STABILITY PROBLEMS IN CRAMÉR-TYPE CHARACTERIZATION IN CASE OF I.I.D. SUMMANDS[*]

S. G. BOBKOV[†], G. P. CHISTYAKOV[‡], AND F. GÖTZE[‡]

Abstract. The stability property in Cramér's characterization of the normal law is considered in the case of identically distributed summands. As opposite results, instability is shown with respect to strong distances including the entropic distance to normality (addressing a question of M. Kac).

Key words. Cramér's theorem, Cramér's characterization of the normal law, stability problems

1. Introduction. A well-known theorem due to Cramér [7] states that if the sum, say $X + Y$, of two independent random variables X and Y has a normal distribution, then both summands X and Y have to be normal. Lévy showed that this characterization property is stable in the topology of weak convergence of probability distributions on the real line. Using Lévy's metric, $L(F, G)$, to measure distances in weak convergence between any distribution functions F, G, there exists for any $\varepsilon > 0$ a number δ_ε, depending on ε only, such that $\delta_\varepsilon \to 0$ as $\varepsilon \to 0$ and

$$L(F * G, \Phi) < \varepsilon \quad \text{implies} \quad L(F, \Phi_{a_1,\sigma_1}) < \delta_\varepsilon \quad \text{and} \quad L(G, \Phi_{a_2,\sigma_2}) < \delta_\varepsilon$$

for some $a_1, a_2 \in \mathbf{R}$ and $\sigma_1, \sigma_2 > 0$.

Here $\Phi_{a,\sigma}$ denotes the distribution function of the normal law $N(a, \sigma^2)$ with mean a and variance σ^2. In the standard case, where $a = 0$ and $\sigma = 1$, we shall omit these indices. Furthermore, let $F * G$ denote the convolution of the corresponding distributions.

There are similar results in terms of the Kolmogorov distance, including sharp quantitative estimates for the rate function δ_ε. Here we refer the reader to [2], [3], [4], [5], [12], [13], [14], [17], [18], [19], [21], [22], [23], [24], [25]. Other results in this context study characterizations and stability problems for distributions of more complicated functionals of independent summands, including linear and quadratic forms (cf., e.g., [11], [6]).

In the mid 1960s Kac, and then McKean, raised the question of the stability in Cramér's theorem with respect to the entropic distance to normality (see [15, pp. 365, 366]). Let us recall that, if a random variable X with finite second moment has density $p(x)$, then the entropic distance of the distribution of X to a normal distribution, or to Gaussianity, is defined by

$$D(X) = h(Z) - h(X) = \int_{-\infty}^{+\infty} p(x) \log \frac{p(x)}{\varphi_{a,\sigma}(x)} \, dx,$$

where $\varphi_{a,\sigma}$ stands for the density of a normal random variable Z with parameters

[†]School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (bobkov@math.umn.edu).

[‡]Faculty of Mathematics, University of Bielefeld, 33615 Bielefeld, Germany (chistyak@math.uni-bielefeld.de, goetze@math.uni-bielefeld.de).

$a = \mathbf{E}\,X$, $\sigma^2 = \mathbf{D}X$, and where

$$h(X) = -\int_{-\infty}^{+\infty} p(x)\log p(x)\,dx$$

denotes the classical entropy functional. The quantity $D(X)$ represents the shortest distance from the distribution F of X to the family of all normal laws on the line with respect to the Kullback–Leibler distance.

In general, $0 \leqq D(X) \leqq +\infty$. As in the case of the Kolmogorov (or uniform) distance $\|F - \Phi_{a,\sigma}\| = \sup_x |F(x) - \Phi_{a,\sigma}(x)|$ and in the case of the total variation distance $\|F - \Phi_{a,\sigma}\|_{\mathrm{TV}}$, the value of $D(X)$ depends on neither the mean nor the variance of X. This distance is stronger than the total variation in view of the Pinsker (or Pinsker–Csiszár–Kullback) inequality, which yields

$$D(X) \geqq \frac{1}{2}\,\|F - \Phi_{a,\sigma}\|_{\mathrm{TV}}^2.$$

Thus, Kac's question may be rephrased as to whether or not smallness of $D(X+Y)$ for independent summands implies the smallness of $D(X)$ and $D(Y)$. A negative answer to this question with nonidentically distributed summands has recently been given in [1]. Surprisingly, the random variables X and Y in another counterexample to the Kac conjecture can be chosen to be identically distributed. More precisely, we prove the following theorem.

THEOREM 1.1. *For any $\varepsilon > 0$, there exist independent random variables X and Y with a common absolutely continuous symmetric distribution F such that*

(a) $D(X + Y) < \varepsilon$;

(b) $\|F - \Phi_{a,\sigma}\|_{\mathrm{TV}} > c$ *for all $a \in \mathbf{R}$ and $\sigma > 0$, where $c > 0$ is an absolute constant*; *in particular, $D(X)$ is separated from zero.*

As shown below, the distribution F in Theorem 1.1 may be chosen to be more or less well behaved. For example, in addition to properties (a) and (b) (with constants c depending on α below) it may have an almost standard normal density p in the sense that, for a prescribed $\alpha \in (0,1)$,

$$(1 - \alpha)\,\varphi(x) \leqq p(x) \leqq (1 + \alpha)\varphi(x) \qquad \forall\, x \in \mathbf{R}.$$

In other words, a rather small oscillation of p around φ suffices to separate F and Φ, while at the same time moving $F * F$ arbitrarily close to $\Phi * \Phi$ in a strong metric, like the total variation and even entropic distance.

Let us discuss this result in view of previous variants of stability in the Cramér theorem for the Kolmogorov distance, which is essentially due to Sapogov (cf. [17], [18], [2]). If independent random variables X and Y have mean zero and (equal) variances $\mathbf{D}X = \mathbf{D}Y = 1$, but not necessarily equal distributions, say, F and G, then the condition

$$\|F * G - \Phi * \Phi\| \leqq \varepsilon < 1$$

ensures that

$$\|F - \Phi\| \leqq \frac{C}{\sqrt{\log(1/\varepsilon)}} \quad \text{and} \quad \|G - \Phi\| \leqq \frac{C}{\sqrt{\log(1/\varepsilon)}}$$

with some absolute constant C. This bound is asymptotically optimal for small values of ε (cf. [13], [3]). However, in case of identically distributed summands, it is rather far

from being optimal, since the bound can be improved to a polynomial rate. Il'in [10] observed that, if F has a median at zero (which is actually not important), the condition $\|F * F - \Phi * \Phi\| \leqq \varepsilon$ with ε small guarantees that

$$\|F - \Phi\| \leqq C \, \varepsilon^{1/36} \log \frac{1}{\varepsilon}.$$

We shall first sharpen the exponent in this estimate. In contrast to Sapogov's result, this can be done without using heavy tools from complex analysis, but rather using classical probabilistic arguments that involve Berry–Esseen's bound and Plancherel's formula.

THEOREM 1.2. *If a distribution function F satisfies $\|F * F - \Phi * \Phi\| \leqq \varepsilon < 1/e$, then we obtain with some absolute constant C*

$$\|F - \Phi\| \leqq C \left(\varepsilon \log \frac{1}{\varepsilon} \right)^{2/5}.$$

Perhaps this bound is still not asymptotically attainable, and the exponent $2/5$ might be replaced with $1/2$ modulo a logarithmic factor. In order to see that the exponent $1/2$ is unimprovable, we construct another family of probability distributions and prove the following theorem.

THEOREM 1.3. *For any $\varepsilon \in (0, 1/e)$, there exists a probability distribution F on the real line, such that $\|F * F - \Phi * \Phi\| \leqq \varepsilon$, while*

$$\|F - \Phi\| \geqq c \, \varepsilon^{1/2} \left(\log \frac{1}{\varepsilon} \right)^{-1/4},$$

where $c > 0$ is an absolute constant.

In fact, even in the case of a stronger hypothesis, such as $\|F * F - \Phi * \Phi\|_{\mathrm{TV}} \leqq \varepsilon$, one can obtain a similar inequality $\|F - \Phi\| \geqq c \varepsilon^{1/2} \log^{-1/2}(1/\varepsilon)$.

Finally, let us emphasize that the stability property itself for identically distributed independent random variables (similarly to Theorem 1.2) remains valid for a large class of probability distributions on the line, viewed as "points of attraction" (although, in general, the Lévy distance should replace the Kolmogorov distance for an improved control of weak convergence). In other words, there is nothing special about the normal law in Theorem 1.2 beyond the problem of optimal rates. This can be seen from the following more general result.

