DOI 10.1007/s10958-024-07134-9
Journal of Mathematical Sciences, Vol. 281, No. 4, May, 2024

ON THE REMAINDER TERM IN THE APPROXIMATE FOURIER INVERSION FORMULA FOR DISTRIBUTION FUNCTIONS

S. G. Bobkov
University of Minnesota
127 Vincent Hall, Minneapolis, MN 55455, USA
bobkov@math.umn.edu

We discuss several uniform bounds on the remainder term in the Fourier inversion formula for increments of distribution functions. These bounds are illustrated by some discrete examples related to the binomial distribution. Bibliography: 10 *titles.*

Dedicated to N. N. Uraltseva

1 Introduction

The transition from characteristic functions to corresponding distribution functions is commonly performed with the help of the Fourier inversion formula

$$F(x) - F(y) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\, dt, \tag{1.1}$$

where

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x) \tag{1.2}$$

denotes the Fourier–Stieltjes transform (the characteristic function) of an arbitrary Borel probability measure μ on the real line with the associated distribution function $F(x) = \mu((-\infty, x])$ and $x, y \in \mathbb{R}$ are points of continuity of F.

Although the convergence in (1.1) might not be uniform with respect to x, y, in various asymptotic problems it is desirable to have a uniform bound for the error of approximation

$$\delta_F(T) = \sup_{x,y} \left| (F(x) - F(y)) - \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\, dt \right| \tag{1.3}$$

Translated from *Problemy Matematicheskogo Analiza* **127**, 2024, pp. 83-96.

for large values of T. One natural bound which immediately follows from (1.1) is given by

$$\delta_F(T) \leqslant \frac{2}{\pi} \int\limits_T^\infty \frac{|f(t)|}{t}\, dt. \tag{1.4}$$

If the measure μ is absolutely continuous and has a density of bounded total variation, then $f(t) = O(1/t)$ as $t \to \infty$, and (1.4) yields $\delta_F(T) = O(1/T)$. However, in general, the integral in (1.4) can be divergent.

For quantified statements, one can also use the Lévy (maximal) concentration function

$$Q_F(h) = \sup_x \mathbf{P}\{x \leqslant X \leqslant x + h\} = \sup_x (F(x+h) - F(x-)), \quad h \geqslant 0,$$

where X is a random variable with distribution μ. For example, suppose that μ is unimodal (i.e., it has a density $p(x)$ which is nondecreasing for $x < a$ and is nonincreasing for $x > a$ for some point $a \in \mathbb{R}$). In this case, it was shown by Ushakov [1] that, for all $t > 0$,

$$|f(t)| \leqslant Q_F(\pi/t)$$

(see also [2, p. 95]). Using this pointwise bound in (1.4), we find

$$\delta_F(T) \leqslant \frac{2}{\pi} \int\limits_0^{\pi/T} \frac{Q_F(h)}{h}\, dh. \tag{1.5}$$

In this paper, we consider a general situation (including discrete probability distributions), thus removing any constraint on the shape of the distribution.

Proposition 1.1. *Given a distribution function F, for all $T > 0$*

$$\delta_F(T) \leqslant \frac{2}{1+T} + 4T \int\limits_0^1 \frac{Q_F(h)}{(1+Th)^2}\, dh. \tag{1.6}$$

Under quasi-Lipschitz conditions posed on F, the last integral can be further estimated.

Corollary 1.2. *If the distribution function F satisfies*

$$|F(x) - F(y)| \leqslant M\,(\varepsilon + |x-y|), \quad x, y \in \mathbb{R}, \tag{1.7}$$

with some $M \geqslant 0$ and $\varepsilon \geqslant 0$, then for all $T \geqslant 2$

$$\delta_F(T) \leqslant \frac{2}{T} + 4M\Big(\varepsilon + \frac{\log T}{T}\Big). \tag{1.8}$$

If $M \geqslant 1$, one can simplify the above inequality as the representation

$$F(x) - F(y) = \frac{1}{2\pi} \int\limits_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\, dt + \theta M \left(\varepsilon + \frac{\log T}{T}\right)$$

with a quantity θ bounded in absolute value.

The logarithmic term in (1.8) cannot be removed under the condition (1.7), even with $\varepsilon = 0$, i.e., when μ has a bounded density p. In this case, let us introduce the functional

$$M(F) = \|F\|_{\mathrm{Lip}} = \operatorname*{ess\,sup}_x p(x),$$

where $\|F\|_{\mathrm{Lip}}$ denotes the Lipschitz seminorm (with respect to the Euclidean distance).

Proposition 1.3. *For $M > 0$ and $T \geqslant 2M$*

$$c_0 \, \frac{\log(T/M)}{T/M} \leqslant \sup_{M(F)=M} \delta_F(T) \leqslant c_1 \, \frac{\log(T/M)}{T/M} \tag{1.9}$$

with some absolute constants $c_1 > c_0 > 0$.

These relations are invariant under linear transformations: (1.9) does not change when the random variable X with distribution function F is multiplied by any positive constant.

As for distribution functions of class $\mathrm{Lip}(\alpha)$ with parameter $\alpha < 1$, there is a similar upper bound, but without the logarithmic term.

Corollary 1.4. *Let $0 < \alpha < 1$. If the distribution function F satisfies*

$$|F(x) - F(y)| \leqslant M\,(\varepsilon + |x-y|^\alpha), \quad x, y \in \mathbb{R},$$

with some $M \geqslant 0$ and $\varepsilon \geqslant 0$, then for all $T > 0$,

$$\delta_F(T) \leqslant \frac{2}{T} + 4M\Big(\varepsilon + \frac{1}{(1-\alpha)\,T^\alpha}\Big).$$

If $\varepsilon = 0$, this bound is consistent with what is obtained on the basis of the inequality (1.5), up to an α-depending factor.

The right-hand side of (1.6) can also be related to the characteristic function f associated to F, by applying Esseen's upper bound

$$Q_F(h) \leqslant ch \int_0^{1/h} |f(t)|\,dt, \quad h > 0,$$

where c is an absolute constant (see [3]). This leads to the inequality

$$\delta_F(T) \leqslant \frac{2}{T} + \frac{c \log T}{T} \int_0^T |f(t)|\,dt, \quad T \geqslant 2.$$

However, here the logarithmic term can be removed. One smoothing type result by Prawitz [4] implies the following sharpening of the upper bound (1.4).

