Gamma(N/2)(2k)^{s+N/2}(4\pi)^{N/2}\right) t^{-(s+N/2)}. \qquad (10)
\end{aligned}$$

Optimality of the obtained constant may be seen by considering Dirac mass as initial data. Note also that the result reduces to the usual $t^{-N/2}$ rate of convergence in L^2 as $s \to 0$. This finishes the proof. □

2.3 A non optimal rate of convergence in H^k norm

The interest of Fourier-based metric for studying rates of convergence of S_n is best understood when we realize that the Fourier transform of the density f_n is given by

$$\widehat{f_n}(\xi) = \left(\widehat{f}(\xi/\sqrt{n})\right)^n. \qquad (11)$$

This shows that the main contribution comes from small values of ξ, for which $\widehat{f_n}(\xi)$ is close to 1. Furthermore, provided $s > 2$, through a repeated application of property (8) we obtain

$$d_s(F_n, \Omega_K) \le \left(\frac{1}{\sqrt{n}}\right)^{s-2} d_s(F, \Omega_K). \qquad (12)$$

The rate of convergence in the metric d_s is clearly optimal. By means of (12), and requiring the initial density in a Sobolev space H^{k+p}, with $k, p > 0$, we can easily obtain convergence in lower order Sobolev spaces H^k at a suboptimal rate. The proof is immediate. Indeed, we wish to investigate the behaviour of the following (square of) H^k norm

$$I_n = \int_{\mathbb{R}^N} (1 + |\xi|^2)^k \left|\widehat{f_n}(\xi) - \widehat{\omega_K}(\xi)\right|^2 d\xi$$

as n goes to ∞. Let $p > 0$. For any given $R > 0$, we obtain an upper bound for I_n by splitting the integral in two parts,

$$\begin{aligned} I_n &\leq \int_{|\xi|\leq R} (1+|\xi|^2)^k \left|\widehat{f_n}(\xi) - \widehat{\omega_K}(\xi)\right|^2 d\xi \\ &\quad + \frac{1}{(1+R^2)^p} \int_{|\xi|>R} (1+|\xi|^2)^{k+p} \left|\widehat{f_n}(\xi) - \widehat{\omega_K}(\xi)\right|^2 d\xi \\ &\leq \int_{|\xi|\leq R} (1+|\xi|^2)^k |\xi|^{2s} d_s^2(f_n, \omega_K)\, d\xi + \frac{1}{(1+R^2)^p} \|f_n - \omega_K\|^2_{H^{k+p}} \\ &\leq C_N (1+R^2)^{k+s+N/2} d_s^2(f_n, \omega_K) + \frac{1}{(1+R^2)^p} \|f_n - \omega_K\|^2_{H^{k+p}}. \end{aligned}$$

Optimizing over R we get

$$\|f_n - \omega_K\|_{H^k} \leq c \|f_n - \omega_K\|_{H^{k+p}}^{\frac{2k+2s+N}{2k+2s+N+2p}} d_s(f_n, \omega_K)^{\frac{2p}{2k+2s+N+2p}} \tag{13}$$

where the constant c is explicitly computable. A simple application of (12) and (13) shows the following

Theorem 3 *Let f satisfy (1), (2) and (3). Then, if f_n is uniformly bounded in H^k, for some $k > 0$, f_n converges to ω_K in any $H^{k'}$ with $k' < k$, and*

$$\|f_n - \omega_K\|_{H^{k'}} \leq C \left(\frac{1}{n}\right)^{\frac{(k-k')(s-2)}{2k+N+2s}}.$$

Unlikely, this rate of convergence in $H^{k'}$ is heavily dependent of the regularity of the initial density, approaching the sharp exponent as soon as the initial density belongs to H^k, for all $k > 0$. Actually, we can improve the rate of convergence obtained in this simple and short way, by a careful investigation of another (quite classical) $n-$dependent splitting of I_n. This will be the object of the next Section.

3 Proof of Theorem 1

In the rest of the paper, we will study upper bounds of I_n, by considering separately integrals on high and low frequencies. To this aim, let us split I_n in two parts,

$$\begin{cases} I_n = H_n + L_n, \\ H_n = \displaystyle\int_{|\xi|\geq\eta\sqrt{n}} \ldots d\xi, \qquad L_n = \int_{|\xi|\leq\eta\sqrt{n}} \ldots d\xi \end{cases}$$

where $\eta > 0$ is a parameter to be determined. As we shall see, the regularity assumption (2) allows us to control the decay of high frequencies H_n, while the extra-moment condition (3) gives both estimates of the decay of low frequencies L_n, and of the rate of convergence.

3.1 High frequencies estimates

Let us first study the high frequencies integral. If the regularity assumption (2) is satisfied, this integral decreases with arbitrarily fast rate. No extra moment conditions are required for this result to hold.

Lemma 1 *Let f satisfy (1) and (2). Then, for any $\alpha > 0$, $n \geq M/2$, one has*

$$H_n \leq C/n^{\alpha}, \tag{14}$$

where C depends on $C_{M,k}$. If (2) is replaced by (4), then (14) holds in any H^k, $k \geq 0$.

Proof. Proceeding as in [9] we obtain

$$\begin{aligned} H_n &\leq 2\int_{|\xi|\geq\eta\sqrt{n}} (1+|\xi|^2)^k |\widehat{f_n}(\xi)|^2\, d\xi + 2\int_{|\xi|\geq\eta\sqrt{n}} (1+|\xi|^2)^k |\widehat{\omega_K}(\xi)|^2\, d\xi \\ &\leq 2\int_{|\xi|\geq\eta\sqrt{n}} (1+|\xi|^2)^k |\widehat{f}(\xi/\sqrt{n})|^{2n}\, d\xi + 2\int_{|\xi|\geq\eta\sqrt{n}} (1+|\xi|^2)^k |\widehat{\omega_K}(\xi/\sqrt{n})|^{2n}\, d\xi \end{aligned}$$

where we used (11) and the stability property $|\widehat{\omega_K}(\xi)|^2 = |\widehat{\omega_K}(\xi/\sqrt{n})|^{2n}$. Hence, it suffices to deal with

$$\widetilde{H}_n = \int_{|\xi|\geq\eta\sqrt{n}} (1+|\xi|^2)^k |\widehat{\varphi}(\xi/\sqrt{n})|^{2n}\, d\xi = n^{N/2} \int_{|\zeta|\geq\eta} (1+n|\zeta|^2)^k |\widehat{\varphi}(\zeta)|^{2n}\, d\zeta,$$

where φ verifies (1) and (2). We have

$$\begin{aligned} \widetilde{H}_n &= n^{N/2+k} \int_{|\zeta|\geq\eta} \left(\frac{(1+n|\zeta|^2)}{n(1+|\zeta|^2)}\right)^k (1+|\zeta|^2)^k |\widehat{\varphi}(\zeta)|^M\, |\widehat{\varphi}(\zeta)|^{2n-M}\, d\zeta, \\ &\leq n^{N/2+k} \left(\|\widehat{\varphi}\|_{L^\infty(|\zeta|\geq\eta)}\right)^{2n-M} C_{M,k}. \end{aligned}$$

Let $\alpha > 0$. We claim that

$$\lim_{n\to\infty} n^{N/2+k+\alpha} \left(\|\widehat{\varphi}\|_{L^\infty(|\zeta|\geq\eta)}\right)^{2n-M} = 0, \tag{15}$$

which will prove (14). Indeed, (15) follows from the following elementary result.