THEOREM 1.4. *Assume a distribution function G has a real positive characteristic function. There exists a positive, increasing function $\alpha(\varepsilon) = \alpha_G(\varepsilon)$ such that $\alpha(\varepsilon) \to 0$ as $\varepsilon \to 0$ and the following property holds. For any distribution function F satisfying*

$$L(F * F, G * G) \leqq \varepsilon \leqq 1,$$

we have $L(F, G) \leqq \alpha(\varepsilon)$.

Here, the assumption on the positivity of the characteristic function of G cannot be removed. For example, it is well known (cf. [9, Chap. XV]) that there exist two different distributions F and G on the real line, such that $F * F = G * G$. Hence, G contradicts the conclusion of Theorem 1.4. One of these distributions, say G, has a nonnegative 2-periodic characteristic function g defined by $g(t) = 1 - |t|$ for $|t| \leqq 1$, while F has the characteristic function $f(t) = 2g(t/2) - 1$.

The organization of the paper is as follows. In section 2 we derive an elementary, but useful, bound on the deviations of characteristic functions in terms of the Lévy

metric. In section 3 we consider some properties of squares of characteristic functions. Theorem 1.4 is proved in section 4. Some additional technical bounds needed for the Gaussian case are collected in section 5. The proof of Theorem 1.2 is completed in section 6. Sections 7–9 are devoted to the proof of Theorem 1.1. Finally, Theorem 1.3 is proved in section 10.

2. A bound on deviations of characteristic functions. Throughout the paper we deal with the following classical metrics in the space of all probability distributions on the real line:

(1) The Kolmogorov distance $\|F - G\| = \sup_x |F(x) - G(x)|$;

(2) the Lévy distance

$$L(F, G) = \min \big\{ h \geqq 0 \colon\ G(x - h) - h \leqq F(x) \leqq G(x + h) + h\ \forall x \in \mathbf{R} \big\};$$

(3) the total variation distance

$$\|F - G\|_{\mathrm{TV}} = \sup \sum |(F(x_k) - G(x_k)) - (F(y_k) - G(y_k))|,$$

where the supremum is taken over all finite collections of points $y_1 < x_1 < \cdots < y_n < x_n$.

Here, F and G denote arbitrary distribution functions.

We introduce the characteristic functions

$$f(t) = \int_{-\infty}^{+\infty} e^{itx}\, dF(x), \quad g(t) = \int_{-\infty}^{+\infty} e^{itx}\, dG(x) \qquad (t \in \mathbf{R}).$$

LEMMA 2.1. *For all $t \in \mathbf{R}$ and $N > 0$,*

$$(2.1) \qquad |f(t) - g(t)| \leqq 2L(F, G)(2 + (N + 2)|t|) + 3G\{|x| > N\}.$$

Here $G\{|x| > N\} = \int_{\{|x| > N\}} dG(x)$, and we use similar notation below.

Proof. The argument is rather standard, and we include it for completeness.

Put $h = L(F, G)$. Without loss of generality, one may assume that N and $N \pm h$ are points of continuity of both F and G. Given $t \in \mathbf{R}$, write

$$(2.2) \quad f(t) - g(t) = \int_{-N-h}^{N+h} e^{itx}\, d\left(F(x) - G(x)\right) + \int_{\{|x| > N+h\}} e^{itx}\, d\left(F(x) - G(x)\right).$$

The last integral is bounded from above in absolute value by

$$F\{|x| > N + h\} + G\{|x| > N + h\}.$$

Also, by the definition of the Lévy distance,

$$F\{|x| > N + h\} = F(-N - h) + (1 - F(N + h))$$
$$\leqq (G(-N) + h) + (1 - G(N)) + h = G\{|x| \geqq N\} + 2h.$$

Hence, the last integral in (2.2) is bounded by

$$(2.3) \qquad\qquad 2\, G\{|x| \geqq N\} + 2h.$$

Further, integrating by parts, we have

$$(2.4) \qquad \int_{-N-h}^{N+h} e^{itx} \, d\left(F(x) - G(x)\right) = e^{itx} \left(F(x) - G(x)\right)\Big|_{x=-N-h}^{x=N+h}$$

$$(2.5) \qquad\qquad\qquad\qquad\qquad - it \int_{-N-h}^{N+h} e^{itx} \left(F(x) - G(x)\right) dx.$$

Since

$$F(-N-h) - G(-N-h) \leqq F(-N-h) \leqq G(-N) + h$$

and $G(-N-h) - F(-N-h) \leqq G(-N)$, we have that

$$|F(-N-h) - G(-N-h)| \leqq G(-N) + h.$$

Similarly, $|F(N+h) - G(N+h)| \leqq (1 - G(N)) + h$, so the increment term in (2.4) is bounded in absolute value by

$$(2.6) \qquad\qquad\qquad\qquad G\{|x| \geqq N\} + 2h.$$

Now, let us turn to the integral in (2.5) over the interval $(-N-h, N+h)$. Again, by the definition of the Lévy distance,

$$F(x) - G(x) = (F(x) - G(x+h)) + (G(x+h) - G(x)) \leqq h + (G(x+h) - G(x)),$$

and similarly,

$$G(x) - F(x) \leqq h + (F(x+h) - F(x)).$$

Hence,

$$|F(x) - G(x)| \leqq h + (G(x+h) - G(x)) + (F(x+h) - F(x)),$$

which implies that

$$\int_{-N-h}^{N+h} |F(x) - G(x)| \, dx \leqq 2(N+h)h$$
$$+ \int_{-\infty}^{+\infty} (G(x+h) - G(x)) \, dx + \int_{-\infty}^{+\infty} (F(x+h) - F(x)) \, dx.$$

Each of the last two integrals does not exceed h, so

$$\left| \int_{-N-h}^{N+h} e^{itx} \left(F(x) - G(x)\right) dx \right| \leqq 2(N+1+h)h \leqq 2(N+2)h.$$

Using the bounds (2.3) and (2.6), we obtain that

$$|f(t) - g(t)| \leqq 3G\{|x| \geqq N\} + 4h + 2|t| \, (N+2)h,$$

which is exactly (2.1).

Remark 2.1. With the same argument one may derive a similar inequality in terms of the classical β-metric defined in the space of all probability distributions via

$$\beta(F, G) = \sup_{\|u\|_{\mathrm{BL}} \leqq 1} \left| \int u(x) \, dF(x) - \int u(x) \, dG(x) \right|,$$

where $\|u\|_{\mathrm{BL}} = \sup_x |u(x)| + \|u\|_{\mathrm{Lip}}$; cf. [8, Chap. 11]. Replacing the function e^{itx} in the proof of Lemma 2.1 with $u(x)$ and maximizing over all admissible u, one arrives at

$$\beta(F, G) \leqq 2\,L(F, G)\,(N + 4) + 3\,G\{|x| > N\}.$$

There is also an elementary reverse bound as well, $L(F, G) \leqq 2\sqrt{\beta(F, G)}$.

Note that there is a slight refinement in Lemma 2.1 when switching to the Kolmogorov distance.

LEMMA 2.2. *For all $t \in \mathbf{R}$ and $N > 0$,*

$$(2.7) \qquad |f(t) - g(t)| \leqq 2\|F - G\|(2 + N|t|) + 2G\{|x| > N\}.$$

Indeed, write

$$f(t) - g(t) = \int_{-N}^{N} e^{itx}\,d(F(x) - G(x)) + \int_{\{|x|>N\}} e^{itx}\,d(F(x) - G(x)).$$

The last integral is bounded in absolute value from above by

$$F\{|x| > N\} + G\{|x| > N\} \leqq 2G\{|x| > N\} + 2\|F - G\|.$$

In addition, integrating by parts, we have

$$\int_{-N}^{N} e^{itx}\,d(F(x) - G(x)) = e^{itx}\,(F(x) - G(x))\Big|_{x=-N}^{x=N}$$
$$- it \int_{-N}^{N} e^{itx}(F(x) - G(x))\,dx,$$

which is bounded in absolute value by $2\|F - G\|(1 + N|t|)$. The two bounds lead to (2.7).

3. Squares of characteristic functions. Let F and G be two distribution functions with characteristic functions f and g, respectively. We want to estimate the difference $f(t) - g(t)$ on large intervals in terms of the Lévy distance $L(F * F, G * G)$ between these convolutions. When g is real and positive, we shall use the following special quantity, which we denote by T_g.

Notation. Define

$$T_g(\varepsilon) = \min\{t \geqq 0\colon\ g(t) \leqq \varepsilon\}, \qquad 0 \leqq \varepsilon \leqq 1,$$

where we put $T_g(\varepsilon) = +\infty$ in case $g(t) > \varepsilon$ for all $t > 0$.

