The Annals of Probability
1996, Vol. 24, No. 1, 466–490

OPTIMAL RATES OF CONVERGENCE IN THE CLT FOR QUADRATIC FORMS[1]

BY V. BENTKUS AND F. GÖTZE

University of Bielefeld

We prove optimal convergence rates in the central limit theorem for sums in $\mathbf{R}^k$. Assuming a fourth moment, we obtain a Berry–Esseen type bound of $O(N^{-1})$ for the probability of hitting a ball provided that $k \geq 5$. The proof still requires a technical assumption related to the independence of coordinates of sums.

1. Introduction and results. Let $\mathbf{R}^k$ denote the real k-dimensional Euclidean space with scalar product $(\cdot, \cdot)$ and norm $|\cdot|$, $1 \leq k < \infty$. For $k = \infty$ regard $\mathbf{R}^k$ as a real separable Hilbert space.

Let $X, X_1, X_2, \dots$ be a sequence of independent and identically distributed (i.i.d.) random variables taking values in $\mathbf{R}^k$. Write

$$S_N = (X_1 + \cdots + X_N)/\sqrt{N}.$$

Assume that $\mathbf{E}X = 0$ and that $\mathbf{E}|X|^2 < \infty$. Then X is pre-Gaussian; that is, there exists a centered Gaussian random variable G such that the covariances of X and G are equal.

Let us consider a quadratic form

$$Q(x) = (Qx, x) \quad \text{for } x \in \mathbf{R}^k,$$

where $Q\colon \mathbf{R}^k \to \mathbf{R}^k$ denotes a symmetric linear bounded operator.

Assuming that the fourth moment $\mathbf{E}|X|^4 < \infty$, we shall prove the following estimate of the convergence rate in the central limit theorem:

$$\sup_r \left|\mathbf{P}\{Q(S_N) < r\} - \mathbf{P}\{Q(G) < r\}\right| = O(N^{-1}). \tag{1.1}$$

It is important to emphasize that we do *not* require any smoothness conditions like Cramér's (C) condition for the characteristic function of X or its coordinates. Related lower bounds in the lattice point problem [see, for instance, Walfisz (1957) and Fricker (1982)] suggest that (1.1) may not be valid in dimensions smaller than 5. We shall prove that (1.1) holds for $k \geq 5$. For dimension $k = 4$ we get $O(N^{-1} \ln^\delta N)$ with some $\delta > 0$. The dimensions $k = 2, 3$ are not studied here since they require different methods. Unfortunately, in order to prove (1.1) we still need certain assumptions on the independence of coordinates of X. For example, the following conditions are

---

Received July 1994; revised January 1995.

[1]Research supported by the SFB 343 in Bielefeld.

AMS 1991 *subject classifications.* Primary 62E20; secondary 60F05.

Key words and phrases. Key words and phrases. Central limit theorem, convergence rates, Berry–Esseen bound, multidimensional spaces, Hilbert space, quadratic forms, ellipsoids.

sufficient: the first five coordinates of X have positive variances, are mutually independent and do not depend on other coordinates; the operator Q is diagonal and its first five eigenvalues are nonzero.

Let us formulate the result more precisely. Without loss of generality we shall assume that X is not concentrated in a proper subspace of $\mathbf{R}^k$. That is, $\ker B = \{0\}$, where the covariance operator B^2 of X is defined by

$$(B^2x, y) = \mathbf{E}(X, x)(X, y) = \mathbf{E}(G, x)(G, y) \quad \text{for } x, y \in \mathbf{R}^k.$$

In the finite-dimensional case $\ker B = \{0\}$ implies that B has rank k. In the case $k = \infty$ the operator B^2 is nuclear and $\ker B = \{0\}$ implies that all eigenvalues of B^2 are positive.

Without loss of generality we may assume that $\ker Q = \{0\}$, which means that $\mathbf{R}^k = \overline{\operatorname{Im} Q}$ (the closure of the image of Q) since Q is symmetric. Otherwise we may replace X by its projection onto $\overline{\operatorname{Im} Q}$ without changing the distribution of (QX, X) and without violating other conditions on X.

A consequence of these assumptions is that $\ker BQB = \{0\}$.