Lemma 2 *i) Let* $\psi \in L^1(\mathbb{R}^N, \mathbb{C})$ *satisfy* $\left| \int_{\mathbb{R}^N} \psi(x)\, dx \right| = \int_{\mathbb{R}^N} |\psi(x)|\, dx$. *Then there exists* $\theta \in (0, 2\pi)$ *such that* $\psi(x) = e^{-i\theta}|\psi(x)|$ *a.e.*

ii) Let $\varphi \in L^1(\mathbb{R}^N)$ *satisfy* $\varphi \geq 0$ *and* $\int_{\mathbb{R}^N} \varphi(x)\, dx = 1$. *Then, for all* $\eta > 0$, *one has* $\sup_{|\xi|\geq\eta} |\widehat{\varphi}(\xi)| < 1$.

Lemma 2 is very standard and we omit the proof. We only mention that *ii)* is a consequence of *i)*. Applying Lemma 2 *ii)* to $\varphi = f$ or ω_K gives $\|\widehat{\varphi}\|_{L^\infty(|\zeta|\geq\eta)} < 1$ for all $\eta > 0$, which in turn leads easily to (15). □

Remark 1 *Note that it is crucial to assume that the common density* f *lies in* $L^1(\mathbb{R}^N)$; *if* f *is a measure, Lemma 2 breaks down in general, as one can immediately verify choosing the Dirac mass* δ_0, *which gives* $\widehat{\delta_0} = 1$.

3.2 Low frequencies estimates

In order to deal with the low frequencies integral, we are going to use the Fourier based distance (6).

Lemma 3 *Let* $s > 2$. *If there exists* $\eta > 0$ *and* $\sigma > 0$ *such that, for* $|\xi| \leq \eta$,

$$|\widehat{f}(\xi)| \leq \exp\left\{-\frac{\sigma}{4}|\xi|^2\right\}, \tag{16}$$

then, for all $n \geq 2$, *one has*

$$\sup_{|\xi|\leq\eta\sqrt{n}} \left(\exp\left\{\frac{\sigma}{8}|\xi|^2\right\} \frac{\left|\widehat{f_n}(\xi) - \widehat{\omega_K}(\xi)\right|}{|\xi|^s} \right) \leq \left(\frac{1}{\sqrt{n}}\right)^{(s-2)} d_s(f, \omega_K).$$

Assume temporarily that the previous lemma holds true, and let us check that f satisfies the requirement (16). Indeed, the matrix K is clearly positive definite; we thus have $0 < \sigma I \leq K$, σ being the smallest eigenvalue of K. Hence

$$\widehat{\omega_K}(\xi) + \frac{\sigma}{8}|\xi|^2 \leq \exp\left\{-\frac{\sigma}{4}|\xi|^2\right\} \tag{17}$$

holds for $|\xi| \leq \epsilon$, $\epsilon > 0$ being sufficiently small. On the other hand, since f and ω_K have the same moments up to the order two, by expanding $\widehat{f}$, we get

$$\widehat{f}(\xi) = \widehat{\omega_K}(\xi) + o(|\xi|^2)$$

Combining this to (17), we deduce that there exists $0 < \eta < \epsilon$ such that

$$|\widehat{f}(\xi)| \leq \exp\left\{-\frac{\sigma}{4}|\xi|^2\right\}$$

holds for $|\xi| \leq \eta$.

Now, we can take advantage of the bound in Lemma 3 to estimate the decay of the low frequency integral and we obtain the announced rate of convergence. Indeed, by applying Lemma 3,

$$\begin{aligned}
L_n &= \int_{|\xi|\leq \eta\sqrt{n}} (1+|\xi|^2)^k \left|\widehat{f}_n(\xi) - \widehat{\omega_K}(\xi)\right|^2 d\xi \\
&\leq \left(\frac{1}{\sqrt{n}}\right)^{2(s-2)} d_s^2(f,\omega_K) \int_{|\xi|\leq \eta\sqrt{n}} (1+|\xi|^2)^k \exp\left\{-\frac{\sigma}{4}|\xi|^2\right\} |\xi|^{2s}\, d\xi \\
&\leq \frac{C}{n^{s-2}}\, d_s^2(f,\omega_K)
\end{aligned}$$

with $C := \int_{\mathbb{R}^N} |\xi|^{2s}(1+|\xi|^2)^k \exp\left\{-\frac{\sigma}{4}|\xi|^2\right\} d\xi$. Thus we proved

Lemma 4 *Let f satisfy (1) and (3). Then, one has*

$$L_n \leq C\left(\frac{1}{\sqrt{n}}\right)^{s-2},$$

where C depends on K, M, s and k.

To complete the proof of Theorem 1, it remains to prove Lemma 3.

Proof of Lemma 3. We write

$$\begin{aligned}
\frac{\left|\widehat{f}_n(\xi) - \widehat{\omega_K}(\xi)\right|}{|\xi|^s} &= \left(\frac{1}{\sqrt{n}}\right)^s \frac{\left|\widehat{f}(\xi/\sqrt{n})^n - \widehat{\omega_K}(\xi/\sqrt{n})^n\right|}{|\xi/\sqrt{n}|^s} \\
&= \left(\frac{1}{\sqrt{n}}\right)^s \frac{\left|\widehat{f}(\xi/\sqrt{n}) - \widehat{\omega_K}(\xi/\sqrt{n})\right|}{|\xi/\sqrt{n}|^s} \left|\sum_{k=0}^{n-1} \widehat{f}(\xi/\sqrt{n})^k \widehat{\omega_K}(\xi/\sqrt{n})^{n-1-k}\right|.
\end{aligned}$$

By hypothesis, if $|\xi| \le \eta\sqrt{n}$, we can estimate

$$
\begin{aligned}
\left|\sum_{k=0}^{n-1} \widehat{f}(\xi/\sqrt{n})^k \widehat{\omega_K}(\xi/\sqrt{n})^{n-1-k}\right| &\le \sum_{k=0}^{n-1} \left|\widehat{f}(\xi/\sqrt{n})^k \widehat{\omega_K}(\xi/\sqrt{n})^{n-1-k}\right| \\
&\le \sum_{k=0}^{n-1} \left\{\exp\left(-\frac{\sigma}{4}|\xi|^2\frac{k}{n}\right)\exp\left(-\frac{\sigma}{2}|\xi|^2\frac{n-1-k}{n}\right)\right\} \\
&\le n\exp\left(-\frac{\sigma}{4}|\xi|^2\frac{n-1}{n}\right).
\end{aligned}
$$