By Lemma 2.1, if $L(F * F, G * G) \leqq \varepsilon \leqq 1$, then

$$(3.1) \qquad |f(t)^2 - g(t)^2| \leqq 2\varepsilon\,(2 + (N + 2)|t|) + 3\,G * G\{|x| > N\}.$$

To simplify the right-hand side, let X, X' be independent random variables with distribution function G. Then

$$G * G\{|x| > N\} = \mathbf{P}\{|X + X|' > N\}$$
$$\leqq \mathbf{P}\left\{|X| > \frac{N}{2}\right\} + \mathbf{P}\left\{|X'| > \frac{N}{2}\right\} = 2G\left\{X > \frac{N}{2}\right\}.$$

Replacing N with $2N$, we thus get from (3.1) that

$$|f(t)^2 - g(t)^2| \leqq \varepsilon' \equiv 4\varepsilon \left(1 + (N+1)|t|\right) + 6\,G\{|x| > N\}.$$

Now, assume that $g(t)$ is real and positive.

Since $\mathrm{Re}[f(t)^2] = (\mathrm{Re}\, f(t))^2 - (\mathrm{Im}\, f(t))^2$, we then get that

$$\left| (\mathrm{Re}\, f(t))^2 - (\mathrm{Im}\, f(t))^2 - g(t)^2 \right| \leqq \varepsilon'.$$

This ensures that $\mathrm{Re}\, f(t) \geqq 0$ on a certain large interval $[-a, a]$. Indeed, otherwise for some $t_0 \in [-a, a]$, we would have that $\mathrm{Re}\, f(t_0) = 0$, which implies that

$$(\mathrm{Im}\, f(t_0))^2 + g(t_0)^2 \leqq \varepsilon',$$

and, therefore, $g(t_0) \leqq \sqrt{\varepsilon'}$. But, if ε' is small enough, the inequality $g(t_0) \leqq \sqrt{\varepsilon'}$ is possible for sufficiently large t_0 only.

Thus, we conclude that

$$\mathrm{Re}\, f(t) < 0 \quad \Rightarrow \quad \exists\, t_0 \text{ such that } 0 < t_0 < |t| \text{ and } \mathrm{Re}\, f(t_0) = 0$$
$$\Rightarrow \quad g(t_0) \leqq \sqrt{\varepsilon'} \quad \Rightarrow \quad |t| > t_0 \geqq T_g(\sqrt{\varepsilon'}),$$

provided that $\varepsilon' \leqq 1$. Hence,

$$|f(t)^2 - g(t)^2| \leqq \varepsilon' \leqq 1 \quad \Rightarrow \quad \mathrm{Re}\, f(t) \geqq 0 \text{ on the interval } |t| \leqq T_g(\sqrt{\varepsilon'}).$$

On the other hand, the inequality $|f(t)^2 - g(t)^2| \leqq \varepsilon'$ requires that

$$|f(t) - g(t)| \leqq \sqrt{\varepsilon'} \quad \text{or} \quad |f(t) + g(t)| \leqq \sqrt{\varepsilon'}.$$

But for $|t| \leqq T_g(\sqrt{\varepsilon'})$ we have $\mathrm{Re}\, f(t) \geqq 0$, so the second assertion would imply that $g(t) \leqq \sqrt{\varepsilon'}$, which is only possible for $|t| \geqq T_g(\sqrt{\varepsilon'})$. Hence,

$$|f(t)^2 - g(t)^2| \leqq \varepsilon' \leqq 1 \quad \Rightarrow \quad |f(t) - g(t)| \leqq \sqrt{\varepsilon'}$$
$$\text{on the interval} \quad |t| \leqq T_g(\sqrt{\varepsilon'}).$$

We summarize the above arguments in the following lemma.

LEMMA 3.1. *Assume that G has a real positive characteristic function g. Suppose that $L(F * F, G * G) \leqq \varepsilon$, and for $N > 0$ and $t \in \mathbf{R}$, let*

$$(3.2) \qquad \varepsilon' = 4\varepsilon(1 + (N+1)|t|) + 6G\{|x| > N\} \leqq 1.$$

If $|t| \leqq T_g(\sqrt{\varepsilon'})$, then

$$(3.3) \qquad |f(t) - g(t)| \leqq \sqrt{\varepsilon'}.$$

Remark 3.1. Using similar arguments, and taking into account Lemma 2.2 and inequality (2.7), the stronger assumption $\|F * F - G * G\| \leqq \varepsilon$ implies (3.3) for a smaller number

$$(3.4) \qquad \varepsilon' = 4\varepsilon \left(1 + N|t|\right) + 4\,G\{|x| > N\},$$

provided that $\varepsilon' \leqq 1$ and $|t| \leqq T_g(\sqrt{\varepsilon'})$.

4. Stability theorem. We are now ready to establish the stability property of convolutions for equal factors with respect to the Lévy distance, as described in Theorem 1.4.

It is enough to show that, for any $\delta > 0$, one can find $\varepsilon > 0$, depending on ε and G only, such that $L(F, G) \leqq \delta$, as long as $L(F * F, G * G) \leqq \varepsilon$.

We may additionally assume that the characteristic function g of G vanishes at infinity (otherwise, a slight modification of the argument is needed). Under this assumption, the function $\varepsilon \mapsto T_g(\varepsilon)$ is finite for $0 < \varepsilon < 1$. In addition, it is nonincreasing and satisfies $T_g(\varepsilon) \to +\infty$ as $\varepsilon \to 0$.

We shall use the so-called λ-metric, introduced by Zolotarev and Senatov [20]:

$$\lambda(F, G) = \min_{T > 0} \max_{|t| \leqq T} \left\{ \frac{1}{2} |f(t) - g(t)|, \ \frac{1}{T} \right\}.$$

Thus,

$$\forall |t| \leqq T |f(t) - g(t)| \leqq a \quad \Rightarrow \quad \lambda(F, G) \leqq \max\left\{ \frac{a}{2}, \frac{1}{T} \right\}.$$

Assume $L(F * F, G * G) \leqq \varepsilon$ and apply Lemma 3.1 with, say, $N = \varepsilon^{-1/3}$. If $|t| \leqq \varepsilon^{-1/3}$, then we may bound ε' in (3.2) by

$$\varepsilon' \leqq \varepsilon'_G(\varepsilon) \to 0 \quad \text{as } \varepsilon \to 0,$$

with some positive function ε'_G, depending on G only. Put

$$T(\varepsilon) = \min\left\{ \varepsilon^{-1/3}, \ T_g\left(\sqrt{\varepsilon'_G(\varepsilon)} \right) \right\}.$$

Then the conclusion (3.3) in Lemma 3.1 takes the form

$$\max_{|t| \leqq T(\varepsilon)} |f(t) - g(t)| \leqq \sqrt{\varepsilon'_G(\varepsilon)},$$

so

$$\lambda(F, G) \leqq \lambda_G(\varepsilon) \equiv \max\left\{ \frac{1}{2} \sqrt{\varepsilon'_G(\varepsilon)}, \ \frac{1}{T(\varepsilon)} \right\} \to 0 \quad \text{as } \varepsilon \to 0.$$

But, according to Theorem 2 in [20], the Lévy distance $L = L(F, G)$ is estimated in terms of $\lambda = \lambda(F, G)$ by virtue of the inequality

$$L \leqq 8\lambda \log \frac{Y}{\lambda},$$

where $Y \geqq e$ may be any number, satisfying $G\{|x| > Y/e\} \leqq \lambda \log(Y/\lambda)$. This ensures that $L \leqq \psi_G(\lambda)$ for some nonnegative increasing function ψ_G such that $\psi_G(\lambda) \to 0$ as $\lambda \to 0$. Therefore,

$$L(F, G) \leqq \psi_G(\lambda_G(\varepsilon)) \to 0 \quad \text{as } \varepsilon \to 0.$$

Thus, Theorem 1.4 is proved.

5. The Gaussian case. In order to find an explicit rate function $\alpha(\varepsilon)$ in the Gaussian case, one could use similar arguments as in the proof of the more general Theorem 1.4. However, this would lead us to inappropriate rates. So, we have to return to Lemma 3.1 (or Remark 3.1 concerning the Kolmogorov distance) and apply it in a more sophisticated way.

Thus, let $G = \Phi$ denote the standard normal distribution, and assume that, for a given distribution function F, we have

$$\text{(5.1)} \qquad \|F * F - \Phi * \Phi\| \leqq \varepsilon$$

with $\varepsilon \in (0, 1/e]$. Note that, since Φ has a bounded density, it does not matter whether the Lévy or Kolmogorov distance is used to measure distance.

We need to find appropriate bounds for the difference of the characteristic functions, $f(t) - g(t)$, where now $g(t) = e^{-t^2/2}$, on different intervals of t-axis.