Proposition 1.5. *Let X be a random variable with distribution function F and characteristic function f. For any $T > 0$,*

$$\delta_F(T) \leqslant \frac{2}{T} \int_0^T |f(t)|\,dt. \tag{1.10}$$

In particular, if $f(t)$ is nonnegative, then with some absolute constant $c > 0$,

$$\delta_F(T) \leqslant c\,\mathbf{P}\{|X| \leqslant 1/T\}. \tag{1.11}$$

If, additionally, X has a bounded density (which is equivalent to the integrability of f when this function is nonnegative), the latter inequality yields

$$\delta_F(T) \leqslant 2c\,\frac{M(F)}{T}. \tag{1.12}$$

This improves upon (1.8).

2 Functions of Bounded Total Variation

Proposition 1.1 is a consequence of a more general assertion for the class of functions F of bounded total variation on the real line. Denote by $|dF(z)|$ the variation of F viewed as a finite positive Borel measure on the real line with total variation norm $\|F\|_{\mathrm{TV}}$.

Proposition 2.1. *Let F be a function of bounded total variation with the Fourier–Stieltjes transform f defined by* (1.2)*. For all $x, y \in \mathbb{R}$ and $T > 0$*

$$\begin{aligned} F(x) - F(y) = {} & \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\,dt \\ & + \theta_1 \int_{-\infty}^{\infty} \frac{|dF(z)|}{1 + T\,|z - x|} + \theta_2 \int_{-\infty}^{\infty} \frac{|dF(z)|}{1 + T\,|z - y|} \end{aligned} \tag{2.1}$$

with some complex numbers θ_1 and θ_2 such that $|\theta_j| \leqslant 1$.

The last two integrals in (2.1) are bounded by $\|F\|_{\mathrm{TV}}$. Since also $|F(x) - F(y)| \leqslant \|F\|_{\mathrm{TV}}$, we see that the error function (1.3) is uniformly bounded, namely,

$$\delta_F(T) \leqslant 3\,\|F\|_{\mathrm{TV}}, \quad T \geqslant 0.$$

Moreover, by the Lebesgue dominated convergence theorem, these integrals are convergent to zero as $T \to \infty$, as long as x and y are points of continuity of F, and then in the limit we return to (1.1). Hence (2.1) can serve as a quantification of the Fourier inversion formula.

Now, introduce the function

$$R(t) = \int_0^t \frac{\sin u}{u}\,du, \quad t \in \mathbb{R}.$$

It satisfies $R(t) \to \pi/2$ as $t \to \infty$ and $R(-t) = -R(t)$ for any $t > 0$. Also, put

$$r(t) = \int_t^{\infty} \frac{\sin u}{u}\,du = \frac{\pi}{2} - R(t). \tag{2.2}$$

As a preliminary step towards the proof of Proposition 2.1, first let us prove the following assertion.

Lemma 2.2. *For all $t \geqslant 0$*

$$|r(t)| \leqslant \frac{\pi}{1+t}\,. \tag{2.3}$$

Proof. Integrating by parts with $t > 0$, we have

$$\int_t^\infty \frac{\sin u}{u}\,du = \frac{\cos t}{t} - \int_t^\infty \frac{\cos u}{u^2}\,du \tag{2.4}$$

which implies

$$|r(t)| \leqslant \frac{2}{t} \leqslant \frac{\pi}{1+t}, \quad t \geqslant t_0 \equiv \frac{1}{\frac{\pi}{2}-1} \sim 1.752\ldots\,.$$

To treat the values $0 \leqslant t \leqslant t_0$, consider the function

$$\psi(t) = r(t) - \frac{\pi}{1+t} = \frac{\pi}{2} - \int_0^t \frac{\sin u}{u}\,du - \frac{\pi}{1+t}.$$

Using the inequality

$$\sin u \geqslant u - \frac{u^3}{6} \quad (u > 0),$$

we get

$$\psi(t) \leqslant v(t) \equiv \frac{\pi}{2} - t + \frac{t^3}{18} - \frac{\pi}{1+t}, \quad t \geqslant 0.$$

To show that $v(t) \leqslant 0$ in the interval $0 \leqslant t \leqslant t_0$, consider the polynomial

$$P(t) = (1+t)v(t) = (1+t)\left(\frac{\pi}{2} - t + \frac{t^3}{18}\right) - \pi.$$

We have $P(0) = v(0) = -\pi/2$ and

$$P'(t) = \frac{\pi}{2} - 1 - 2t + \frac{t^2}{6} + \frac{2t^3}{9}, \quad P'(0) = \frac{\pi}{2} - 1.$$

Since also

$$P''(t) = \frac{2}{3}\,(t+2)\Big(t - \frac{3}{2}\Big),$$

we conclude that $P(t)$ is concave in $0 \leqslant t \leqslant 3/2$ and is convex in $t \geqslant 3/2$. This implies that, on the first interval,

$$P(t) \leqslant P(0) + P'(0)t \leqslant -\frac{\pi}{2} + \left(\frac{\pi}{2} - 1\right)\frac{3}{2} = \frac{\pi}{4} - \frac{3}{2} < 0.$$

Since $P(t_0) = -2.82... < 0$, we also have, by convexity, $P(t) \leqslant 0$ in $3/2 \leqslant t \leqslant t_0$. Thus, $P(t) \leqslant 0$ for all $0 \leqslant t \leqslant t_0$, and the same is true for $v(t)$ and $\psi(t)$ as well, i.e., $r(t) \leqslant \pi/(1+t)$.