Since $\dfrac{n-1}{n} \ge \dfrac{1}{2}$ if $n \ge 2$, we get for $|\xi| \le \eta\sqrt{n}$,

$$
\frac{\left|\widehat{f}_n(\xi) - \widehat{\omega_K}(\xi)\right|}{|\xi|^s} \le \left(\frac{1}{\sqrt{n}}\right)^{s-2} \exp\left\{-\frac{\sigma}{8}|\xi|^2\right\} \frac{\left|\widehat{f}(\xi/\sqrt{n}) - \widehat{\omega_K}(\xi/\sqrt{n})\right|}{|\xi/\sqrt{n}|^s}
$$

which implies the result. □

Remark 2 *The evaluation of the optimal constant in the estimate of Theorem 1, remains open. On the other hand, optimality of the rate in the case $N = 1$ and $\delta = 1$ can be seen easily through a simple example. We set*

$$
\begin{cases} f(x) = \overline{\widehat{\phi}}(x) = \dfrac{1}{2\pi}\displaystyle\int_{\mathbb{R}} e^{ix\xi}\phi(\xi)\,d\xi, \\ \phi(\xi) = e^{-|\xi|^p}(1 + i\alpha\xi^3 + \alpha\xi^4) \end{cases}
$$

where $0 < \alpha << 1$. Obviously, moments of f up to order 2 are those of $e^{-x^2/2}$ and $f(x)$ takes value in $\mathbb{R}^+$ (since f behaves as $C\alpha e^{-x^2/4}x^4$ as $|x|$ is large, see [5] p. 382 and p. 1093, and, then, we choose α small enough). Clearly, discussion on optimality of the obtained decay relies on the study of the low frequencies integral L_n (since the other term decay faster). Therefore, optimality of the asserted rate of convergence follows from the following property

$$
\lim_{n\to\infty} \Gamma_n = \lim_{n\to\infty}\left(n\int_0^{n^{1/6}} \left|\phi^n(\xi/\sqrt{n}) - e^{-|\xi|^p}\right|^2 d\xi\right) = C > 0. \tag{18}
$$

Indeed, set $u = \alpha\dfrac{\xi^3}{n\sqrt{n}}\left(i + \dfrac{\xi}{\sqrt{n}}\right)$ and write

$$
\phi^n(\xi/\sqrt{n}) - e^{-|\xi|^p} = e^{-|\xi|^p}((1+u)^n - 1) = e^{-|\xi|^p}\, nu \int_0^1 (1+su)^{n-1}\,ds.
$$

Then, we get

$$
\begin{aligned}
\Gamma_n &= n \int_0^{n^{1/6}} \left| e^{-|\xi|^p} nu \int_0^1 (1+su)^{n-1}\, ds \right|^2 d\xi \\
&= \int_0^1 \exp(-2n^{1/3}\zeta^{2/3})\; n^3\; \alpha^2\; (\zeta/n)^2\; \left| i + (\zeta/n)^{1/3} \right|^2 \\
&\quad \left| \int_0^1 \left(1 + s\alpha(\zeta/n)\; \left[i + (\zeta/n)^{1/3} \right] \right)^{n-1} ds \right|^2 \frac{n^{1/6}}{3\zeta^{2/3}}\, d\zeta \\
&= \int_0^1 f_n(\zeta) g_n^2(\zeta)\, d\zeta
\end{aligned}
$$

by using the change of variable $\zeta = \xi^3/\sqrt{n}$ *and the notation*

$$
\begin{cases}
f_n(\zeta) = (\alpha^2/3)\; n^{7/6}\; \exp(-2n^{1/3}\zeta^{2/3})\; \zeta^{4/3} \left| i + (\zeta/n)^{1/3} \right|^2, \\
g_n(\zeta) = \displaystyle\int_0^1 \left(1 + s\alpha(\zeta/n)\; \left[i + (\zeta/n)^{1/3} \right] \right)^{n-1} ds.
\end{cases}
$$

Then, we can check easily that

$$
\begin{cases}
f_n(\zeta) \longrightarrow C\delta_{\zeta=0}, \qquad C = \alpha^2 \displaystyle\int_0^\infty e^{-2|\zeta|^2} |\zeta|^6 d\zeta, \\
g_n(\zeta) \longrightarrow \displaystyle\int_0^1 e^{i\alpha s\zeta} ds = g(\zeta) \qquad uniformly
\end{cases}
$$

as n *goes to infinity. It follows that*

$$
\Gamma_n \longrightarrow Cg(0) = C > 0
$$

which ends the proof of (18).

4 Stable laws

The aim of this section is to show how our method of proof extends easily to cover the rate of convergence problem in the case of stable distributions [3, 7, 6]. For the sake of simplicity, also in this section we restrict our attention to sums of independent and identically distributed random variables, and to strictly stable distributions. By a strictly p-stable random variable, we simply mean a random variable θ, associated to the characteristic function

$$
\Theta(\xi) = e^{-\sigma|\xi|^p} = \widehat{f}_p(\xi), \qquad 0 < p \leq 2, \qquad \sigma > 0. \tag{19}
$$

It is well–known that a Gaussian law satisfying (1) is a strictly stable distribution of order 2, while the Cauchy law is 1–stable [7]. The normal domain of attraction (NDA) of a p–stable law consists of the set of distributions F such that, when dealing with the the normalized sum

$$S_n = \frac{X_1 + ... + X_n}{n^{1/p}}$$

of the independent and identically distributed random variables X_i with common distribution function F, S_n converges to θ. The normal domain of attraction is characterized by requiring that (see [3]), for all $\xi \in \mathbb{R}^N$,

$$\lim_{n\to\infty} n\big(\widehat{f}(\xi/n^{1/p}) - 1\big) = -\sigma|\xi|^p.$$

Equivalent definitions of the NDA in terms of properties of the distribution function can be found in [7].

As in the classical problem for a normal law, informations on the rate of convergence and restriction of the set of NDA (in order to get a rate of convergence) depend on the choice of the metric. We refer to the recent book [12] for an exhaustive discussion of this aspect. According to the previous section, one still considers this problem in terms of the Fourier-based metric

$$d_p(f, g) = \sup_{\xi\neq 0} \frac{|\widehat{f}(\xi) - \widehat{g}(\xi)|}{|\xi|^p}, \quad p < 2.$$

Let us introduce the set

$$D_{\sigma,p} = \left\{ f \in \mathcal{P},\ \widehat{f}(\xi) = 1 - \sigma|\xi|^p + \psi_f(\xi), \text{ with } \frac{\psi_f(\xi)}{|\xi|^p} \in L^\infty(\mathbb{R}^N) \right\}. \tag{20}$$

The importance of such a set for the stable limit theorem is easily understood if one realizes that for the normal case ($p = 2$) the key point of the analysis was the expansion $\widehat{f}(\xi) = 1 - \sigma|\xi|^2 + \psi_f(\xi)$ with a precise control of the remainder ψ_f, which has to decay faster than $|\xi|^2$.