First, in order to control the relatively small values of t, we apply Remark 3.1 with the value $N = \sqrt{2\log(1/\varepsilon)}$. In particular, $\Phi\{|x| > N\} \leqq e^{-N^2/2} = \varepsilon$, so the quantity ε' in (3.4) for $\varepsilon \leqq 1/e$ is bounded by

$$4\varepsilon\left(1 + \sqrt{2\log\frac{1}{\varepsilon}}\,|t|\right) + 4\varepsilon = 8\varepsilon + 4\varepsilon\sqrt{2\log\frac{1}{\varepsilon}}\,|t|.$$

For $|t| \leqq N$, the last expression may be bounded from above by

$$8\,\varepsilon + 8\,\varepsilon\log\frac{1}{\varepsilon} \leqq 16\,\varepsilon\log\frac{1}{\varepsilon}.$$

Thus, in case $\varepsilon \leqq 1/e$, the concluding inequality (3.3), $|f(t) - g(t)| \leqq \sqrt{\varepsilon'}$, holds with

$$\varepsilon' = 16\,\varepsilon\log\frac{1}{\varepsilon},$$

provided that $\varepsilon' \leqq 1$ and $|t| \leqq \min\{N, T_g(\sqrt{\varepsilon'})\}$. But

$$N \geqq T_g(\sqrt{\varepsilon'}) = \sqrt{\log\frac{1}{\varepsilon'}} = \sqrt{\log\frac{1}{\varepsilon} - \log\log(\varepsilon^{-16})},$$

since the inequality on the left is equivalent to $e^{-N^2/2} \leqq \sqrt{\varepsilon'}$, that is, $16\,\varepsilon\log(1/\varepsilon) \geqq \varepsilon^2$, which is true for $\varepsilon \leqq 1/e$. By Lemma 3.1 and Remark 3.1, we obtain the following result.

LEMMA 5.1. *Assuming* (5.1), *with* $\varepsilon' = 16\,\varepsilon\log(1/\varepsilon) \leqq 1$ *and* $|t| \leqq \sqrt{\log(1/\varepsilon')}$, *we have*

$$|f(t) - e^{-t^2/2}| \leqq 4\sqrt{\varepsilon\log\frac{1}{\varepsilon}}.$$

Furthermore, in order to control the deviations of $f(t)$ around $g(t)$ for very small values of t, or for relatively large t, a truncation procedure will be needed.

DEFINITION. *Given a random variable X with distribution function F, put $X^* = X$ in case $|X| \leqq M$ and $X^* = 0$ in case $|X| > M$, where*

$$M = M(\varepsilon) = 1 + 2\sqrt{\log\frac{1}{\varepsilon}}.$$

Denote by F^* the distribution function and by f^* the characteristic function of X^*.

Since F^* is supported on the interval $[-M, M]$, its characteristic function satisfies, for all t,

$$|1 - f^*(t)| \leqq \int_{-M}^{M} |1 - e^{itx}| \, dF^*(x) \leqq |t| \int_{-M}^{M} |x| \, dF^*(x) \leqq M|t|.$$

Using a similar inequality for the normal distribution, we have that

$$(5.2) \qquad\qquad |f^*(t) - e^{-t^2/2}| \leqq 2M(\varepsilon) \, |t|.$$

Before using this bound, one needs to quantify the distance between F^* and F in the Kolmogorov metric. More generally, the next lemma allows one to reduce Theorem 1.2 to the case of compactly supported distributions.

LEMMA 5.2. *Under* (5.1),

$$(5.3) \qquad\qquad \|F^* - F\| \leqq 13 \, \varepsilon \quad and \quad \|F^* * F^* - \Phi * \Phi\| \leqq 13 \, \varepsilon.$$

Proof. The argument is rather standard. With a slight modification we borrow it from [17] and [19], where the non-i.i.d. case is treated.

Assume that ε is small enough, say, $\varepsilon \leqq \varepsilon_0 = 1/4 - \Phi(-1) = 0.091345\ldots$. This assumption allows one to bound a median $m = m(X)$. Indeed, taking an independent copy Y of X and applying (5.1), we get

$$\frac{1}{4} \leqq \mathbf{P}\{X \leqq m, \, Y \leqq m\} \leqq \mathbf{P}\{X + Y \leqq 2m\} \leqq (\Phi * \Phi)(2m) + \varepsilon = \Phi(\sqrt{2}\,m) + \varepsilon.$$

Hence, $\Phi(\sqrt{2}\,m) \geqq 1/4 - \varepsilon_0 = \Phi(-1)$. So, $m \geqq -1/\sqrt{2}$. With a similar bound for $-m$, we obtain that

$$(5.4) \qquad\qquad |m(X)| < 1.$$

To simplify matters, we may assume that M is a point of continuity of F. By definition,

$$(\Phi * \Phi)(-(M-1)) = 1 - (\Phi * \Phi)(M - 1) \leqq \frac{1}{2}\, e^{-(M-1)^2/4} = \frac{1}{2}\, \varepsilon.$$

Hence, by (5.1) once more,

$$\mathbf{P}\{X \leqq 1, \, Y \leqq -M\} \leqq \mathbf{P}\{X + Y \leqq -(M-1)\}$$
$$\leqq (\Phi * \Phi)(-(M-1)) + \varepsilon \leqq \frac{3}{2}\, \varepsilon,$$

and, by (5.4), $F(-M) \leqq 3\varepsilon$. Analogously, $1 - F(M) \leqq 3\varepsilon$. Thus, $F\{|x| \geqq M\} \leqq 6\varepsilon$, where F is treated as a measure.

Now, for $x < -M$, we have $|F^*(x) - F(x)| = F(x) \leqq 6\varepsilon$, and similarly for $x > M$.

If $-M < x < 0$, then $F^*(x) = F(x) - F(-M)$, and if $0 < x < M$, then we have $F^*(x) = F(x) + (1 - F(M))$. In both cases, $|F^*(x) - F(x)| \leqq 6\varepsilon$. Therefore,

$$\|F^* - F\| \leqq 6\varepsilon.$$

From this, by the triangle inequality,

$$\|F^* * F^* - F * F\| \leqq \|F^* * F^* - F^* * F\| + \|F^* * F - F * F\|$$
$$\leqq \|F^* - F\| + \|F^* - F\| \leqq 12\,\varepsilon.$$

Finally, by (5.1) and the triangle inequality again, we get

$$\|F^* * F^* - \Phi * \Phi\| \leqq \|F^* * F^* - F * F\| + \|F * F - \Phi * \Phi\| \leqq 13\,\varepsilon.$$

Thus, the bounds (5.3) hold for the values $\varepsilon \leqq \varepsilon_0$. For the range $\varepsilon > \varepsilon_0$, they hold automatically, since $13\,\varepsilon_0 > 1$. Thus, Lemma 5.2 is proved.

LEMMA 5.3. *Let* $\varepsilon' = 16\,\varepsilon \log(1/\varepsilon) \leqq 1/e$ *and* $T_1 = \sqrt{\log(1/\varepsilon')} < T$. *If* F *is supported on* $[-M(\varepsilon), M(\varepsilon)]$, *then, assuming* (5.1), *we have*

$$\int_{T_1}^{T} \frac{|f(t) - e^{-t^2/2}|}{t}\, dt \leqq CT^{1/4}\sqrt{\varepsilon \log \frac{1}{\varepsilon}},$$

where C *is an absolute constant.*

Proof. Recall that $g(t) = e^{-t^2/2}$. Since

$$\frac{f(t)^2 - g(t)^2}{-it} = \int_{-\infty}^{+\infty} e^{itx}\Big((F * F)(x) - (\Phi * \Phi)(x)\Big)\, dx,$$

one may apply Plancherel's formula

$$(5.5) \qquad \frac{1}{\pi}\int_{0}^{+\infty} \frac{|f(t)^2 - g(t)^2|^2}{t^2}\, dt = \int_{-\infty}^{+\infty} \Big|(F * F)(x) - (\Phi * \Phi)(x)\Big|^2 dx.$$

By the assumption (5.1), the integrand on the right-hand side is bounded by ε^2, so

$$\int_{-2M}^{2M} \Big|(F * F)(x) - (\Phi * \Phi)(x)\Big|^2 dx \leqq 4M\varepsilon^2.$$

On the other hand, since F is supported on $[-M, M]$, for the region $x > 2M$ we have $(F * F)(x) = 1$, while $(F * F)(x) = 0$ in $x < -2M$. Hence, in terms of a standard normal random variable Z,

$$\int_{2M}^{+\infty} \big|(F * F)(x) - (\Phi * \Phi)(x)\big|^2 dx = \int_{2M}^{+\infty} \big|1 - (\Phi * \Phi)(x)\big|^2 dx$$
$$= \int_{2M}^{+\infty} \Big[\mathbf{P}\Big\{Z > \frac{x}{\sqrt{2}}\Big\}\Big]^2 dx \leqq \frac{1}{4}\int_{2M}^{+\infty} e^{-x^2/2}\, dx$$
$$= \frac{\sqrt{2\pi}}{4}\,\mathbf{P}\{Z > 2M\} < \frac{1}{2}\,e^{-2M^2} < \frac{1}{2}\,\varepsilon^2.$$