As the next step, consider the function

$$\psi(t) = -r(t) - \frac{\pi}{1+t} = -\frac{\pi}{2} + \int\limits_0^t \frac{\sin u}{u}\,du - \frac{\pi}{1+t}.$$

Using $\sin u \leqslant u$, we have

$$\psi(t) \leqslant v(t) \equiv -\frac{\pi}{2} + t - \frac{\pi}{1+t}.$$

The function $v(t)$ is increasing, so that

$$v(t) \leqslant v(t_0) < v(2) = 2 - \frac{5\pi}{6} < 0.$$

Thus, $\psi(t) \leqslant 0$, i.e., $-r(t) \leqslant \pi/(1+t)$. The two bounds yield the desired inequality (2.3). $\square$

Proof of Proposition 2.1. By the Fubini theorem,

$$I \equiv \int\limits_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\,dt = \int\limits_{-\infty}^{\infty} \left[\int\limits_{-T}^{T} \frac{e^{it(z-x)} - e^{it(z-y)}}{-it}\,dt\right] dF(z)$$

$$= -2\int\limits_{-\infty}^{\infty} \left[\int\limits_{0}^{T} \frac{\sin(t(z-x)) - \sin(t(z-y))}{t}\,dt\right] dF(z).$$

Hence, in terms of the function R, we obtain the general representation

$$\frac{1}{2}\,I = \int\limits_{-\infty}^{\infty} [R(T(z-y)) - R(T(z-x))]\,dF(z).$$

We can assume that x, y are points of continuity of F and $x > y$. Splitting the integration into the three regions, write

$$\frac{1}{2}\,I = \int\limits_{-\infty}^{y} [R(T(x-z)) - R(T(y-z))]\,dF(z)$$

$$+ \int\limits_{x}^{\infty} [R(T(z-y)) - R(T(z-x))]\,dF(z) + \int\limits_{y}^{x} [R(T(z-y)) + R(T(x-z))]\,dF(z).$$

Equivalently, by the definition (2.2),

$$\frac{1}{2}\,I = \int\limits_{-\infty}^{y} [r(T(y-z)) - r(T(x-z))]\,dF(z) + \int\limits_{x}^{\infty} [r(T(z-x)) - r(T(z-y))]\,dF(z)$$

$$+ \int\limits_{y}^{x} [\pi - r(T(z-y)) - r(T(x-z))]\,dF(z).$$

Let us rewrite this equality as

$$\frac{1}{2}\,I - \pi(F(x) - F(y)) = \int\limits_{-\infty}^{y} [r(T(y-z)) - r(T(x-z))]\,dF(z)$$

$$+ \int\limits_{x}^{\infty} [r(T(z-x)) - r(T(z-y))]\,dF(z) - \int\limits_{y}^{x} [r(T(z-y)) + r(T(x-z))]\,dF(z). \qquad (2.5)$$

Applying the bound (2.3), we get

$$\Big|\frac{1}{2}\,I - \pi(F(x) - F(y))\Big| \leqslant \int\limits_{-\infty}^{y} \frac{\pi}{1+T\,(y-z)}\,dF(z) + \int\limits_{-\infty}^{y} \frac{\pi}{1+T\,(x-z)}\,dF(z)$$

$$+ \int\limits_{x}^{\infty} \frac{\pi}{1+T\,(z-x)}\,dF(z) + \int\limits_{x}^{\infty} \frac{\pi}{1+T\,(z-y)}\,dF(z) + \int\limits_{y}^{x} \Big(\frac{\pi}{1+T\,(z-y)} + \frac{\pi}{1+T\,(x-z)}\Big)\,dF(z).$$

As a result,

$$\Big|\frac{1}{2\pi}\,I - (F(x) - F(y))\Big| \leqslant \int\limits_{-\infty}^{\infty} \Big(\frac{1}{1+T\,|z-y|} + \frac{1}{1+T\,|x-z|}\Big)\,dF(z).$$

The proposition is proved. □

3 Proof of Proposition 1.1, Corollaries 1.2 and 1.4

From now on, let F be a distribution function. In this case, the relation (2.1) is simplified to

$$F(x) - F(y) = \frac{1}{2\pi} \int\limits_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\,f(t)\,dt + \theta_1 \int\limits_{-\infty}^{\infty} \frac{dF(z)}{1+T\,|z-x|} + \theta_2 \int\limits_{-\infty}^{\infty} \frac{dF(z)}{1+T\,|z-y|} \qquad (3.1)$$

with some complex numbers θ_j such that $|\theta_j| \leqslant 1$.

Proof of Proposition 1.1. To estimate the last integral in (3.1), assume without loss of generality that $y = 0$ and that it is the point of continuity of F. First, note that

$$\int\limits_{a}^{\infty} \frac{1}{1+Tz}\,dF(z) \leqslant \frac{1}{1+Ta}\,(1 - F(a)),$$

where $a > 0$ is a point of continuity of F. On the other hand, integrating by parts, we have

$$\int\limits_{0}^{a} \frac{1}{1+Tz}\,dF(z) = \frac{1}{1+Ta}\,(F(a) - F(0)) + T \int\limits_{0}^{a} \frac{F(z) - F(0)}{(1+Tz)^2}\,dz$$

$$\leqslant \frac{1}{1+Ta}\,(F(a) - F(0)) + T \int\limits_{0}^{a} \frac{Q_F(z)}{(1+Tz)^2}\,dz.$$

Combining the two estimates and letting $a \to 1$, we get

$$\int_0^\infty \frac{1}{1+Tz}\,dF(z) \leqslant \frac{1}{1+T}\,(1-F(0)) + T\int_0^1 \frac{Q_F(z)}{(1+Tz)^2}\,dz.$$

By a similar argument,

$$\int_{-\infty}^0 \frac{1}{1+T|z|}\,dF(z) \leqslant \frac{1}{1+T}\,F(0) + T\int_0^1 \frac{Q_F(z)}{(1+Tz)^2}\,dz,$$

so that

$$\int_{-\infty}^\infty \frac{1}{1+T|z|}\,dF(z) \leqslant \frac{1}{1+T} + 2T\int_0^1 \frac{Q_F(z)}{(1+Tz)^2}\,dz.$$

More generally, for all $y \in \mathbb{R}$

$$\int_{-\infty}^\infty \frac{1}{1+T|z-y|}\,dF(z) \leqslant \frac{1}{1+T} + 2T\int_0^1 \frac{Q_F(z)}{(1+Tz)^2}\,dz.$$

By (3.1), the error function (1.3) admits the upper bound (1.6). □

Proof of Corollaries 1.2 and 1.4. In the setting of Corollary 1.2, $Q_F(h) \leqslant M(\varepsilon + h)$ for all $h \geqslant 0$. Hence the integral in (1.6) does not exceed

$$M\int_0^1 \frac{\varepsilon+h}{(1+Th)^2}\,dh \leqslant \frac{M\varepsilon}{T} + \frac{M}{T^2}\Big(\log(1+T) - \frac{T}{1+T}\Big).$$

Here, the expression in the brackets is smaller than $\log T$ for $T \geqslant 2$.