It is clear that the $d_p(f, g)$ is finite when f and g belong to the set $D_{\sigma,p}$. Notice also that the stable law f_p belongs to $D_{\sigma,p}$ (and actually the remainder goes to 0 faster than $|\xi|^p$). Moreover, if $g \in D_{\sigma,p}$ and $d_p(f, g) < \infty$, then $f \in D_{\sigma,p}$. Indeed,

$$\begin{aligned} \frac{|\widehat{f}(\xi) - 1 + \sigma|\xi|^p|}{|\xi|^p} &\le \frac{|\widehat{f}(\xi) - \widehat{g}(\xi)|}{|\xi|^p} + \frac{|\widehat{g}(\xi) - 1 + \sigma|\xi|^p|}{|\xi|^p} \\ &\le d_p(f, g) + \psi_g(\xi), \end{aligned}$$

where $\psi_g(\xi) \in L^\infty(\mathbb{R}^N)$.

We naturally introduce the following restriction of $D_{\sigma,p}$, that we call the Fourier domain of attraction (FDA)

$$D^0_{\sigma,p} = \Big\{ f \in D_{\sigma,p}, \quad \frac{\psi_f(\xi)}{|\xi|^p} \xrightarrow[\xi \to 0]{} 0 \Big\}.$$

We first compare convergence in $D^0_{\sigma,p}$ for the d_p metric to the usual convergence and FDA to NDA.

Lemma 5 *Let f_n be a sequence in $D^0_{\sigma,p}$ such that*

$$\begin{cases} \sup_n \dfrac{|\psi_n(\xi)|}{|\xi|^p} \xrightarrow[\xi \to 0]{} 0, \\ d_p(f_n, f) \xrightarrow[n \to \infty]{} 0 \quad \text{for some } f \in \mathcal{P}. \end{cases} \tag{21}$$

Then, $f_n \rightharpoonup f$ and $f \in D^0_{\sigma,p}$.

Proof. By definition, one has $|\widehat{f}_n(\xi) - \widehat{f}(\xi)| \le d_p(f_n, f)\ |\xi|^p$. This implies that $\widehat{f}_n(\xi)$ converges to $\widehat{f}(\xi)$ for $\xi \neq 0$ when $n \to \infty$. On the other hand, since f_n and f belong to $\mathcal{P}$, one has $\widehat{f}_n(0) = 1 = \widehat{f}(0)$. One deduces that $f_n \rightharpoonup f$.

Next, let $\varepsilon > 0$, $\xi \neq 0$. Then, provided $n \ge N(\varepsilon)$ is large enough, thanks to (21),

$$\begin{aligned} \frac{|\widehat{f}_n(\xi) - \widehat{f}(\xi)|}{|\xi|^p} &= \frac{|1 - \sigma|\xi|^p + \psi_n(\xi) - \widehat{f}(\xi)|}{|\xi|^p} \\ &= \left| \frac{1 - \sigma|\xi|^p - \widehat{f}(\xi)}{|\xi|^p} + \frac{\psi_n(\xi)}{|\xi|^p} \right| \\ &\le d_p(f_n, f) \le \varepsilon. \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} \frac{|\widehat{f}(\xi) - (1 - \sigma|\xi|^p)|}{|\xi|^p} &= \frac{|\psi_f(\xi)|}{|\xi|^p} = \frac{|\widehat{f}(\xi) - \widehat{f}_n(\xi) + \psi_n(\xi)|}{|\xi|^p} \\ &\le d_p(f_n, f) + \sup_n \frac{|\psi_n(\xi)|}{|\xi|^p} \le 2\varepsilon \end{aligned}$$

holds by choosing $n \ge N(\varepsilon)$ and then $|\xi| \le \eta(\varepsilon)$ to control the second term. This proves that $f \in D^0_{\sigma,p}$.

It is worth remarking that the uniform control on the remainder is necessary in order that the limit f belongs to $D^0_{\sigma,p}$. Indeed, assume that there exist $\varepsilon > 0$, a sequence ξ_n that tends to 0, and a subsequence, still labeled by n, such that $|\psi_n(\xi_n)|/|\xi_n|^p \geq 2\varepsilon$. Then, we are led to

$$\begin{aligned}
\frac{|\widehat{f}(\xi_n) - (1-\sigma|\xi_n|^p)|}{|\xi_n|^p} &= \frac{|\widehat{f}(\xi_n) - \widehat{f}_n(\xi_n) + \psi_n(\xi_n)|}{|\xi_n|^p} \\
&\geq \left| \frac{|\psi_n(\xi_n)|}{|\xi_n|^p} - \frac{|\widehat{f}(\xi_n) - \widehat{f}_n(\xi_n)|}{|\xi_n|^p} \right| \\
&\geq 2\varepsilon - d_p(f_n, f)
\end{aligned}$$

that remains $> \varepsilon$ for large enough n's. □

Our next result is concerned with the relationship between $D^0_{\sigma,p}$ and the normal domain of attraction of the stable law. We prove:

Lemma 6 *If $N > 1$, $D^0_{\sigma,p}$ is a subset of NDA. If $N = 1$, one actually has $D^0_{\sigma,p} = NDA$.*

Proof. Let $f \in D^0_{\sigma,p}$. Then, one has

$$\begin{aligned}
\sigma|\xi|^p + n\big(\widehat{f}(\xi/a_n) - 1\big) &= \sigma|\xi|^p - n\sigma\big(|\xi|/a_n\big)^p + n\psi(\xi/a_n) \\
&= \sigma|\xi|^p\Big(1 - \frac{n}{a_n^p}\Big) + \frac{\psi(\xi/a_n)}{\big(|\xi|/a_n\big)^p}\ \frac{n}{a_n^p}\ |\xi|^p \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

for any $\xi \in \mathbb{R}^N\backslash\{0\}$ and any sequence $a_n \sim n^{1/p}$. Therefore $f \in NDA$.

Conversely, for $f \in NDA$, we set $\psi(\xi) = \widehat{f}(\xi) - 1 + \sigma|\xi|^p$, so that $n\psi(\xi/n^{1/p}) = |\xi|^p \frac{\psi(\xi/n^{1/p})}{|\xi|^p/n}$ tends to 0, for any $\xi \neq 0$ when $n \to \infty$. Therefore, the problem reduces to deduce from $\lim_{n\to\infty} g(s/n) = 0$ for any $s > 0$ that $\lim_{s\to 0} g(s) = 0$. Of course this is false in general, but here one knows that g is continuous on $\mathbb{R}^{*+}$ since it is obtained from the Fourier transform of a probability measure. Suppose that $g(s)$ does not tend to 0 when s vanishes: there exist $\varepsilon > 0$ and a sequence s_n such that $\lim_{n\to 0} s_n = 0$ while $|g(s_n)| > \varepsilon$. Since g is continuous in $s_1 > 0$, there exists $\gamma_1 > 0$ such that $|g(x)| > \varepsilon/2$ for any $x \in [s_1; s_1 + \gamma_1] = F_1$. Set $n_1 = 1 = k_1$, so that for all $x \in F_1$ we have $|g(x/n_1)| \geq \varepsilon/2$. Next, by using $\lim_{n\to\infty} s_n = 0$, we can find some $s_{k_2} \in]0, \min(s_{k_1}, \gamma_1)[$. One has $|g(s_{k_2})| > \varepsilon$ and, by continuity, there exists some $\gamma_2 > 0$ such that $|g(x)| > \varepsilon/2$ for any $x \in [s_{k_2}; s_{k_2} + \gamma_2]$. Furthermore, $\mathbb{R}$ being archimedian, there exists an integer $n_2 = \big[s_{k_1}/s_{k_2}\big]$ verifying $s_{k_1} \leq n_2 s_{k_2} \leq s_{k_1} + s_{k_2} \leq s_{k_1} + \gamma_1$. Hence, we can choose $\gamma_2 > 0$ small enough to guarantee that $s_{k_1} \leq n_2 s_{k_2} \leq n_2(s_{k_2} + \gamma_2) < s_{k_1} + \gamma_1$. Then, we set