With a similar bound for the region $x < -2M$, we can estimate the right integral in (5.5) by $(1 + 4M)\varepsilon^2 < 5M\varepsilon^2$, which results in

$$(5.6) \qquad \int_{0}^{+\infty} \frac{|f(t)^2 - g(t)^2|^2}{t^2}\, dt \leqq 5M\pi\varepsilon^2.$$

Now, just write

$$|f(t)|^4 = \left| (f(t)^2 - g(t)^2) + g(t)^2 \right|^2 \leqq 2|f(t)^2 - g(t)^2|^2 + 2g(t)^4,$$

so that (5.6) implies

$$\int_{T_1}^{+\infty} \frac{|f(t)|^4}{t^2}\, dt \leqq 10\, M\pi\varepsilon^2 + 2 \int_{T_1}^{+\infty} \frac{g(t)^4}{t^2}\, dt.$$

Since $T_1 \geqq 1$, the last integral does not exceed

$$\int_{T_1}^{+\infty} e^{-2t^2}\, dt = \frac{\sqrt{2\pi}}{2}\, \mathbf{P}\{Z > 2T_1\} \leqq \frac{\sqrt{2\pi}}{2}\, e^{-2T_1^2} = \frac{\sqrt{2\pi}}{2}\, (\varepsilon')^2 \leqq C\varepsilon^2 \left(\log \frac{1}{\varepsilon} \right)^2$$

with some absolute constant C. Hence we get (using a different constant), for any $T > T_1$,

$$(5.7) \qquad \int_{T_1}^{T} \frac{|f(t)|^4}{t^2}\, dt \leqq C\varepsilon^2 \left(\log \frac{1}{\varepsilon} \right)^2.$$

Applying Hölder's inequality with exponents $p = 4$, $q = 4/3$ with respect to the measure $d\mu(t) = dt/t^2$ on the interval $[T_1, T]$, and using (5.7), we get that

$$\begin{aligned}
\int_{T_1}^{T} \frac{|f(t)|}{t}\, dt &= \int_{T_1}^{T} |f(t)|\, t\, d\mu(t) \\
&\leqq \left(\int_{T_1}^{T} |f(t)|^4\, d\mu(t) \right)^{1/4} \left(\int_{T_1}^{T} t^{4/3}\, d\mu(t) \right)^{3/4} \\
&\leqq C\sqrt{\varepsilon \log \frac{1}{\varepsilon}}\, \left(T^{1/3} - T_1^{1/3} \right)^{3/4}.
\end{aligned}$$

Thus, with some absolute constant C,

$$\int_{T_1}^{T} \frac{|f(t)|}{t}\, dt \leqq C T^{1/4} \sqrt{\varepsilon \log \frac{1}{\varepsilon}}.$$

It remains to note that

$$\frac{1}{\sqrt{2\pi}} \int_{T_1}^{+\infty} \frac{g(t)}{t}\, dt \leqq \mathbf{P}\{Z > T_1\} \leqq e^{-T_1^2/2} = \sqrt{\varepsilon'} = 4\sqrt{\varepsilon \log \frac{1}{\varepsilon}}.$$

Thus, Lemma 5.3 is proved.

6. Proof of Theorem 1.2. According to Lemma 5.2, we may assume that the distribution F is supported on the interval $[-M, M]$, where

$$M = M(\varepsilon) = 1 + 2\sqrt{\log \frac{1}{\varepsilon}}$$

with $\varepsilon > 0$, such that $\varepsilon' = 16\, \varepsilon \log(1/\varepsilon) \leqq 1/e$.

In particular, the characteristic function f of F satisfies the bound (5.2).

At this point we shall employ the Berry–Esseen bound (cf., e.g., [16]): For any $T > 0$,

$$\text{(6.1)} \qquad \|F - \Phi\| \leqq c_1 \int_0^T \frac{|f(t) - g(t)|}{t} \, dt + c_2 \frac{1}{T},$$

where $c_1, c_2 > 0$ are universal constants and $g(t) = e^{-t^2/2}$.

We split the integral into three parts. Put

$$T_0 = \sqrt{\varepsilon}, \quad T_1 = \sqrt{\log \frac{1}{\varepsilon}}.$$

Clearly, $T_0 < T_1$. The value $T > T_1$ will be of order $\varepsilon^{-2/5}$ modulo a logarithmic term and will be specified somewhat later. Anyway, $\log T \leqq C \log(1/\varepsilon)$ with some absolute constant.

By (5.2), the part of the integral (6.1) over the interval $0 < t < T_0$ does not exceed

$$2M(\varepsilon) \sqrt{\varepsilon} \leqq 8 \sqrt{\varepsilon \log \frac{1}{\varepsilon}}.$$

By Lemma 5.1, the integral along the interval $T_0 < t < T_1$ does not exceed

$$4 \sqrt{\varepsilon \log \frac{1}{\varepsilon}} \, (\log T_1 - \log T_0).$$

Since $-\log T_0 = (1/2) \log(1/\varepsilon)$ and $\log T_1$ is of order $\log \log(1/\varepsilon)$, we conclude that

$$\int_{T_0}^{T_1} \frac{|f(t) - g(t)|}{t} \, dt \leqq C \sqrt{\varepsilon} \left(\log \frac{1}{\varepsilon} \right)^{3/2}$$

with some absolute constant C.

To bound the integral along the interval $T_1 < t < T$, we apply Lemma 5.3, which yields the bound $C T^{1/4} \sqrt{\varepsilon \log(1/\varepsilon)}$.

Collecting all these bounds and applying (6.1), we arrive at

$$\|F - \Phi\| \leq C \left[\sqrt{\varepsilon} \left(\log \frac{1}{\varepsilon} \right)^{3/2} + T^{1/4} \sqrt{\varepsilon \log \frac{1}{\varepsilon}} + \frac{1}{T} \right]$$

with some absolute constant C. It remains to choose the value $T = (\varepsilon \log(1/\varepsilon))^{-2/5}$, which is almost optimal. Thus, Theorem 1.2 is proved.

7. Oscillating densities. Let us now focus on Theorem 1.1. First, we shall describe the class of probability distributions on the real line, which provide a counterexample to Kac's conjecture. Given parameters $T > 0$ and $\alpha \in (0, 1]$, put

$$\text{(7.1)} \qquad p_T(x) = \frac{1}{c} \left(1 - \alpha \cos(Tx^2) \right) \varphi(x), \qquad x \in \mathbf{R},$$

where φ denotes the density of the standard normal law, and $c = c_T > 0$ is a normalizing constant (meaning that c depends on both T and α). Denote by F_T the corresponding distribution function,

$$\text{(7.2)} \qquad F_T(x) = \int_{-\infty}^x p_T(y) \, dy = \frac{1}{c} \left[\Phi(x) - \alpha \int_{-\infty}^x \cos(Ty^2) \, d\Phi(y) \right].$$

By construction, the density p_T has oscillations around φ, bounded by a multiple of $\alpha\varphi$. When T is large, this type of oscillation, on the one hand, provides closeness of p_T to φ in the weak sense, but, on the other hand, also separates p_T from the normal in the strong sense.

To carefully develop the participating Fourier transforms, we start with a well-known elementary identity.

LEMMA 7.1. *For all complex numbers a and z such that* $\operatorname{Re} z > 0$,

$$(7.3) \qquad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ax - zx^2/2}\, dx = \frac{1}{\sqrt{z}}\, e^{a^2/(2z)},$$

where $\sqrt{z}$ is the principal branch of the square root function in the half-plane $\operatorname{Re} z > 0$.

For example, applying this identity with $a = 0$, $z = 1 - 2iT$, we find that $\int_{-\infty}^{+\infty} e^{iTx^2} \varphi(x)\, dx = 1/\sqrt{1 - 2iT}$.

Write $\cos(Tx^2) = e^{iTx^2}/2 + e^{-iTx^2}/2$, so that to express the normalizing constant $c = c_T$ in (7.1) explicitly in terms of T and α,

$$c_T = 1 - \frac{\alpha}{2}\left(\frac{1}{\sqrt{1 - 2iT}} + \frac{1}{\sqrt{1 + 2iT}} \right).$$

As an immediate consequence, we obtain the following lemma.