In Corollary 1.4, we assume that $Q_F(h) \leqslant M(\varepsilon + h^\alpha)$, $h \geqslant 0$. Then the integral in (1.6) is bounded by

$$\begin{aligned} M\int_0^\infty \frac{\varepsilon+h^\alpha}{(1+Th)^2}\,dh &= \frac{M\varepsilon}{T} + \frac{M}{T^{\alpha+1}}\int_0^\infty \frac{u^\alpha}{(1+u)^2}\,du \\ &< \frac{M\varepsilon}{T} + \frac{M}{T^{\alpha+1}}\int_0^\infty \frac{du}{(1+u)^{2-\alpha}} = \frac{M\varepsilon}{T} + \frac{M}{(1-\alpha)\,T^{\alpha+1}}. \end{aligned}$$

The corollaries are proved. □

Remark 3.1. In connection with the use of the function Q_F in Proposition 1.1, one can also recall the Kawata mean concentration function

$$C_F(h) = \frac{1}{h}\int_{-\infty}^\infty (F(x+h) - F(x))^2\,dx, \quad h \geqslant 0,$$

which is related to the maximal concentration function via the inequalities

$$\frac{1}{2}\,Q_F(h/2)^2 \leqslant C_F(h) \leqslant Q_F(h).$$

The relationship between the behavior of $Q_F(h)$ and $C_F(h)$ at $h = 0$ in the form of Lipschitz properties of F and that of the characteristic function $f(t)$ at infinity were studied by Kawata [5] and Makabe [6]. Some portion of connections is based on the Parseval identity

$$C_F(2h) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\sin^2(ht)}{ht^2}\,|f(t)|^2\,dt.$$

4 Proof of Proposition 1.3

First, let us verify that the inequality (1.9) is invariant with respect to linear transformations of a random variable X with distribution functions F. Define

$$I_{F,T}(x,y) = \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\,f(t)\,dt,$$

where f is the characteristic function of X. For $\lambda > 0$ the random variable λX has respectively the distribution and characteristic functions $F_\lambda(x) = F(x/\lambda)$, $f_\lambda(t) = f(\lambda t)$, $(x, t \in \mathbb{R})$. Hence

$$I_{F_\lambda,T}(x,y) = \int_{-\lambda T}^{\lambda T} \frac{e^{-itx/\lambda} - e^{-ity/\lambda}}{-it}\,f(t)\,dt = I_{F,\lambda T}(x/\lambda, y/\lambda),$$

and it follows from the definition (1.3) that $\delta_{F_\lambda}(T) = \delta_F(\lambda T)$.

In addition, $M(F_\lambda) = M(F)/\lambda$. Therefore, if (1.9) holds for F with an arbitrary value $T \geqslant 2M(F)$, it will hold automatically for F_λ with $T \geqslant 2M(F_\lambda)$.

As a consequence, to prove the upper bound in (1.9), we can assume without loss of generality that $M = 1$. But then, by Corollary 1.2, for any $T \geqslant 2$

$$\delta_F(T) \leqslant \frac{2}{T} + 4\,\frac{\log T}{T} \leqslant 7\,\frac{\log T}{T},$$

i.e., we obtain (1.9) with $c_1 = 7$.

Let us now turn to the lower bound. By the homogeneity with respect to X, assume again that $M = 1$. Then we need to show that

$$\delta_F(T) \geqslant c_0\,\frac{\log T}{T} \tag{4.1}$$

for some distribution function F such that $M(F) = 1$. So, fix $T \geqslant 2$.

Suppose that F corresponds to the probability measure μ which is supported on the interval $(0, 2\pi)$ and is symmetric about the point π. In particular, $x = 2\pi + 2\pi m/T$ and $y = -2\pi m/T$ are points of continuity of F for any integer $m \geqslant 1$ (which will be chosen later on), with $F(x) = 1$, $F(y) = 0$, so that $F(x) - F(y) = 1$.

As in the proof of Proposition 2.1, define

$$I = \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\, dt.$$

Note that the first two integrals in (2.5) are vanishing, and this identity is simplified to

$$\frac{1}{2}\, I - \pi = -\int_0^{2\pi} (r(T(z-y) + r(T(x-z)))\, dF(z) = -2\int_0^{2\pi} r(T(z-y))\, dF(z),$$

where we used the symmetry assumption at the last step. This gives

$$\delta_F(T) \geqslant \frac{2}{\pi} \int_0^{2\pi} r(T(z-y))\, dF(z). \tag{4.2}$$

Put $T_0 = [T]$ and define Δ to be the union of the intervals of the form

$$\Delta_k = \frac{2\pi}{T}\,(k-h, k+h), \quad k = 1, \dots, T_0 - 1,$$

with $0 < h < 1/2$, so that these intervals are disjoint. In this case, Δ is contained in $(0, 2\pi)$ and has the Lebesgue measure

$$|\Delta| = \sum_{k=1}^{T_0-1} |\Delta_k| = 4\pi h\, \frac{T_0 - 1}{T}.$$

Moreover, let us require that $|\Delta| = 1$, i.e.,

$$h = \frac{1}{4\pi}\, \frac{T}{T_0 - 1}.$$

Since the last ratio is maximized for $T \uparrow 3$, we have

$$\frac{1}{4\pi} \leqslant h \leqslant \frac{3}{4\pi}. \tag{4.3}$$

Now, define μ to be the uniform distribution on Δ, so that $M(F) = 1$ and, by (4.2),

$$\delta_F(T) \geqslant \frac{2}{\pi} \int_{\Delta} r(T(z-y))\, dz. \tag{4.4}$$

It remains to properly estimate the above integrand. For this aim, let us integrate in (2.4) once more, which leads to

$$r(t) = \frac{\cos t}{t} + \frac{\sin t}{t^2} - 2\int_t^{\infty} \frac{\sin u}{u^3}\, du.$$