$F_2 = F_1 \cap [n_2 s_{k_2}; n_2(s_{k_2} + \gamma_2)]$. Notice that the interior of F_2 is non empty and for any $x \in F_2$, one has $|g(x/n_2)| \geq \varepsilon/2$. Repeating this reasoning, one constructs a decreasing subsequence $\left(s_{k_\ell}\right)_{\ell \in \mathbb{N}^*}$ which tends to 0 as $\ell \to \infty$, a sequence of positive reals γ_ℓ, a sequence of closed intervals F_ℓ having non empty interiors and verifying $F_{\ell+1} \subset F_\ell$, and a sequence of integers n_ℓ (given by $\left[n_{\ell-1} s_{k_{\ell-1}}/s_{k_\ell}\right]$) that increases to $+\infty$ as $\ell \to \infty$, such that for all $\ell \in \mathbb{N}^*$, all $x \in F_\ell$, $|g(x/n_\ell)| \geq \varepsilon/2$. However $F = \bigcap_\ell F_\ell$ is a non empty interval. Thus, for $x \in F$, one has $|g(x/n_\ell)| \geq \varepsilon/2$ that contradicts the assumption $\lim_{n\to 0} g(x/n) = 0$. $\square$

The rate of convergence problem for a p-stable law, will be discussed by introducing a suitable subset of the set $D^0_{\sigma,p}$. To this extent, we consider, for $\delta > 0$, the set

$$D^\delta_{\sigma,p} = \{f \in D^0_{\sigma,p},\ |\psi(\xi)|/|\xi|^p \leq C_f\ |\xi|^\delta\}.$$

Let f_p be the density such that $\widehat{f_p}(\xi) = e^{-\sigma|\xi|^p}$. It is worth noticing that f_p belongs to $D^\delta_{\sigma,p}$ for any $0 \leq \delta \leq p$. Furthermore, for any $f \in D^\delta_{\sigma,p}$, one has $d_{p+\delta}(f, f_p) \leq C < \infty$. According to Lemma 5, we also remark that, if f_n is bounded in $D^\delta_{\sigma,p}$ and $d_p(f_n, f)$ tends to 0 for some $f \in \mathcal{P}$, then $f_n \rightharpoonup f$ and $f \in D^\delta_{\sigma,p}$.

One naturally asks for a easy-to-check criterion ensuring that a given density f lies in $D^\delta_{\sigma,p}$. A simple criterion relies on finiteness of some pseudo-moment. For the normal case $p = 2$, f_p has all moments finite, so that integrability of $x^\alpha(f - f_p)$ is equivalent to integrability for $x^\alpha\ f$. This of course does not apply when $p < 2$. Note also that, for $p = 2$, the set considered in the previous section satisfies

$$\left\{f \in \mathcal{P}, \int |x|^{2+\delta}\, f(x)\, dx < \infty\right\} \subset D^\delta_{\sigma,2}$$

with a strict inclusion.

Lemma 7 *Let $0 \leq p < 1$ and $0 < \delta \leq 1 - p$. Then, the following embedding holds*

$$\left\{f \in \mathcal{P}, \int |x|^{p+\delta}\, |f(x) - f_p(x)|\, dx < \infty\right\} \subset D^\delta_{\sigma,p}.$$

Let $1 \leq p < 2$ and $0 \leq \delta \leq 2 - p$. Then

$$\left\{f \in \mathcal{P}, \int |x|^{p+\delta}\, |f(x) - f_p(x)|\, dx < \infty,\ \int x\, (f(x) - f_p(x))\, dx = 0\right\} \subset D^\delta_{\sigma,p}.$$

Proof. The result of the lemma is a consequence of a quite simple property of Fourier transforms. Let $g \in L^1(\mathbb{R}^N)$ be such that $|x|^\gamma g \in L^1(\mathbb{R}^N)$ for a given constant γ, with $0 < \gamma < 1$, then

$$\widehat{g}(\xi) = \widehat{g}(0) + \psi_g(\xi), \qquad \text{with} \quad \psi_g(\xi)/|\xi|^\gamma \xrightarrow[\xi \to 0]{} 0. \tag{22}$$

If $\gamma = 1$, then $\widehat{g}$ belongs to C^1 and has bounded gradient. In this case we can write an expansion up to the first order, in which the remainder is only bounded. Applying the property (22) with $\gamma = p + \delta$ and $g = f - f_p$ gives the first embedding. The second one follows from the same reasoning, by considering an higher order expansion, since now we deal with C^1 functions.

The proof of (22) relies on the following simple remark

$$\begin{aligned}
\frac{|\widehat{g}(\xi) - \widehat{g}(0)|}{|\xi|^\delta} &= \left| \int_{\mathbb{R}} g(x) |x|^\delta \, \frac{e^{-ix\cdot\xi} - 1}{(|x| \; |\xi|)^\delta} \, dx \right| \\
&\leq \left| \int_{|x||\xi| \leq R} |g(x)| \; |x|^\delta \, \frac{e^{-ix\cdot\xi} - 1}{(|x| \; |\xi|)^\delta} \, dx \right| + \left| \int_{|x||\xi| \geq R} |g(x)| \; |x|^\delta \, \frac{e^{-ix\cdot\xi} - 1}{(|x| \; |\xi|)^\delta} \, dx \right| \\
&\leq \int_{|x| \leq R/|\xi|} |g(x)| \; |x|^\delta \, \frac{|e^{-ix\cdot\xi} - 1|}{(|x| \; |\xi|)^\delta} \, dx + \frac{2}{R^\delta} \int_{\mathbb{R}^N} |g(x)| \; |x|^\delta \, dx.
\end{aligned}$$

Let $\varepsilon > 0$. Choose $R = R(\varepsilon)$ so that the second term becomes $\leq \varepsilon$. Then, by using the dominated convergence theorem, one sees that there exists $\eta = \eta(\varepsilon) > 0$ so that for $|\xi| \leq \eta(\varepsilon)$, the first term is also $\leq \varepsilon$. In turn, one gets

$$\frac{|\widehat{g}(\xi) - \widehat{g}(0)|}{|\xi|^\delta} \leq 2\varepsilon$$

provided $|\xi| \leq \eta(\varepsilon)$. □

Lemma 8 *If $d_{p+\delta}(f, f_p) < \infty$ with $\delta + p > 1$, then f has a vanishing first moment, in the sense that*