LEMMA 7.2. *We have*

$$(7.4) \qquad |1 - c_T| \leqq \frac{\alpha}{(1 + 4T^2)^{1/4}} < \frac{\alpha}{\sqrt{T}}.$$

Thus, $c_T \to 1$ as $T \to +\infty$, uniformly over all allowed values of α.

From (7.4), whenever $T \geqq 4$, one easily has

$$(7.5) \qquad \frac{\alpha}{4\sqrt{T}} \leqq \left| \frac{1}{c_T} - 1 \right| \leqq \frac{2\alpha}{\sqrt{T}}.$$

For $T \geqq 4$, one also has a uniform bound

$$(7.6) \qquad |p_T(x) - \varphi(x)| \leqq 3\alpha\, \varphi(x), \qquad x \in \mathbf{R}.$$

Moreover, one can easily compute the characteristic function f_T of F_T by applying (7.3) with $a = it$ and $z = 1 \pm 2iT$. Namely, for all $t \in \mathbf{R}$, we get

$$c_T\, f_T(t) = e^{-t^2/2} - \frac{\alpha}{2}\left(\frac{e^{-t^2/(2(1 - 2iT))}}{\sqrt{1 - 2iT}} + \frac{e^{-t^2/(2(1 + 2iT))}}{\sqrt{1 + 2iT}} \right).$$

This implies

$$|c_T\, f_T(t) - e^{-t^2/2}| \leqq \frac{\alpha}{(1 + 4T^2)^{1/4}} < \frac{1}{\sqrt{T}},$$

and using Lemma 7.2, we arrive at the following conclusion.

PROPOSITION 7.1. *Uniformly for all $\alpha \in (0, 1]$, we have $F_T \to \Phi$ weakly as* $T \to +\infty$.

Equivalently, $L(F_T, \Phi) \to 0$ for the Lévy distance.

To test the closeness with respect to the total variation norm, write

$$\int_{-\infty}^{+\infty} |c_T \, p_T(x) - \varphi(x)| \, dx = \frac{\alpha}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} |\cos(Tx^2)| \, e^{-x^2/2} \, dx$$

$$= \frac{\alpha}{\sqrt{2\pi T}} \int_0^{+\infty} \frac{|\cos y|}{\sqrt{y}} \, e^{-y/(2T)} \, dy.$$

Restricting the integration to $0 < y < T$ and assuming $T \geqq 1$, we obtain that

$$\int_{-\infty}^{+\infty} |c_T \, p_T(x) - \varphi(x)| \, dx > \frac{\alpha}{\sqrt{2\pi e \, T}} \int_0^T \frac{|\cos y|}{\sqrt{y}} \, dy > c_T \kappa \alpha$$

for some absolute constant $\kappa > 0$. Hence, using the upper bound in (7.5), we find that

$$\|F_T - \Phi\|_{\mathrm{TV}} = \int_{-\infty}^{+\infty} |p_T(x) - \varphi(x)| \, dx$$

$$\geqq \int_{-\infty}^{+\infty} \left| p_T(x) - \frac{1}{c_T} \, \varphi(x) \right| dx - \left| \frac{1}{c_T} - 1 \right| \geqq \alpha \left(\kappa - \frac{2}{\sqrt{T}} \right).$$

One may now combine this bound with (7.6) and summarize.

PROPOSITION 7.2. *For some absolute constants $T_0 > 0$ and $\kappa > 0$,*

$$\kappa \alpha \leqq \|F_T - \Phi\|_{\mathrm{TV}} \leqq 3\alpha \qquad (T \geqq T_0).$$

8. Computations and bounds for convolutions. In the following we shall study the behavior of the convolutions $p_T * p_T$ for the growing parameter T. Again applying Lemma 7.1 as in the previous section, one can obtain explicit formulas for the densities of convolutions (so, the formula for the characteristic function f_T is not needed).

By analogy to notation for distribution functions, for integrable functions $p(x)$ and $q(x)$ we write $(p * q)(x) = p(x) * q(x) = \int_{-\infty}^{+\infty} p(x - y) \, q(y) \, dy$.

For T real, define

$$u_T(x) = e^{iTx^2} \varphi(x), \qquad x \in \mathbf{R}.$$

In particular, $u_0 = \varphi$, so

$$u_0(x) * u_0(x) = \varphi(x) * \varphi(x) = \frac{1}{2\sqrt{\pi}} \, e^{-x^2/4}.$$

Step 1. More generally, in the integral

$$(8.1) \qquad u_T(x) * \varphi(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{iTy^2} \, e^{-((x-y)^2 + y^2)/2} \, dy$$

we change variable via $y = x/2 + s/\sqrt{2}$. Then the integral (8.1) becomes

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left\{ \frac{iTx^2}{4} \right\} \exp\left\{ \frac{iTx}{\sqrt{2}} \, s + \frac{iTs^2}{2} \right\} \exp\left\{ -\frac{x^2}{4} - \frac{s^2}{2} \right\} \frac{ds}{\sqrt{2}}$$

$$= \frac{e^{-x^2/4}}{2\pi\sqrt{2}} \exp\left\{ \frac{iTx^2}{4} \right\} \int_{-\infty}^{+\infty} \exp\left\{ \frac{iTx}{\sqrt{2}} \, s - \frac{1 - iT}{2} \, s^2 \right\} ds$$

$$= \varphi(x) * \varphi(x) e^{iTx^2/4} \frac{1}{\sqrt{1 - iT}} \exp\left\{ -\frac{T^2 x^2}{4(1 - iT)} \right\},$$

where we used (7.3) with $a = iTx/\sqrt{2}$ and $z = 1 - iT$. Thus,

$$\frac{u_T(x) * \varphi(x)}{\varphi(x) * \varphi(x)} = e^{\frac{iTx^2}{4}} \frac{1}{\sqrt{1 - iT}} \exp\left\{ -\frac{T^2 x^2}{4(1 - iT)} \right\}.$$

Since $1/(1 - iT) = (1 + iT)/(1 + T^2)$ has a positive real part, we obtain that

$$\frac{|u_T(x) * \varphi(x)|}{\varphi(x) * \varphi(x)} \leqq \frac{1}{(1 + T^2)^{1/4}} < \frac{1}{\sqrt{|T|}}.$$

Applying a similar representation to $-T$ and using $\cos(Tx^2)\,\varphi(x) = u_T(x)/2 + u_{-T}(x)/2$, we conclude that

$$(8.2) \qquad \frac{|[\cos(Tx^2)\,\varphi(x)] * \varphi(x)|}{\varphi(x) * \varphi(x)} < \frac{1}{\sqrt{|T|}}.$$

Step 2. Here, as before, consider the convolution

$$u_T(x) * u_T(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{iT\,((x-y)^2 + y^2)}\, e^{-((x-y)^2 + y^2)/2}\, dy$$

$$= \varphi(x) * \varphi(x) e^{iTx^2/2} \frac{1}{\sqrt{1 - 2iT}},$$

which gives $|u_T(x) * u_T(x)| = \varphi(x) * \varphi(x)/(1 + 4T^2)^{1/4}$, and again

$$(8.3) \qquad \frac{|u_T(x) * u_T(x)|}{\varphi(x) * \varphi(x)} < \frac{1}{\sqrt{|T|}}.$$

Step 3. Using $(x - y)^2 - y^2 = -\sqrt{2}\,xs$, we get

$$u_T(x) * u_{-T}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{iT\,((x-y)^2 - y^2)}\, e^{-((x-y)^2 + y^2)/2}\, dy$$

$$= \frac{1}{2\pi\sqrt{2}} e^{-x^4/4} \int_{-\infty}^{+\infty} e^{-iT\sqrt{2}\,xs - s^2/2}\, ds$$

$$= \varphi(x) * \varphi(x) e^{-T^2 x^2},$$

where now we applied (7.3) with $a = -iTx\sqrt{2}$ and $z = 1$. Hence,

$$(8.4) \qquad \frac{u_T(x) * u_{-T}(x)}{\varphi(x) * \varphi(x)} = e^{-T^2 x^2}.$$

Step 4. Furthermore, for $T > 0$, consider the convolution

$$\psi_T(x) = [\cos(Tx^2)\,\varphi(x)] * [\cos(Tx^2)\,\varphi(x)].$$

Representing

$$\psi_T = \frac{u_T + u_{-T}}{2} * \frac{u_T + u_{-T}}{2} = \frac{u_T * u_T + 2u_T * u_{-T} + u_{-T} * u_{-T}}{4}$$

and applying Steps 2 and 3 (i.e., relations (8.3) and (8.4)), we obtain that, for all $x \in \mathbf{R}$,

$$(8.5) \qquad \frac{|\psi_T(x)|}{\varphi(x) * \varphi(x)} < \frac{1}{2} \left(\frac{1}{\sqrt{T}} + e^{-T^2 x^2} \right).$$