The independence condition. We shall assume that the random variable X is a sum of two independent random variables,

$$X = U + V, \quad \text{where } U \text{ and } V \text{ are independent.} \tag{1.2}$$

Furthermore, we require that $\mathbf{E}U = 0$, that U assumes values in an invariant subspace of Q, say $\mathbf{R}_X$, with dimension at least 5, that Q is diagonal in some orthonormal basis of $\mathbf{R}_X$ and that the coordinates of U in this basis are independent and have positive variances. Without loss of generality we may assume that the dimension of $\mathbf{R}_X$ is 5.

The independence condition means that there exists a basis of $\mathbf{R}^k$ such that U may be written as $U = (U_1, \dots, U_5, 0, 0, \dots)$, where the coordinates $U_1, \dots, U_5$ of U are centered independent random variables which have positive variances and such that

$$(QU, U) = \sum_{j=1}^{5} \lambda_j U_j^2 \quad \text{with some } \lambda_j \neq 0. \tag{1.3}$$

Note that (1.2) may hold in cases where coordinates of X are not independent. Note as well that $\mathbf{E}|X|^4 < \infty$ implies $\mathbf{E}|U|^4 < \infty$ and $\mathbf{E}|V|^4 < \infty$. Furthermore, it follows that k is at least 5.

Set $\alpha_k = 6$, for $k = \infty$, and $\alpha_k = 3$, otherwise.

THEOREM 1.1. *Assume that* (1.2) *holds and that* $\mathbf{E}|X|^4 < \infty$. *Let* $a \in \mathbf{R}^k$. *Then*

$$\begin{aligned} &\sup_r \left| \mathbf{P}\{Q(S_N - a) < r\} - \mathbf{P}\{Q(G - a) < r\} - \frac{\kappa(r)}{6\sqrt{N}} \right| \\ &\qquad = O\left(\frac{1 + |a|^{\alpha_\kappa}}{N} \right), \end{aligned} \tag{1.4}$$

where the coefficient κ in the Edgeworth correction $\kappa(r)/(6\sqrt{N})$ depends only on Q, B, r, a and on the third moment of X.

For the precise definition of $\kappa(r)$, see (2.4) in the case $k < \infty$ and Lemma 2.2 when $k = \infty$. Furthermore, note that $\kappa = 0$ whenever $a = 0$ or X is symmetric.

REMARK. A rather straightforward inspection of proofs shows that condition (1.2) with four independent coordinates instead of five implies (1.4) with $O(N^{-1} \ln^{\delta} N)$ instead of $O(N^{-1})$, for some $\delta > 0$.

EXAMPLE [Bentkus and Zalesskii (1985)]. Let Q be the identity operator $Q(x) = |x|^2$. For any $k \geq 1$ there exists a bounded random variable X with independent coordinates satisfying conditions of Theorem 1.1 such that the distribution function $\mathbf{P}\{Q(S_N) < r\}$ has a jump of $O(1/N)$, for all N.

Thus the bound $O(1/N)$ in (1.1) and (1.4) is best possible since the distribution function $\mathbf{P}\{Q(G) < r\}$ is continuous.

Construction of the example. Choose any bounded centered random variable X such that $\mathbf{P}\{X \in \mathbb{Z}^k\} = 1$. The central limit theorem implies that $\mathbf{P}\{Q(S_N) < r_0\} \geq c > 0$, for some $r_0 \geq 1$ and c, and for all $N \geq 1$. However, the random variable $NQ(S_N) = N|S_N|^2$ assumes only integer values $0, 1, 2, \ldots$ since $\mathbf{P}\{X \in \mathbb{Z}^k\} = 1$. The interval $[0, Nr_0)$ contains at most $2Nr_0$ integers. Thus $NQ(S_N)$ is equal to at least one of these integers with a probability not less than $c/(2Nr_0)$ since $\mathbf{P}\{Q(S_N) < r_p\} \geq c$.