The last integral is smaller than $1/(2t^2)$, so,

$$r(t) \geqslant \frac{\cos t}{t} - \frac{2}{t^2} = \frac{1}{t}\left(\cos t - \frac{2}{t}\right), \quad t > 0. \tag{4.5}$$

Let $t = T(z-y)$ for $z \in \Delta_k$, $1 \leqslant k \leqslant T_0$. Then $t = 2\pi(k+\theta) + 2\pi m$ for some $\theta \in (-h,h)$, so that

$$\cos t = \cos(2\pi\theta) \geqslant \cos(2\pi h) \geqslant \cos(3/2) = 0.0707...,$$

where we made use of the upper bound in (4.3). On the other hand,

$$t \geqslant 2\pi(k-h) + 2\pi m \geqslant 2\pi m.$$

It follows that

$$\cos t - \frac{2}{t} \geqslant \cos(3/2) - \frac{1}{\pi m} > 0.01,$$

where, at the last step, we choose $m = 6$. Then $t \leqslant 2\pi(k+h) + 2\pi m < 2\pi(k+7)$ and, by (4.5),

$$r(t) \geqslant \frac{0.01}{2\pi(k+7)}, \quad t = T(z-y), \quad z \in \Delta_k.$$

Returning to (4.4), this gives with some absolute constant $c_0 > 0$

$$\delta_F(T) \geqslant \frac{2}{\pi} \sum_{k=1}^{T_0} \frac{0.01}{2\pi(k+7)}\, |\Delta_k| = \frac{0.04h}{\pi T} \sum_{k=1}^{T_0} \frac{1}{k+7} \geqslant c_0\, \frac{\log T}{T},$$

where we made use of the lower bound in (4.3). This proves (4.1).

5 Proof of Proposition 1.5

We apply smoothing inequalities due to Prawitz [4]: For an arbitrary distribution function F with characteristic function f and any point $x \in \mathbb{R}$

$$\frac{1}{2} - \int_{-T}^{T} e^{-itx}\, K_T(-t) f(t)\, dt \leqslant F(x) \leqslant \frac{1}{2} + \int_{-T}^{T} e^{-itx}\, K_T(t) f(t)\, dt. \tag{5.1}$$

Here, for a fixed value $T > 0$ the kernel is defined by

$$K_T(t) = \frac{1}{T}\, K\Big(\frac{t}{T}\Big),$$

where

$$K(t) = \frac{1}{2}\,(1-|t|) + \frac{i}{2}\, \Big[(1-|t|)\cot(\pi t) + \frac{\mathrm{sgn}(t)}{\pi}\Big], \quad |t| < 1.$$

The integrals in (5.1) are understood as principal values, i.e., as limits of the integrals over the regions $\varepsilon < |t| < T$ as $\varepsilon \downarrow 0$. It was also mentioned in [4] that

$$\Big|K(t) - \frac{i}{2\pi t}\Big|^2 = \frac{1}{4}\,(1-|t|)^2\, \Big[1 + \Big(\frac{1}{\pi t} - \cot(\pi t)\Big)\Big]^2,$$

which can be estimated by means of the elementary bound

$$\cot x \geqslant \frac{1}{x} - \frac{x}{3}\, \frac{\pi^2}{\pi^2 - x^2}, \quad 0 < x < \pi.$$

It is easy to see that this leads to

$$\Big|K(t) - \frac{i}{2\pi t}\Big| \leqslant \frac{1}{2}, \quad |t| \leqslant 1.$$

Applying this bound to (5.1), we arrive at the representation

$$F(x) = \frac{1}{2} + \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx}}{-it}\, f(t)\, dt + R \tag{5.2}$$

with the remainder term satisfying

$$|R| \leqslant \frac{1}{2T} \int_{-T}^{T} |f(t)|\, dt. \tag{5.3}$$

Thus, for all $x, y \in \mathbb{R}$

$$F(x) - F(y) = \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)\, dt + \frac{\theta}{T} \int_{-T}^{T} |f(t)|\, dt$$

with some complex number $\theta = \theta(x, y, T)$ such that $|\theta| \leqslant 1$. As a consequence, similarly to the Esseen bound with $h = 1/T$, we obtain the desired inequality (1.10).

If $f(t)$ is nonnegative, the normalized integral in (1.10) is equivalent to $\mathbf{P}\{|X| \leqslant 1/T\}$, assuming that the random variable X has the distribution function F (see, for example, [7, p. 27]). Therefore, in this case, (1.10) can be written up to some absolute constant c as (1.11).

Remark 5.1. With the factor $1/T$ in front of the integral in (5.3), the representation (5.2) appeared in [8, Lemma 4.1].

Let us explain why the inequality (1.11) improves upon (1.4). Consider the function

$$I(t) = \int_{0}^{t} |f(s)|\, ds, \quad t \geqslant 0,$$

assuming for a moment that $I(t) = o(t)$ as $t \to \infty$. Then, integrating by parts, we have

$$\int_{T}^{\infty} \frac{|f(t)|}{t}\, dt = \int_{T}^{\infty} \frac{1}{t}\, dI(t) = \frac{I(T)}{T} + \int_{T}^{\infty} \frac{I(t)}{t^2}\, dt \geqslant \frac{I(T)}{T}.$$

At this step, the assumption on the growth of $I(t)$ can be dropped. Hence (1.11) implies

$$\delta_F(T) \leqslant 2 \int_{T}^{\infty} \frac{|f(t)|}{t}\, dt,$$

i.e., (1.4) with an extra factor.

6 Squares of Bernoulli Sums

To illustrate Corollaries 1.2 and 1.4 by specific examples, let us fix an integer $d \geqslant 1$ and consider the normalized sums

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

of independent random vectors X_k uniformly distributed in the discrete cube $\{-1,1\}^d$. By the central limit theorem, the distributions of $Z_n^{(d)}$ are weakly convergent as $n \to \infty$ to the distribution of the random vector $Z^{(d)}$ in $\mathbb{R}^d$ with the standard normal law.