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}^N} x f(x) \; e^{-\varepsilon x^2/2} \, dx = 0.$$

Proof. By a parity argument, the integral of $x f_p(x) \; e^{-\varepsilon x^2/2}$ vanishes. Therefore, we will show the result for $g = f - f_p$, that fulfills $|\widehat{g}(\xi)| \leq C|\xi|^s$ for $s = p + \delta > 1$. Of course, this gives $\widehat{g}(0) = 0 = \int g(x) \, dx$ but we also have

$$\frac{|\widehat{g}(\xi) - \widehat{g}(0)|}{|\xi|} \leq C|\xi|^{s-1} \xrightarrow[|\xi| \to 0]{} 0,$$

which says that $\widehat{g}$ is derivable in 0 with $\nabla\widehat{g}(0)=0$. Then, we use the relation $\widehat{ixg}=\nabla(\widehat{g})$ which holds in $\mathcal{S}'$ with the mollifiers $\varphi_\varepsilon(\xi)=(2\pi\varepsilon)^{-N/2}\ e^{-\xi^2/(2\varepsilon)}$ as test functions. One gets

$$\begin{aligned}
\langle\widehat{ixg},\varphi_\varepsilon\rangle &= \int_{\mathbb{R}^N} ixg(x)\ \widehat{\varphi_\varepsilon}(x)\,dx=\int_{\mathbb{R}^N} ixg(x)\ (2\pi)^{-N/2}\ e^{-\varepsilon x^2/2}\,dx \\
&= \langle\widehat{g},\nabla\varphi_\varepsilon\rangle=\int_{\mathbb{R}^N}\widehat{g}(\xi)\ (2\pi\varepsilon)^{-N/2}\ e^{-\xi^2/(2\varepsilon)}\ (-\xi/\varepsilon)\,d\xi \\
&= \int_{\mathbb{R}^N}\widehat{g}(\sqrt{\varepsilon}\zeta)\ (2\pi)^{-N/2}\ e^{-\zeta^2/2}\ (-\zeta/\sqrt{\varepsilon})\,d\zeta.
\end{aligned}$$

The previous integral can be estimated by

$$C\ \sqrt{\varepsilon}^{p+\delta-1}\quad\int_{\mathbb{R}^N}|\zeta|^{p+\delta+1}\ (2\pi)^{-N/2}\ e^{-\zeta^2/2}\,d\zeta=C^{st}\ \varepsilon^{(s-1)/2},$$

with $s>1$. □

We are now in a position to prove the rate of convergence result for the p-stable law. Clearly we can write the law as

$$\widehat{f}_p(\xi)=e^{-\sigma|\xi|^p}=1-\sigma|\xi|^p+\psi_p(\xi),\qquad \psi_p(\xi)/|\xi|^p\xrightarrow[\xi\to 0]{}0.$$

Then, for $|\xi|\leq\varepsilon$ small enough, one obtains

$$\widehat{f}_p(\xi)+\sigma|\xi|^p/8=1-7\sigma|\xi|^p/8+\psi_p(\xi)\leq 1-\sigma|\xi|^p/4\leq e^{-\sigma|\xi|^p/4}.$$

Let us consider $f\in D^0_{\sigma,p}$. We write

$$\widehat{f}(\xi)=1-\sigma|\xi|^p+\psi_f(\xi)=\widehat{f}_p(\xi)+r_f(\xi),$$

where $r_f(\xi)/|\xi|^p\xrightarrow[\xi\to 0]{}0$. Thus, we can choose $0<\eta<\varepsilon$ sufficiently small to ensure that

$$|\widehat{f}(\xi)|\leq\exp\left\{-\frac{\sigma}{4}|\xi|^p\right\}$$

holds whenever $|\xi|\leq\eta$. To this point, we simply repeat the proof of Lemma 3, which generalizes to give the following upper bound

$$\frac{|\widehat{f}_n(\xi)-\widehat{f}_p(\xi)|}{|\xi|^{p+\delta}}\leq\frac{1}{n^{\delta/p}}\ d_{p+\delta}(f,f_p)\ \exp\left\{-\frac{\sigma}{8}|\xi|^p\right\},\tag{23}$$

for all $n \geq 2$, and $|\xi| \leq \eta n^{1/p}$. From (23) one obtains the decay estimate on the low frequencies integral. In fact,

$$\begin{aligned} L_n &= \int_{|\xi|\leq \eta n^{1/p}} (1+|\xi|^2)^k \left| \widehat{f}_n(\xi) - \widehat{f}_p(\xi) \right|^2 d\xi \\ &\leq \frac{1}{n^{2\delta/p}} d^2_{p+\delta}(f, f_p) \int_{|\xi|\leq \eta n^{1/p}} (1+|\xi|^2)^k \exp\left\{-\frac{\sigma}{4}|\xi|^p\right\} |\xi|^{2(p+\delta)} d\xi \\ &\leq \frac{C}{n^{2\delta/p}} d^2_{p+\delta}(f, f_p), \end{aligned}$$

where

$$C := \int_{\mathbb{R}^N} |\xi|^{2(p+\delta)} (1+|\xi|^2)^k \exp\left\{-\frac{\sigma}{4}|\xi|^p\right\} d\xi.$$

On the other hand, *mutatis mutandis*, the high frequency integral can be estimated in the same way of the Gaussian case. We finally obtain the main result of this section.

Theorem 4 *Let f be a density function belonging to the normal domain of attraction of a strictly p-stable law. Moreover, let f satisfy (2), and $f \in D^\delta_{\sigma,p}$ for some $\delta \in]0,p]$. Then $\|f_n - f_p\|_{H^k} \leq C\, n^{-\delta/p}$, the constant C depending on p, δ, M and k.*

Appendix

We present in this Appendix further results on the Fourier-based metric, including a proof of the central limit theorem for stable laws which allows for the explicit control of the distance. Analogous computations for the Gaussian case have been done in [14]. For this reason, we skip most of the proofs, since they result in simple extensions of the case $p = 2$.

Let $\{X_0, Y_0\}$ and $\{X_1, Y_1\}$ be two independent pairs of random variables such that $\{X_0, Y_0\} \in D_{\sigma_0,p}$, $\{X_1, Y_1\} \in D_{\sigma_1,p}$ for some constants σ_0 and σ_1. Let F_i (resp. G_i) denote the law of X_i (resp. Y_i). For $0 < \lambda < 1$, $0 < p \leq 2$, denote by F_λ the law of

$$X_\lambda = \lambda^{1/p} X_0 + (1-\lambda)^{1/p} X_1$$

and G_λ is similarly defined from Y_0, Y_1.