Bounds for $\sup_r |\mathbf{P}\{Q(S_N) < r\} - \mathbf{P}\{Q(G) < r\}|$ were studied by a number of authors. For a discussion of the related literature, see Bentkus, Götze, Pavlauskas and Račkauskas (1991). Assuming a fourth moment, the best results obtained were of the type $O(N^{-1+\varepsilon_k})$, where $0 < \varepsilon_k \leq 1/2$, and $\varepsilon_k \searrow 0$ as $k \nearrow \infty$. Theorem 1.1 improves these results, and the improvement is final in the sense that the distribution function $\mathbf{P}\{Q(S_N) < r\}$ may have jumps of magnitude $O(N^{-1})$ (cf. the Example below).

Theorem 1.1 is related to certain results in number theory. Let $\mathbb{Z}^k$ denote the standard lattice of integer points in $\mathbf{R}^k$. Assume that Q is a positive definite operator and let $k < \infty$. Consider the ellipsoid $E_r = \{x\colon (Qx, x) \leq r\}$ and let $V_Q(r)$ [resp. $A_Q(r)$] denote the volume of E_r (resp. the number of lattice points within E_r). A well-known result in the theory of lattice points [see Landau (1915), Walfisz (1927), Jarnik (1928), as well as Landau (1924), Landau (1962), Kendall (1948) and Krätzel (1988)] states that

$$(1.5) \qquad r^{-k/2}\big|A_Q(r) - V_Q(r)\big| = O(r^{-1}) \quad \text{whenever } k \geq 5,\ r \geq 1,$$

for ellipsoids with principal axes parallel to the coordinate directions. The estimate (1.4) is similar to (1.5). The independence assumption (1.2) thereby corresponds to the assumption that principal axes are parallel to the coordinate system of $\mathbf{R}^k$.

Estimates like (1.5) for ellipsoids with general orientation have, to the best of our knowledge, not been proved yet. There is, however, an estimate $O(r^{-k/(k+1)})$ due to Landau (1915), which corresponds to a seminal result of Esseen (1945), who proved that the convergence rate in the central limit theorem is $O(N^{-k/(k+1)})$, for balls with center at the origin and for summands having identity covariance matrix. Without the independence condition, Theorem 1.1 would correspond to an estimate (1.5) for ellipsoids in general position.

To prove (1.4) we apply a variety of techniques. We use a symmetrization inequality [Götze (1979)] analogous to the Weyl–van der Corput inequality in analytic number theory [Weyl (1915/1916), and Graham and Kolesnik (1991)] and related techniques. These techniques turned out to be very useful for the investigation of the convergence rates and Edgeworth expansions in the central limit theorem in Hilbert and Banach spaces [see papers of the authors cited in Bentkus, Götze, Palauskas and Račkauskas (1991) as well as for the investigation of the asymptotic properties of certain statistics [see, e.g., Bentkus, Götze, and Zitikis (1993)].

Similar methods lead to the extension [Götze, Prohorov and Ulyanov (1994)] of Vinogradov (1934) bounds on trigonometric sums for characteristic functions of polynomials of sums of independent random variables. Related probabilistic problems connected to number theory are considered by Sinai (1991), Major (1992) and Blecher and Lebowitz (1993).

We use a new smoothing inequality [going back to Beurling, see Graham and Kolesnik (1991)] proved by Prawitz (1972) since our proof requires the inversion of the Fourier transform; that is, at a certain stage of the proof we pass to distributions again.

The classical techniques to prove convergence rates in finite-dimensional spaces [see, e.g., Bhattacharya and Rao (1986) and Sazonov (1981)] are useful to estimate characteristic functions for frequencies of magnitude $O(\sqrt{N/\ln N})$. For higher frequences of magnitude up to $O(N)$ we apply a reduction to sums of symmetric Bernoulli random variables, that is, to sums taking values in a lattice. For the estimation of integrals of the characteristic functions of those sums we use some ideas related to the circle method of Hardy and Littlewood (1920) in analytic number theory, and as a core result we prove the estimate (see Lemma 3.5)

$$\int_{\sqrt{N/\ln N}}^{N/4} \Bigl(\mathbf{E}\exp\bigl\{itB_{2N}\overline{B}_{2N}\bigr\}\Bigr)^{k/2}\frac{dt}{t} = O\left(\frac{1}{N}\right) \quad \text{for } k \geq 5,$$

where $B_N = (\varepsilon_1 + \cdots + \varepsilon_N)/\sqrt{N}$, the random variable $\overline{B}_N$ denotes an independent copy of B_N and $\varepsilon_1, \varepsilon_2, \dots$ is a sequence of i.i.d. symmetric Bernoulli random variables such that $\mathbf{P}\{|\varepsilon_1| = 1\} = 1$.