Step 5. Given $T > 0$ and $\alpha \in [0, 1]$, at last consider

$$\zeta_T(x) = [(1 - \alpha \cos(Tx^2))\,\varphi(x)] * [(1 - \alpha \cos(Tx^2))\,\varphi(x)]$$
$$= \varphi(x) * \varphi(x) - 2\alpha\,[\cos(Tx^2)\,\varphi(x)] * \varphi(x) + \alpha^2 \psi_T(x).$$

By Step 1 (inequality (8.2)) and Step 4 (inequality (8.5)), we get

$$(8.6) \qquad \left| \frac{\zeta_T(x)}{\varphi(x) * \varphi(x)} - 1 \right| < \frac{5}{2\sqrt{T}} + \frac{1}{2}\,e^{-T^2 x^2}.$$

Finally, for the density p_T, introduced in (7.1), we have $\zeta_T = c_T^2\,p_T * p_T$. Using (8.6), we obtain that

$$(8.7) \qquad \left| \frac{p_T(x) * p_T(x)}{\varphi(x) * \varphi(x)} - 1 \right| < \left| \frac{1}{c_T^2} - 1 \right| + \frac{1}{c_T^2}\left(\frac{5}{2\sqrt{T}} + \frac{1}{2}\,e^{-T^2 x^2} \right).$$

By (7.4) and (7.5), if $T \geqq 4$, we have $|1/c_T - 1| \leqq 2/\sqrt{T}$ and $c_T \geqq 1/2$. Hence, the right-hand side of (8.7) does not exceed

$$\frac{2}{\sqrt{T}} \cdot 3 + 4\left(\frac{5}{2\sqrt{T}} + \frac{1}{2}\,e^{-T^2 x^2} \right) = \frac{16}{\sqrt{T}} + 2\,e^{-T^2 x^2}.$$

Thus, we have arrived at the following result.

LEMMA 8.1. *Given $T \geqq 4$ and $\alpha \in [0, 1]$, for all $x \in \mathbf{R}$,*

$$(8.8) \qquad \left| \frac{p_T(x) * p_T(x)}{\varphi(x) * \varphi(x)} - 1 \right| < \frac{16}{\sqrt{T}} + 2\,e^{-T^2 x^2}.$$

9. Proof of Theorem 1.1. Having established Proposition 7.2 and Lemma 8.1, we may proceed with the last step of the proof of Theorem 1.1. More precisely, for large values of T, we need to verify statement (a) in Theorem 1.1 for the distributions $F = F_T$ described in (7.1)–(7.2). Note that statement (b) is provided by Proposition 7.2, when α is separated from zero (take $\alpha = 1$, for example).

Let us recall that in the setting of an arbitrary probability space (Ω, μ), the Kullback–Leibler distance (also called informational divergence) is well defined for any probability measure ν, which is absolutely continuous with respect to μ. Namely, if u is density of ν with respect to μ, the distance is given by

$$D(\nu \| \mu) = \int_\Omega u\,\log u\,d\mu.$$

LEMMA 9.1. *If $|u - 1| \leqq A$ with some constant $A \geqq 0$, then*

$$(9.1) \qquad D(\nu \| \mu) \leqq \log((A + 1)e) \int_\Omega |u - 1|\,d\mu.$$

Indeed, if $u \geqq 1/e$, we have $|u \log u| \leqq \log((A + 1)e)\,|u - 1|$ by the mean value theorem. For $0 \leqq u \leqq 1/e$ the latter inequality holds automatically, since $u \log u$ is decreasing on that interval.

Proof of Theorem 1.1. On the real line $\Omega = \mathbf{R}$, taking μ to be the normal law $N(0, 2)$ with mean zero and variance 2, and taking ν to be the distribution $F_T * F_T$, the corresponding density u in Lemma 9.1 is given by

$$u = \frac{d\nu}{d\mu} = \frac{p_T * p_T}{\varphi * \varphi}.$$

By Lemma 8.1 with $T \geqq 4$, we have $|u-1| \leqq 8.5$. Hence, the assumption of Lemma 9.1 is fulfilled with $A = 8.5$. Using $\log((A+1)e) < 4$, we obtain by (9.1) and (8.8) that

$$(9.2) \qquad D(F_T * F_T \| N(0,2)) \leqq 4 \int_{-\infty}^{+\infty} \left(\frac{16}{\sqrt{T}} + 2\, e^{-T^2 x^2} \right) d\mu(x) \leqq \frac{C}{\sqrt{T}},$$

where C is an absolute constant.

Finally, let us apply the well-known fact that, if a random variable ξ has a distribution ν, which is absolutely continuous with respect to the Lebesgue measure and has a finite second moment, then the entropic distance $D(\xi)$ to normality represents the shortest Kullback–Leibler distance from ν to the class of all normal laws. That is,

$$(9.3) \qquad \min_{a,\sigma} D\big(\nu \| N(a, \sigma^2)\big) = D(\xi)$$

with minimum being attained for $a = \mathbf{E}\,\xi$, $\sigma^2 = \mathbf{D}\xi$. Hence, if X_T and Y_T are independent random variables distributed according to F_T, then by (9.2)–(9.3), applied with $\xi = X_T + Y_T$, we obtain that $D(X_T + Y_T) \leqq C/\sqrt{T} \to 0$ as $T \to +\infty$. Note that there is no need to compute the parameters a and σ^2 for ξ.

Hence, Theorem 1.1 is now proved.

10. Proof of Theorem 1.3. To prove Theorem 1.3, we introduce another family of probability distributions on the real line, which oscillate around the standard normal law and can be described by explicit formulas for their distribution functions. Using distribution functions seems more convenient for estimating the Kolmogorov distances compared to an "explicit density approach." This family has a much simpler structure than the family described in the proof of Theorem 1.1, and it can also be used to construct a counterexample to Kac's question—however, not in the i.i.d. case.

Given T large enough (say, $T \geqq 6$), consider the distribution function

$$F_T(x) = \Phi(x) + \frac{1}{2T}\, \cos(T_0 x)\, \varphi(x)\, 1_{\{|x| < T_0\}},$$

where $T_0 = 4\sqrt{\log T}$ and, as before, Φ and φ stand for the density and the distribution function of the standard normal law. If $\cos(T_0^2) = 0$, which we may assume from now on, F_T has density

$$p_T(x) = \varphi(x) - \frac{1}{2} \left(\frac{T_0}{T}\, \sin(T_0 x) + \frac{x}{T}\, \cos(T_0 x) \right) \varphi(x)\, 1_{\{|x| < T_0\}}.$$

Since $T_0 < T$, the function p_T is positive and continuous everywhere (except for $|x| = T_0$). Since in addition $F_T(-\infty) = 0$ and $F_T(+\infty) = 1$, F_T is indeed a distribution function.

It is easy to check that, if X_T is a random variable with distribution F_T, we have $\mathbf{E}\, X_T \to 0$ and $\mathbf{D}X_T \to 1$ as $T \to +\infty$.

It follows immediately from the definition that $|F_T(x) - \Phi(x)|$ attains its maximum at the origin, which we formulate below in terms of the Kolmogorov distance.

LEMMA 10.1. *For all $T > 0$, we have $\|F_T - \Phi\| = 1/(2T\sqrt{2\pi})$.*

Hence, to prove Theorem 1.3, it will be sufficient to show that for the corresponding convolutions the distance $\|F_T * F_T - \Phi * \Phi\|$ is of order $1/T^2$ modulo a logarithmic factor. Consequently, one may take ε to be of order $1/T^2$.

The remaining part of this section is devoted to the confirmation of this conjecture. Here and below, the convolution operator is applied to functions of bounded variation that are regarded as positive or signed measures. For example, $(F_T * F_T)(x) = \int_{-\infty}^{+\infty} F_T(x - y)\, dF_T(y)$.

Rewrite F_T in the form $F_T = \Phi + G_T - H_T$, where

$$(10.1) \qquad G_T(x) = \frac{1}{2T}\,\cos(T_0 x)\,\varphi(x), \qquad H_T(x) = \frac{1}{2T}\,\cos(T_0 x)\,\varphi(x)\, 1_{\{|x| \geqq T_0\}}.$$

Hence,

$$(10.2) \qquad F_T * F_T = (\Phi + G_T) * (\Phi + G_T) - 2\,(\Phi + G_T) * H_T + H_T * H_T.$$

Obviously, we have the bounds $\|G_T\|_{\mathrm{TV}} < 1$ and $\|H_T\|_{\mathrm{TV}} < e^{-T_0^2/2}$, and hence the sum of the last two terms in (10.2) has total variation less than $5\,e^{-T_0^2/2} = 5/T^8$.