Let F_n^{*d} and F^{*d} denote respectively the distribution functions of the random variables

$$\xi_n^{(d)} = \frac{1}{2}\,|Z_n^{(d)}|^2, \quad \xi^{(d)} = \frac{1}{2}\,|Z^{(d)}|^2.$$

If $d=1$, we simplify the notation: $Z_n = Z_n^{(1)}$, $Z = Z^{(1)}$, $\xi_n = \xi_n^{(1)}$, $\xi = \xi^{(1)}$, and similarly for the distribution functions $F_n = F_n^{(1)}$, $F = F^{(1)}$.

Note that $2\xi_n$ is the square of the sum of n independent Bernoulli random variables taking the values ± 1 with probability $1/2$, and $\xi_n^{(d)}$ is the sum of d independent copies of ξ_n. Hence F_n^{*d} and F^{*d} represent the d-th convolution power of F_n and F, respectively.

First let us look at the one-dimensional case $d=1$. In terms of the distribution functions $\Phi_n(x) = \mathbf{P}\{Z_n \leqslant x\}$ and $\Phi(x) = \mathbf{P}\{Z \leqslant x\}$, we have

$$F_n(x) = \mathbf{P}\{|Z_n| \leqslant \sqrt{2x}\} = 2\,\Phi_n(\sqrt{2x}) - 1,$$
$$F(x) = \mathbf{P}\{|Z| \leqslant \sqrt{2x}\} = 2\,\Phi(\sqrt{2x}) - 1,$$

for all $x \geqslant 0$. It is well-known that, up to some absolute constant $c>0$,

$$|\Phi_n(x) - \Phi_n(y)| \leqslant c\,\Big(\frac{1}{\sqrt{n}} + |x-y|\Big), \quad x,y \in \mathbb{R},$$

and obviously $|\Phi(x) - \Phi(y)| \leqslant |x-y|$. Thus,

$$F_n(x) - F_n(y) \leqslant c\,\Big(\frac{1}{\sqrt{n}} + \sqrt{x} - \sqrt{y}\Big),$$

and $F(x) - F(y) \leqslant \sqrt{x} - \sqrt{y}$ for $x > y \geqslant 0$. Since $\sqrt{x} - \sqrt{y} \leqslant \sqrt{x-y}$, we are in position to apply Corollary 1.4 with $\alpha = 1/2$. Introduce the characteristic functions

$$f_n(t) = \mathbf{E}\,e^{itZ_n^2/2} = \int_{-\infty}^{\infty} e^{itx^2/2}\,d\Phi_n(x), \tag{6.1}$$

$$f(t) = \mathbf{E}e^{itZ^2/2} = \int_{-\infty}^{\infty} e^{itx^2/2}\,d\Phi(x) = \frac{1}{\sqrt{1-it}}, \quad t \in \mathbb{R}, \tag{6.2}$$

associated with the distribution functions F_n and F.

Corollary 6.1. *For all $x,y \in \mathbb{R}$ and $T>0$*

$$F_n(x) - F_n(y) = \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\,f_n(t)\,dt + \theta\,\Big(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{T}}\Big),$$

$$F(x) - F(y) = \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\,f(t)\,dt + \frac{\theta}{\sqrt{T}},$$

where θ is bounded in absolute value by an absolute constant.

Thus, if $T \geqslant n$, then $\delta_{F_n}(T) \leqslant c/\sqrt{n}$ and similarly for F. Note that $F_n(x)$ makes a jump of order $1/\sqrt{n}$ at $x = 0$ for large even values of n.

7 Approximation for Convolutions F_n^{*2}

If $d \geqslant 2$, the remainder term in the Fourier inversion formula is improved for the d-th convolution power F_n^{*d} of the distribution F_n with its characteristic function $f_n(t)^d$ (recall that $f_n(t)$ was defined in (6.1)). To see this, here we focus on the case $d = 2$. In what follows, we use the sequence

$$\varepsilon_N = \frac{\log\log\log N}{\log\log N}, \quad N \geqslant 3$$

(putting $\varepsilon_1 = \varepsilon_2 = 0$ for definiteness).

Corollary 7.1. *For all $x, y \in \mathbb{R}$ and $T \geqslant 2$*

$$F_n^{*2}(x) - F_n^{*2}(y) = \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f_n(t)^2\, dt + \theta n^{\varepsilon_n} \Big(\frac{1}{n} + \frac{\log T}{T}\Big),$$

where the quantity θ is bounded in absolute value by an absolute constant.

Note that the random variable $\xi^{(2)}$ has a standard exponential distribution with the distribution function

$$F^{*2}(x) = \mathbf{P}\{\xi^{(2)} \leqslant x\} = 1 - e^{-x} \quad (x \geqslant 0)$$

and characteristic function $f(t)^2 = \frac{1}{1-it}$, cf. (6.2). Therefore, by (1.4),

$$F^{*2}(x) - F^{*2}(y) = \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f(t)^2\, dt + \frac{\theta}{T}.$$

To prove the corollary, we need an upper bound for the number of representations of a natural number N as the sum of two squares of integers, which is commonly denoted as

$$r_2(N) = \mathrm{card}\,\{(k_1, k_2) : k_1^2 + k_2^2 = N, \quad k_1, k_2 \in \mathbb{Z}\}.$$

It is well known that $r_2(N) = o(N^{\varepsilon})$ for any $\varepsilon > 0$ as N tends to infinity. Let us give a more precise statement which seems to be also known, although we cannot give a precise reference.