2. Proof of Theorem 1.1. By c (resp. C) with (or without) indices we shall denote sufficiently small (resp. sufficiently large) generic absolute constants. The dependence of constants on parameters, say α, we shall indicate by writing $c = c(\alpha)$. By μ we shall denote the distribution of the random variable X. Throughout we shall assume that all random variables are independent in aggregate, if the contrary is not obvious from the context. For a random variable X let $\overline{X}$ denote an independent copy and let $\tilde{X}$ denote a symmetrization, say $\tilde{X} = X - \overline{X}$.

Proof of Theorem 1.1. The proof of Theorem 1.1 is rather complicated and splits into a series of lemmas. We shall apply a smoothing lemma, which reduces the problem to the estimation of the characteristic function of the quadratic form (i.e., we shall apply a kind of one-dimensional Fourier transform). For the estimation of characteristic functions we shall use symmetrization inequalities, conditioning arguments and a reduction to the binomial case, which we call "Bernoullization" for short. The arguments depend on the dimension, and the case of lower dimensions $5 \le k \le 13$ is technically more involved. For lower dimensions we use the inverse Fourier transforms and additionally apply multidimensional Fourier transforms. We can invert the Fourier transform because we use the smoothing Lemma 4.1 instead of conventional smoothing inequalities.

Denote

$$f(t) = \mathbf{E}\exp\{itQ(S_N - a)\}, \qquad g(t) = \mathbf{E}\exp\{itQ(G - a)\}.$$

We denote the principal value of an integral by P.V. (see Section 4 for the definition).

The following lemma provides a reduction of Theorem 1.1 to the investigation of characteristic functions for lower order Fourier frequencies of magnitude up to $O(N^{2/5})$.

Lemma 2.1. *Assume the independence condition* (1.2). *Then there exists a constant* $C = C(Q, \mu)$ *such that*

$$\mathbf{P}\{Q(S_N - a) \le r\} = \frac{1}{2} + \frac{i}{2\pi}\mathrm{P.V.}\int_{|t| \le CN^{2/5}} \exp\{-irt\} f(t)\frac{dt}{t} + R,$$

where the remainder term R *satisfies* $\sup_{r,a}|R| = O(N^{-1})$.

Proof. Let us apply the smoothing lemma, Lemma 4.1. For any $H > 0$, we have

$$\mathbf{P}\{Q(S_N - a) \le r\} = \frac{1}{2} + \frac{i}{2\pi}\mathrm{P.V.}\int_{-H}^{H} \exp\{-irt\} f(t)\frac{dt}{t} + R_1,$$

where

$$|R_1| \le \frac{1}{H}\int_{-H}^{H} |f(t)|\, dt.$$

We choose $H = c_1 N$ with a sufficiently small constant $c_1 = c_1(Q, \mu)$. Lemmas 3.2–3.5 together imply that $\sup_{r,a}|R_1| = O(N^{-1})$. In order to apply Lemma 3.5, we have to change the variable of integration as $t = N\tau$.

Similarly Lemmas 3.2–3.5 imply that, for a sufficiently large $C = C(Q, \mu)$,

$$\sup_{r, a} \int_{CN^{2/5} \le |t| \le H} |f(t)| \frac{dt}{|t|} = O\left(\frac{1}{N}\right),$$

which concludes the proof of the lemma. □

Let us continue the proof of Theorem 1.1. From here on the proof will depend on the dimension k. The following lemma completes the proof in the case $k = \infty$.