The remaining convolutions in (10.2) are readily computable. Changing the variable $y = x/2 + s/\sqrt{2}$, we find by direct computations that

$$\begin{aligned} 2T\,(G_T * \Phi)(x) &= 2T \int_{-\infty}^{+\infty} G_T(x - y)\, d\Phi(y) \\ &= \int_{-\infty}^{+\infty} \cos(T_0(x - y))\,\varphi(x - y)\,\varphi(y)\, dy \\ &= \frac{1}{2\sqrt{\pi}}\,\cos\left(\frac{T_0 x}{2}\right) e^{-x^2/4}\, e^{-T_0^2/4}. \end{aligned}$$

It easily follows that

$$\|G_T * \Phi\|_{\mathrm{TV}} = \frac{e^{-T_0^2/4}}{4T\sqrt{\pi}} \int_{-\infty}^{+\infty} \left| \frac{d}{dx}\,\cos\left(\frac{T_0 x}{2}\right) e^{-x^2/4} \right| dx < e^{-T_0^2/4} = \frac{1}{T^4}.$$

Taking into account the previous bounds, we get the following preliminary conclusion.

LEMMA 10.2. *If $T \geqq 6$, we have $F_T * F_T = \Phi * \Phi + G_T * G_T + U_T$, where the function G_T is defined in* (10.1) *and U_T has a total variation less than $7/T^4$.*

It remains to proceed to the last step. The following simple identity is verified directly, so we omit the proof. Recall that $T_0 = 4\sqrt{\log T}$ appears in the definition of G_T.

LEMMA 10.3. *For all x,*

$$4T^2\,(G_T * G_T)(x) = -\frac{T_0}{4\sqrt{\pi}}\,\sin(T_0 x)\, e^{-x^2/4} - \frac{x e^{-x^2/4}}{8\sqrt{\pi}}\,(\cos(T_0 x) + e^{-T_0^2}).$$

Proof of Theorem 1.3.   In view of the equality of Lemma 10.3, we first note that $x e^{-x^2/4} < 1$ and obtain that $4T^2 |(G_T * G_T)(x)| < T_0$, whenever $T \geqq 6$. By Lemma 10.2,

$$\|F_T * F_T - \Phi * \Phi\| < \frac{T_0}{4T^2} = \frac{\sqrt{\log T}}{T^2}.$$

Hence, choosing $T = \varepsilon^{-1/2} \log^{1/4}(1/\varepsilon)$ with $\varepsilon > 0$ small enough, we get $\|F_T * F_T - \Phi * \Phi\| < \varepsilon$, while, by Lemma 10.1,

$$\|F - \Phi\| = \frac{1}{2T\sqrt{2\pi}} = \frac{1}{2\sqrt{2\pi}}\,\varepsilon^{1/2}\left(\log\frac{1}{\varepsilon}\right)^{-1/4}.$$

Thus, Theorem 1.3 is proved.

Remark 10.1. Using Lemma 10.3, one can see that there is a stronger property, namely, $T^2 \|G_T * G_T\|_{\mathrm{TV}} < CT_0^2$, which, by Lemma 10.2, leads to

$$\|F_T * F_T - \Phi * \Phi\|_{\mathrm{TV}} < C\,\frac{\log T}{T^2}$$

with some absolute constant C. Hence, modulo a logarithmic factor, the statement of Proposition 7.1 may be sharpened: For $\varepsilon > 0$ small enough and T of the same order $\varepsilon^{-1/2}\log^{1/4}(1/\varepsilon)$, we have

$$\|F_T * F_T - \Phi * \Phi\|_{\mathrm{TV}} < \varepsilon, \quad \text{while} \quad \|F - \Phi\| > c\varepsilon^{1/2}\bigg(\log\frac{1}{\varepsilon}\bigg)^{-1/2},$$

where $c > 0$ is an absolute constant.

REFERENCES

[1] S. G. BOBKOV, G. P. CHISTYAKOV, AND F. GÖTZE, *Entropic instability of Cramér's characterization of the normal law*, in Selected Works of Willem van Zwet, Springer, New York, 2012, pp. 231–242.

[2] S. G. BOBKOV, G. P. CHISTYAKOV, AND F. GÖTZE, *Regularized Distributions and Entropic Stability of Cramér's Characterization of the Normal Law*, preprint, University of Bielefeld, Bielefeld, 2010.

[3] G. P. CHISTYAKOV, *The sharpness of the estimates in theorems on the stability of decompositions of a normal distribution and a Poisson distribution*, Teor. Funkciĭ Funksional. Anal. i Prilozhen., 26 (1976), pp. 119–128, iii (in Russian).

[4] G. P. CHISTYAKOV, *A remark on a theorem of N. A. Sapogov on the stability of decompositions of a normal distribution*, in Operators in Function Spaces and Problems in Function Theory, Naukova Dumka, Kiev, 1987, pp. 108–116, 147 (in Russian).

[5] G. P. CHISTYAKOV AND L. B. GOLINSKII, *Order-sharp estimates for the stability of decompositions of the normal distribution in the Lévy metric*, J. Math. Sci., 72 (1994), pp. 2848–2871.

[6] G. CHRISTOPH, YU. V. PROHOROV, AND V. V. ULYANOV, *Characterization and stability problems for finite quadratic forms*, in Asymptotic Methods in Probability and Statistics with Applications, (St. Petersburg, 1998), Birkhäuser Boston, Boston, MA, 2001, pp. 39–50.

[7] H. CRAMÉR, *Über eine Eigenschaft der Normalen Verteilungsfunktion*, Math. Zeitschrift, 41 (1936), pp. 405–414.

[8] R. M. DUDLEY, *Real Analysis and Probability*, Cambridge University Press, Cambridge, UK, 2002.

[9] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed., John Wiley, New York, 1971.

[10] V. V. IL'IN, *On the quantitative stability of the Cramér theorem in a uniform metric in the case of identically distributed random variables*, Dokl. Akad. Nauk SSSR, 260 (1981), pp. 525–526 (in Russian).

[11] A. M. KAGAN, YU. V. LINNIK, AND C. R. RAO, *Characterization Problems in Mathematical Statistics*, John Wiley, New York, 1973.

[12] YU. V. LINNIK AND I. V. OSTROVSKII, *Decompositions of random variables and random vectors*, Amer. Math. Soc., Providence, RI, 1977.

[13] S. G. MALOSHEVSKII, *Sharpness of an estimate of N. A. Sapogov in the stability problem of Cramér's theorem*, Theory Probab. Appl., 13 (1968), pp. 494–496.

[14] S. G. MALOSHEVSKII, *Bounds connected with the stability theorems of H. Cramér's theorem*, Math. Notes, 7 (1970), pp. 171–175.

[15] H. P. MCKEAN, JR., *Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas*, Arch. Rational Mech. Anal., 21 (1966), pp. 343–367.

[16] V. V. PETROV, *Sums of Independent Random Variables*, Springer-Verlag, Berlin, 1975.

[17] N. A. SAPOGOV, *The stability problem for a theorem of Cramér*, Izvestiya Akad. Nauk SSSR.
        Ser. Mat., 15 (1951), pp. 205–218 (in Russian); English translation in Select. Transl. Math.
        Statist. and Probability, Vol. 1, Inst. Math. Statist. and Amer. Math. Soc., Providence,
        RI, 1961, pp. 41–53.
[18] N. A. SAPOGOV, *The problem of stability for a theorem of Cramér*, Vestnik Leningrad. Univ.,
        10 (1955), pp. 61–64 (in Russian).
[19] V. V. SENATOV, *Sharpening stability estimates for Cramér's theorem*, J. Soviet Math., 16
        (1981), pp. 1425–1432.
[20] V. M. ZOLOTAREV AND V. V. SENATOV, *Two-sided estimates of Lévy's metric*, Theory Probab.
        Appl., 20 (1975), pp. 234–245.
[21] I. S. SHIGANOV, *Some estimates related to the stability of a theorem of H. Cramér*, J. Soviet
        Math., 17 (1981), pp. 2351–2358.
[22] I. S. SHIGANOV, *On stability estimates of Cramér's theorem*, in Stability Problems for Stochastic
        Models (Varna, 1985), Lecture Notes in Math. 1233, Springer, Berlin, 1987, pp. 178–181.
[23] V. M. ZOLOTAREV, *On the problem of stability of the decomposition of the normal law into
        components*, Theory Probab. Appl., 13 (1968), pp. 697–700.
[24] V. M. ZOLOTAREV, *Several new probabilistic inequalities connected with the Lévy metric*, Dokl.
        Akad. Nauk SSSR, 190 (1970), pp. 1019–1021 (in Russian).
[25] V. M. ZOLOTAREV, *Estimates of the difference between distributions in the Lévy metric*, Proc.
        Steklov Inst. Math., 112 (1971), pp. 232–240.