Lemma 7.2. *For $\lambda > 1/2$ we have $r_2(N) \leqslant N^{\lambda \varepsilon_N}$ for all N large enough.*

Proof. One can employ the following representation [9]: If

$$N = 2^{\alpha} p_1^{\alpha_1} \dots p_r^{\alpha_r} q_1^{\beta_1} \dots q_s^{\beta_s}$$

is the decomposition of N into prime factors, where $p_i \equiv 1 \pmod 4$, $q_j \equiv 3 \pmod 4$, then

$$r_2(N) = \begin{cases} 4\,(\alpha_1 + 1) \dots (\alpha_r + 1), & \text{all } \beta_j \text{ are even} \\ 0, & \text{some of } \beta_j \text{ is odd.} \end{cases}$$

Therefore, starting from the prime factorization without the above specification

$$N = p_1^{\alpha_1} \dots p_r^{\alpha_r}, \quad 2 \leqslant p_1 < \dots < p_r, \tag{7.1}$$

we have

$$r_2(N) \leqslant 4\,(\alpha_1 + 1) \dots (\alpha_r + 1) \leqslant 2^{r+2}\,\alpha_1 \dots \alpha_r. \tag{7.2}$$

Necessarily, $N \geqslant p_1 \dots p_r \geqslant r!$ implying that for all N large enough

$$r \leqslant \lambda\,\frac{\log N}{\log\log N}. \tag{7.3}$$

Indeed, assume that the opposite inequality holds. Then for a given $\varepsilon > 0$ we would get

$$\log r - 1 > \log \lambda - 1 + \log\log N - \log\log\log N > (1-\varepsilon)\,\log\log N$$

for sufficiently large N. Using $r! \geqslant (r/e)^r \sqrt{r}$ and choosing $\varepsilon = (2\lambda-1)/(2\lambda+1)$, this would lead to

$$\log(r!) \geqslant r\,(\log r - 1) + \frac{1}{2}\log r > \lambda\,\frac{\log N}{\log\log N}\cdot(1-\varepsilon)\,\log\log N + \frac{1-\varepsilon}{2}\,\log\log N = \log N,$$

contradicting to $r! \leqslant N$. Thus, by (7.3) with $\lambda \leqslant 1/\log 2$,

$$2^r = e^{r\log 2} \leqslant \exp\Big\{\frac{\log N}{\log\log N}\Big\},$$

so that, by (7.2),

$$r_2(N) \leqslant 4\,\alpha_1 \dots \alpha_r\,\exp\Big\{\frac{\log N}{\log\log N}\Big\}. \tag{7.4}$$

Now, taking the logarithm in (7.4), let us maximize the concave function in r real variables

$$u(\alpha_1, \dots, \alpha_r) = \log\alpha_1 + \dots + \log\alpha_r, \quad \alpha_1, \dots, \alpha_r \geqslant 0,$$

subject to the linear condition $c_1\alpha_1 + \dots + c_r\alpha_r = c$ with $c_i = \log p_i$ and $c = \log N$, according to (7.1). Treating α_r as a function of the remaining variables and assuming that $r \geqslant 2$, we have

$$\frac{\partial u}{\partial \alpha_i} = \frac{1}{\alpha_i} - \frac{c_i}{c_r}\,\frac{1}{\alpha_r} = 0, \quad 1 \leqslant i \leqslant r-1,$$

which means that the point of maximum of u satisfies $c_i\alpha_i = b$ for all $i \leqslant r$. Since the sum of $c_i\alpha_i$ is c, we get $b = c/r$, $\alpha_i = c/(c_i r)$, so

$$\max u = \log(\alpha_1 \dots \alpha_r) = \log\frac{c^r}{r^r c_1 \dots c_r}.$$

This also holds for $r = 1$. Using $c_1 \dots c_r \geqslant \log 2$, we get

$$\alpha_1 \dots \alpha_r \leqslant \Big(\frac{\log N}{r\log 2}\Big)^r.$$

But the function $((\log N)/(x\log 2))^x$ is positive and increasing for $1 \leqslant x < (1/(e\log 2))\,\log N$. In view of (7.3), our values of r belong to this interval for all N large enough as long as $1/2 < \lambda < (1/(e\log 2)) \sim 0.53...$ which can be assumed. We then get

$$\Big(\frac{\log N}{r\log 2}\Big)^r \leqslant \Big(\frac{\lambda\log\log N}{\log 2}\Big)^{\frac{\lambda\log N}{\log\log N}} = \exp\Big\{\frac{\lambda\log N}{\log\log N}\,(\log\log\log N + \log\lambda - \log\log 2)\Big\}.$$

It remains to recall (7.4) and note that λ can be as close to $1/2$ as we wish. $\square$

Proof of Corollary 7.1. Recall that

$$\xi_n^{(2)} = \frac{1}{2} Z_n^2 + \frac{1}{2} Z_n'^2,$$

where Z_n' is an independent copy of Z_n. By the local limit theorem for the binomial distributions,

$$\mathbf{P}\Big\{Z_n = \frac{k}{\sqrt{n}}\Big\} \leqslant \frac{c}{\sqrt{n}}, \quad k \in \mathbb{Z}, \tag{7.5}$$

with some absolute constant $c > 0$. Since the random variable $\xi_n^{(2)}$ takes values of the form $N/(2n)$ with $N = k_1^2 + k_2^2$ $(k_1, k_2 \in \mathbb{Z})$, the inequality (7.5) yields

$$\mathbf{P}\Big\{\xi_n^{(2)} = \frac{N}{2n}\Big\} = \sum_{k_1^2+k_2^2=N} \mathbf{P}\Big\{Z_n = \frac{k_1}{\sqrt{n}}\Big\} \mathbf{P}\Big\{Z_n = \frac{k_2}{\sqrt{n}}\Big\} \leqslant \frac{c^2}{n}\, r_2(N).$$

Note that $|Z_n| \leqslant \sqrt{n}$. Hence we only need to consider the values $N \leqslant 2n^2$. In this case, since n^{ε_n} is increasing for large n, while $\varepsilon_{2n^2} \sim \varepsilon_n$, from Lemma 7.2 for all n large enough we have $r_2(N) \leqslant n^{\frac{3}{4}\varepsilon_{2n^2}} \leqslant n^{\varepsilon_n}$. Thus,

$$\mathbf{P}\Big\{\xi_n^{(2)} = \frac{N}{2n}\Big\} \leqslant \frac{c}{n}\, n^{\varepsilon_n}. \tag{7.6}$$