LEMMA 2.2. *Let $k \ge 13$ and $\mathbf{E}|X|^4 < \infty$. Then the function*

$$g_1(t) = 4(it)^2 \mathbf{E}\big(3(G - a, QX)(QX, X) + 2it(G - a, QX)^3\big)\exp\{itQ(G - a)\}$$

is the characteristic function of a function of bounded variation, say κ, and the Edgeworth correction

$$\frac{\kappa(r)}{6\sqrt{N}} = \frac{i}{12\pi\sqrt{N}} \int_{\mathbf{R}} \exp(-irt) g_1(t) \frac{dt}{t}.$$

Furthermore,

$$\sup_r \left| \mathbf{P}\{Q(S_N - a) < r\} - \mathbf{P}\{Q(G - a) < r\} - \frac{\kappa(r)}{6\sqrt{N}} \right| = O\left(\frac{1 + |a|^6}{N}\right).$$

PROOF. We shall apply well-known techniques which have been developed for the estimation of the convergence rate in the CLT in infinite-dimensional spaces. We shall restrict ourselves to the case $k = \infty$. The case $k \ge 13$ requires only a more careful and tedious analysis. The dimension 13 is required to ensure the convergence of some integrals. Indeed, the estimates of the characteristic function show that each dimension contributes a factor $1/\sqrt{|t|}$. The asymptotic expansion for the characteristic function yields a factor $|t|^6$, such that $k = 13$ just leads to a converging integral $\int_{|t| \ge 1} dt/|t|^{3/2}$. It seems that the dimension $k = 12$ would suffice for the proof and would imply uniform estimates, just as dimension $k = 6$ yields the uniform estimate $O(N^{-1/2})$ [see Sazonov, Ulyanov and Zalesskii (1989), Nagaev (1988), and Senatov (1989)].

Thus let us assume that $k = \infty$. Then BQB has infinitely many nonzero eigenvalues. Therefore, the functions $|g(t)|$ and $|g_1(t)|$ are bounded from above by $C_s(Q, \mu)|t|^{-s}$, for all $s \ge 0$ and all $t \in \mathbf{R}$ [see Theorem 4.6 and Lemma 2.4 in Bentkus (1984a) and (1984b)]. Hence the smoothing lemma, Lemma 4.1, yields

$$\begin{aligned} &\mathbf{P}\{Q(G - a) \le r\} - \frac{1}{2} \\ &\quad - \frac{i}{2\pi} \text{P.V.} \int_{|t| \le CN^{2/5}} \exp\{-irt\} g(t) \frac{dt}{t} = O\left(\frac{1}{N}\right) \end{aligned} \tag{2.1}$$

uniformly in a and r. It is easy to verify that

$$\begin{aligned}(2.2)\qquad &\frac{i}{2\pi}\int_{\mathbf{R}}\exp\{-irt\}g_1(t)\frac{dt}{t}\\ &=\frac{i}{2\pi}\int_{|t|\le CN^{2/5}}\exp\{-irt\}g_1(t)\frac{dt}{t}+O\left(\frac{1}{N}\right)\end{aligned}$$

uniformly in a and r. Therefore, Lemma 2.1, (2.1) and (2.2) reduce the proof to the estimate

$$(2.3)\qquad \sup_r\int_{|t|\le CN^{2/5}}\left|f(t)-g(t)-\frac{g_1(t)}{6\sqrt{N}}\right|\frac{dt}{|t|}=O\left(\frac{1+|a|^6}{N}\right).$$

The random variables in the characteristic functions in (2.3) may be truncated as in Bentkus (1985). Then, applying Theorem 4.6 and Lemma 2.7 of Bentkus (1984a) and (1984b), we obtain the estimate of the lemma. □

The case $5\le k<\infty$ is technically more complicated. In the first step in Lemmas 2.3 and 2.4 we shall replace the sum S_N by a sum having $m\approx N^{3/10}$ Gaussian summands.