Lemma 9 *For any $0 < p \leq 2$, d_p is superadditive with respect to the rescaled convolution, i.e.*

$$d_p(F_\lambda, G_\lambda) \leq \lambda d_p(F_0, G_0) + (1-\lambda) d_p(F_1, G_1). \tag{24}$$

Corollary 1 *For any* $0 < p \leq 2$, *the rescaled convolution is continuous with respect to* d_p : *with obvious notations, if* $d_p(F_0^n, G_0) \longrightarrow 0$ *and* $d_p(F_1^n, G_1) \longrightarrow 0$, *then* $d_p(F_\lambda^n, G_\lambda) \longrightarrow 0$.

Proof. The proof is an easy adaptation of Theorem 3 of [14].

□

Let X be a random variable with law $F \in D_{\sigma,p}$ for some $\sigma > 0$. Let us consider the functional

$$J(X) = J(F) = d_p(F, F_p) = \sup_{\xi \neq 0} \frac{|\widehat{f}(\xi) - \widehat{f}_p(\xi)|}{|\xi|^p}, \tag{25}$$

where this time F_p is the distribution of the strictly p-stable law in $D_{\sigma,p}$. The same proof as before shows the

Lemma 10 *Let* $0 < p \leq 2$. *Consider any independent random variables* $X_0 \in D^0_{\sigma_0,p}$ *and* $X_1 \in D^0_{\sigma_1,p}$ *for some constants* σ_0 *and* σ_1. *Let* $0 < \lambda < 1$. *One has*

$$J\left(\lambda^{1/p} X_0 + (1-\lambda)^{1/p} X_1\right) \leq \lambda J(X_0) + (1-\lambda) J(X_1) \tag{26}$$

with equality if and only if both X_0 *and* X_1 *are* p*-strictly stable laws (belonging to* $D_{\sigma_0,p}$ *and* $D_{\sigma_1,p}$ *respectively).*

Proof. Suppose that there is equality in Lemma 9, with G_0 and G_1 replaced by the corresponding p-strictly stable laws, denoted by $F_{p,0}, F_{p,1}$. Let us denote by g_0 and g_1 the densities of $F_{p,0}$ and $F_{p,1}$. Let $\left(\xi_n\right)_{n\in\mathbb{N}}$ such that

$$\frac{|\widehat{f}_0(\lambda^{1/p}\xi_n)\widehat{f}_1((1-\lambda)^{1/p}\xi_n) - \widehat{g}_0(\lambda^{1/p}\xi_n)\widehat{g}_1((1-\lambda)^{1/p}\xi_n)|}{|\xi_n|^p} \xrightarrow[n\to\infty]{} J(\lambda^{1/p} X_0 + (1-\lambda)^{1/p} X_1).$$

The left-hand side is bounded by

$$\left\{(1-\lambda)|\widehat{f}_0(\lambda^{1/p}\xi_n)| \frac{|\widehat{f}_1((1-\lambda)^{1/p}\xi_n) - \widehat{g}_1((1-\lambda)^{1/p}\xi_n)|}{((1-\lambda)^{1/p}|\xi_n|)^p} \right.$$
$$\left. + \lambda|\widehat{g}_1((1-\lambda)^{1/p}\xi_n)| \frac{|\widehat{f}_0(\lambda^{1/p}\xi_n) - \widehat{g}_0(\lambda^{1/p}\xi_n)|}{(\lambda^{1/p}|\xi_n|)^p}\right\}.$$

Suppose that $|\xi_n|$ tends to ∞; then $|\widehat{g_1}((1-\lambda)^{1/p}\xi_n)| \leq 1/2$ for large n, and $J(\lambda^{1/p} X_0 + (1-\lambda)^{1/p} X_1) = \lambda J(X_0) + (1-\lambda)J(X_1) \leq (1-\lambda)J(X_1) + \lambda/2J(X_0)$, which is impossible when $J(X_0) \neq 0$. Therefore, extracting a subsequence if necessary, we may assume that $\xi_n \longrightarrow \xi \in \mathbb{R}^d$. The same argument shows then that $\xi = 0$. But by definition,

$$\frac{|\widehat{f_0}(\lambda^{1/p}\,\xi_n) - \widehat{g_0}(\lambda^{1/p}\,\xi_n)|}{(\lambda^{1/p}|\xi_n|)^p} \longrightarrow 0$$

as $n \to \infty$, and a similar property holds with f_1, g_1 since one assumes that $f_i \in D^0_{\sigma_i,p}$. Hence, $J(X_\lambda) = 0$, and $J(X_0) = J(X_1) = 0$.

□

Following ideas from Murata and Tanaka [10] and others, one realizes that a functional having these properties can be used to prove a central limit theorem for stable laws.

Theorem 5 *Let $0 < p < 2$, and let (X_n) be a sequence of independent identically distributed variables in $D^0_{\sigma,p}$. Then, if*

$$S_n = \frac{1}{n^{1/p}}(X_1 + \cdots + X_n), \tag{27}$$

and F_n denotes the distribution function of S_n, then

$$d_p(F_n, F_p) \longrightarrow 0 \quad \text{as } n \to \infty. \tag{28}$$

Moreover, given $\varepsilon > 0$, knowing (in the Fourier space) the remainder $\psi(\xi)$ of the common distribution function, one can compute explicitly n_0 such that for $n \geq n_0$, $d_p(F_n, F_p) \leq \varepsilon$.

Proof. Let $f \in D^0_{\sigma,p}$, and recall that the Fourier transform of f_n is given by

$$\widehat{f_n}(\xi) = \left(\widehat{f}(\xi/n^{1/p})\right)^n.$$

Since f and f_p belong to $D^0_{\sigma,p}$, we have

$$\begin{cases} \dfrac{|\widehat{f}(\xi) - (1-\sigma|\xi|^p)|}{|\xi|^p} = \dfrac{|\psi_f(\xi)|}{|\xi|^p} \leq r(|\xi|), \\ \dfrac{|\widehat{f_p}(\xi) - (1-\sigma|\xi|^p)|}{|\xi|^p} = \dfrac{|\psi_p(\xi)|}{|\xi|^p} \leq r(|\xi|) \end{cases}$$

where r can be supposed non decreasing from 0. Therefore, we get

$$\begin{aligned} \frac{|\widehat{f_n}(\xi) - \widehat{f_p}(\xi)|}{|\xi|^p} &= \frac{1}{|\xi|^p}\left|\left(\widehat{f}(\xi/n^{1/p})\right)^n - \left(\exp(-\sigma|\xi/n^{1/p}|^p)\right)^n\right| \\ &\leq \frac{n}{|\xi|^p}\left|\widehat{f}(\xi/n^{1/p}) - \widehat{f_p}(\xi/n^{1/p})\right| = \frac{\left|\widehat{f}(\xi/n^{1/p}) - \widehat{f_p}(\xi/n^{1/p})\right|}{|\xi/n^{1/p}|^p}, \end{aligned}$$

by using the inequality $|a^n - b^n| \leq |a-b| \ \sum_{k=0}^{n-1} |a|^k |b|^{n-1-k}$, with $a = \widehat{f}(\xi/n^{1/p})$, $b = \widehat{f}_p(\xi/n^{1/p})$ having their modulus ≤ 1. Consequently, we get

$$\frac{|\widehat{f}_n(\xi) - \widehat{f}_p(\xi)|}{|\xi|^p} \leq 2r(|\xi|/n^{1/p}).$$