Now, suppose that $x > y \geqslant 0$ and $1/n \leqslant x - y \leqslant 1$. The interval $[y, x]$ contains at most $[2n(x-y)] + 1 \leqslant 3n(x-y)$ points of the form $N/(2n)$ with integers N. Hence

$$F_n^{*2}(x) - F_n^{*2}(y) = \sum_{y < \frac{N}{2n} \leqslant x} \mathbf{P}\Big\{\xi_n = \frac{N}{2n}\Big\} \leqslant 3cn^{\varepsilon_n}\,(x-y).$$

Combining this with (7.6), it follows that $F_n^{(2)}$ satisfies the quasi-Lipschitz condition

$$|F_n^{*2}(x) - F_n^{*2}(y)| \leqslant cn^{\varepsilon_n}\Big(\frac{1}{n} + |x-y|\Big) \tag{7.7}$$

for all $x, y \in \mathbb{R}$ up to some absolute constant $c > 0$. We are in position to apply Corollary 1.2 to F_n with $\varepsilon = 1/n$ and $M = cn^{2\varepsilon_n}$. $\square$

8 Approximation for Convolution Powers F_n^{*3}

As the last example, consider the distribution functions F_n^{*3} of the random variables

$$\xi_n^{(3)} = \frac{1}{2} Z_n^2 + \frac{1}{2} Z_n'^2 + \frac{1}{2} Z_n''^2,$$

where Z_n', Z_n'' are independent copies of Z_n. The next assertion is analogous to Corollary 7.1.

Corollary 8.1. *For all $x, y \in \mathbb{R}$ and $T \geqslant 2$*

$$F_n^{*3}(x) - F_n^{*3}(y) = \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{-it}\, f_n(t)^3\, dt + \theta\, n^{\varepsilon_n}\Big(\frac{1}{n} + \frac{\log T}{T}\Big),$$

where $f_n(t)$ is the characteristic function of $\frac{1}{2} Z_n^2$ and where θ is bounded in absolute value.

Proof. By Lemma 7.2, the set $\Omega = \{(k_1, k_2, k_3) : k_1^2 + k_2^2 + k_3^2 = N, k_j \in \mathbb{Z},\ |k_j| \leqslant n\}$ has cardinality

$$r_{3,n}(N) = \mathrm{card}\,(\Omega) \leqslant c\sqrt{n}\, N^{3\varepsilon_N/4} \tag{8.1}$$

(where $c > 0$ is an absolute value which can vary from place to place). Since $\xi_n^{(3)}$ takes the values $N/(2n)$, where $N = k_1^2 + k_2^2 + k_3^2$ with $k_j \in \mathbb{Z}$, $|k_j| \leqslant n$, we obtain, by (7.5),

$$\mathbf{P}\Big\{\xi_n^{(3)} = \frac{N}{2n}\Big\} \leqslant \sum_{(k_1,k_2,k_3)\in\Omega} \mathbf{P}\Big\{Z_n = \frac{k_1}{\sqrt{n}}\Big\}\,\mathbf{P}\Big\{Z_n = \frac{k_2}{\sqrt{n}}\Big\}\,\mathbf{P}\Big\{Z_n = \frac{k_3}{\sqrt{n}}\Big\} \leqslant \frac{c}{n^{3/2}}\, r_{3,n}(N).$$

Hence, by (8.1),

$$\mathbf{P}\Big\{\xi_n^{(3)} = \frac{N}{2n}\Big\} \leqslant \frac{c}{n}\, N^{3\varepsilon_N/4}.$$

Since $|Z_n| \leqslant \sqrt{n}$, it is necessary that $N \leqslant 3n^2$. As we noted above, n^{ε_n} is increasing for large n, while $\varepsilon_{3n^2} \sim \varepsilon_n$. Therefore, we arrive at the same bound as in dimension two,

$$\mathbf{P}\Big\{\xi_n^{(3)} = \frac{N}{2n}\Big\} \leqslant \frac{c}{n}\, n^{\varepsilon_n}.$$

With a similar argument, this implies that the distribution functions F_n of the random variables ξ_n satisfy the quasi-Lipschitz condition (7.7). One can therefore apply Corollary 1.2. □

Remark 8.1. For the convolutions F_n^{*k} with larger values of k (at least for $k > 4$) one can derive similar representations as in Corollaries 7.1 and 8.1 without the factor n^{ε_n}. In this case, the number $r_k(n)$ of representations of n as a sum of k squares of integers is approximately $n^{\frac{k}{2}-1}$ within k-dependent factors. There is an intensive literature on this topic (see, for example, [10]).

Funding

The work was partially supported by the NSF grant DMS-2154001.

Declarations

Data availability This manuscript has no associated data.

Ethical conduct Not applicable.

Conflicts of interest The author declares that there is no conflict of interest.

References

1. N. G. Ushakov, "Some inequalities for characteristic functions of unimodal distributions," *Theory Probab. Appl.* **26**, 393–398 (1982).

2. N. G. Ushakov, *Selected Topics in Characteristic Functions*, VSP, Utrecht (1999).

3. V. V. Petrov, *Sums of Independent Random Variables* Springer, New York etc. (1975).

4. H. Prawitz, "Limits for a distribution, if the characteristic function is given in a finite domain," *Skand. Actuarial J.* **55**, 138–154 (1972).

5. T. Kawata, *Fourier Analysis in Probability Theory*, Academic Press, New York etc. (1972).

6. H. Makabe, "A remark on the smoothness of the distribution function," *Yokohama Math. J.* **8**, 59–68 (1960).

7. V. V. Petrov, *Limit Theorems of Probability Theory. Sequences of Independent Random Variables*, Clarendon Press, New York (1995).

8. V. Bentkus and F. Götze, "Optimal rates of convergence in the CLT for quadratic forms," *Ann. Prob.* **24**, 466–490 (1996).

9. M. D. Hirschhorn, "Arithmetic consequences of Jacobi's two-squares theorem," *Ramanujan J.* **4**, No. 1, 51–57 (2000).

10. J. Rouse, "Explicit bounds for sums of squares," *Math. Res. Lett.* **19**, No. 2, 359–376 (2012).

Submitted on March 3, 2024