This shows that $\widehat{f}_n$ converges uniformly to $\widehat{f}_p$ on compact sets since

$$\sup_{|\xi| \leq R} \frac{|\widehat{f}_n(\xi) - \widehat{f}_p(\xi)|}{|\xi|^p} \leq 2r(R/n^{1/p}) \xrightarrow[n \to \infty]{} 0.$$

We also note that $\psi_n(\xi) = \widehat{f}_n(\xi) - (1 - \sigma|\xi|^p)$ satisfies the uniform estimate $|\psi_n(\xi)|/|\xi|^p \leq Cr(|\xi|)$. On the other hand, one clearly has for $|\xi| \geq R$,

$$\frac{|\widehat{f}_n(\xi) - \widehat{f}_p(\xi)|}{|\xi|^p} \leq 2/R^p.$$

One deduces readily that

$$\lim_{n \to \infty} d_p(f_n, f_p) = 0.$$

Remark 3 *If $f \in D^{\delta}_{\sigma,p}$ with $0 \leq \delta \leq p$, one is led similarly to $d_{p+\delta}(f_n, f) \leq C/n^{\delta/p}$. Consequently, one obtains the following rate for the distance d_p*

$$d_p(f_n, f) \leq C/n^{\delta/(p+\delta)}.$$

By Lemma 10, we remark that

$$\begin{aligned} J(S_{n+k}) &= J\left(\frac{X_1 + ... + X_n}{n^{1/p}} \left(\frac{n}{n+k}\right)^{1/p} + \frac{X_{n+1} + ... + X_{n+k}}{k^{1/p}} \left(\frac{k}{n+k}\right)^{1/p}\right) \\ &\leq \frac{n}{n+k} \ J(S_n) + \frac{k}{n+k} \ J(S_k), \end{aligned}$$

since the X_i's obey the same law. It is a classical fact that a sequence satisfying this property converges to some limit ℓ.

Precisely, let us consider the particular case where n, k are successive powers of 2. For X and Y two independent random variables with distribution functions P and Q respectively, we denote by $P \circ Q$ the law of

$$Z = \left(\frac{1}{2}\right)^{1/p} X + \left(\frac{1}{2}\right)^{1/p} Y$$

Then, we get

$$J(S_{2^{n+1}}) = J(P_n \circ P_n) \leq J(S_{2^n}) = J(P_n)$$

where P_n is the probability law of S_{2^n}. If $d_p(P_k, Q)$ tends to 0, since the rescaled convolution is continuous with respect to d_p, we obtain $J(P_k \circ P_k) = J(P_{k+1}) \longrightarrow J(Q \circ Q) = J(Q) = \ell$, and we recover $Q = F_p$, $\ell = 0$.

The interest of the present method is that it can immediately lead to explicit computations. First, note that any integer $n \geq 1$ can be written $\sum_0^\infty \alpha_k 2^k$ with $\alpha_k \in \{0, 1\}$, and

$$J(S_n) \leq \frac{1}{n} \sum \alpha_k 2^k J(S_{2^k}) \longrightarrow 0.$$

Then, let $\varepsilon > 0$, and let us look for n_0 such that for $n \geq n_0$, $d_p(F_n, F_p) \leq \varepsilon$. To start with, since $J(P_k)$ is decreasing, we look for k_0 such that $J(P_{k_0}) \leq \varepsilon$. The proof of Lemma 10 clearly shows that

$$J(P_{k+1}) = J(P_k \circ P_k) \leq \sup_\xi \left\{ \inf \left(\psi(|\xi|), \left(\frac{1 + e^{-\sigma|\xi|^p}}{2} \right) J(P_k) \right) \right\}.$$

Let η such that $\psi(\eta) \leq \varepsilon$. As long as $J(P_k \circ P_k) \geq \varepsilon$, the supremum can only be obtained for $|\xi| \geq \eta$, which implies $J(P_{k+1}) \leq \mu J(P_k)$ with

$$\mu = \frac{1 + e^{-\sigma|\xi|^p}}{2} < 1.$$

Therefore, one can take

$$k_0 = -\frac{\log J(X_1)}{\log \mu}.$$

Now, for $n \geq 2^{k_0}$, one writes

$$J(X_n) \leq \sum_{k<k_0} \frac{\alpha_k 2^k}{n} J(X_1) + \varepsilon \sum_{k \geq k_0} \frac{\alpha_k 2^k}{n} \leq \frac{2^{k_0}}{n} J(X_1) + \varepsilon,$$

and it suffices to choose $n \geq 2^{k_0} J(X_1)/\varepsilon$ for this expression to be less than 2ε.

$\square$

References

[1] Carlen E.A., Gabetta E. and Toscani G., Propagation of smoothness and the rate of exponential convergence to equilibrium for a spatially homogeneous Maxwellian gas, *Commun. Math. Phys.*, **305**, (1999) 521-546.

[2] Carrillo J.A. and Toscani G., Exponential convergence to equilibrium for homogeneous Fokker-Planck type equations, *Math. Meth. Appl. Sci.*, **21**, (1998) 1269-1286.

[3] Feller W., *An introduction to probability theory and its applications; Vol. I–II*, John Wiley and Sons, New York, 1966.

[4] Gabetta E., Toscani G. and Wennberg B., Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation, *J. Stat. Phys.*, **81**, (1995) 901-934.

[5] Gradshteyn I.S. and Ryzhik I.M., *Table of integrals, series and products*, Fifth edition, Academic Press, New York, 1994.

[6] Hoffmann-Jørgensen J., *Probability with a view towards statistics; Vol.I*, Chapman & Hall, New York, 1994.

[7] Laha R.G. and Rohatgi V.K. *Probability Theory*, John Wiley and Sons, New York, 1979.

[8] Linnik Y.V. An information-theoretic proof of the central limit theorem with the Lindeberg condition, *Theory Prob. Appl.* **4** (1959) 288-299.

[9] Lions P.-L. and Toscani G., A strengthened central limit theorem for smooth densities, *J. Funct. Anal.*, **129**, (1995) 148-167.

[10] Murata H. and Tanaka H., An inequality for certain functional of multidimensional probability distributions, *Hiroshima Math. J.*, **4**, (1974) 75–81.

[11] Petrov V., *Sums of independent random variables*, Ergebnisse der Math. und ihrer Grenzgebiete, vol. 82 Springer Verlag, Berlin, 1975.

[12] Rachev S. and Ruschendorf L., *Mass transportation problems, Vol. II: Probability and its applications*, Springer Verlag, New York, 1998.

[13] Senatov V.V. Uniform estimates of the rate of convergence in the multidimensional central limit theorem *Theory Prob. Appl.* **25**, (1980) 745–759.

[14] Toscani G. and Villani C., Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas, *J. Stat. Phys.*, **94**, (1999) 619-637.

[15] Zolotarev V.M., Probability metrics, *Theory Prob. Appl.*, **28**, 1983, 278-302.