LEMMA 2.3. *Let $k\ge 5$ and let m denote the integer nearest to $N^{3/10}$. Assume that $\mathbf{E}|X|^3<\infty$ and denote $\delta=\sqrt{m/N}$. Then*

$$\mathbf{P}\{Q(S_N-a)\le r\}=\frac{1}{2}+\frac{i}{2\pi}\text{P.V.}\int_{|t|\le CN^{2/5}}\exp\{-irt\}h(t)\frac{dt}{t}+R,$$

where

$$h(t)=\mathbf{E}\exp\left\{itQ\left(\delta G-a+(X_1+\cdots+X_{N-m})/\sqrt{N}\right)\right\}$$

and

$$\sup_r|R|\le C(Q,\mu)(1+|a|^3)/N.$$

PROOF. We shall derive the result from Lemma 2.1. Let $G_1,G_2,\dots$ denote a sequence of independent copies of G. Denote

$$f_j(t)=\mathbf{E}\exp\left\{itQ\left(W_j-a+X/\sqrt{N}\right)\right\},$$

$$g_j(t)=f_{j+1}(t)=\mathbf{E}\exp\left\{itQ\left(W_j-a+G/\sqrt{N}\right)\right\},$$

where

$$W_j=\left(G_1+\cdots+G_{j-1}+X_{j+1}+\cdots+X_N\right)/\sqrt{N},\qquad 1\le j\le m.$$

Then $f=f_1$ and $f_{m+1}=h$. We have

$$f=R_1+\cdots+R_m+f_{m+1},\quad \text{where } R_s=f_s-f_{s+1}=f_s-g_s.$$

Expanding in Taylor series, conditioning and applying the symmetrization Lemma 3.1 we get

$$|R_s|\le CN^{-3/2}t^2(1+|t|)(1+|a|^3)\sqrt{\mathbf{E}\exp\left\{it\left(Q\tilde{Z}_n,\bar{\tilde{Z}}_N\right)\right\}},$$

where $Z_n = \sum_{1 \le j \le c_2 N} X_j / \sqrt{N}$ with some small but positive absolute constant $c_2 > 0$. The proof of this inequality follows from well-known standard arguments; see, for example, Bentkus [(1984a), Section 4]. Note that by careful Taylor expansion we may avoid the use of truncation techniques.

Arguing as in Lemmas 3.2–3.4, we get

$$\int_{1 \le |t| \le CN^{2/5}} |t|^3 \sqrt{\mathbf{E} \exp\{it(QZ_N, Z_N)\}} \frac{dt}{|t|} \le C \int_1^{CN^{2/5}} \frac{dt}{\sqrt{t}} \le CN^{1/5}.$$

These estimates together with $m \approx N^{3/10}$ conclude the proof of the lemma. □

LEMMA 2.4. *Let $k \ge 5$ and let m denotes the integer nearest to $N^{3/10}$. Denote $\delta = \sqrt{m/N}$. Then*

$$\sup_r \left| \mathbf{P}\{Q(S_N - a) \le r\} - \mathbf{P}\left\{Q\left(\delta G - a + (X_1 + \cdots + X_{N-m})/\sqrt{N}\right) \le r\right\} \right|$$
$$\le C(Q, \mu)(1 + |a|^3)/N.$$

PROOF. Due to Lemma 2.3 and to the Fourier inversion formula (4.3), it is sufficient to show that each of the integrals

$$J_1 = \int_{CN^{2/5} \le |t| \le cN} \frac{|h(t)|\, dt}{|t|},$$
$$J_2 = \int_{cN \le |t| \le cN^{11/10}} \frac{|h(t)|\, dt}{|t|},$$
$$J_3 = \int_{|t| \ge cN^{11/10}} \frac{|h(t)|\, dt}{|t|}$$

is $O(N^{-1})$ uniformly in a and r.

The Gaussian component δG ensures the desired estimate of J_3. Indeed, applying the symmetrization Lemma 3.1 we get

$$|h(t)| \le \sqrt{\mathbf{E} \exp\{ict(QG_1, G_2)\}} \le C(Q, \mu)\left(|t|\delta^2\right)^{-5/2}$$
$$= C(Q, \mu)\left(N/(|t|m)\right)^{5/2}$$

for dimensions $k \ge 5$ (here G_1 and G_2 denote independent copies of G). Integrating, we arrive at $J_3 = O(N^{-1})$.

The characteristic function $h(t)$ is the characteristic function of a quadratic form in a sum with at least $N/2$ summands from the sequence $X_1, \ldots, X_N$. Hence, arguing as in the proof of Lemma 2.1, we obtain $J_1 = O(N^{-1})$.

It remains to estimate J_2. Let us apply the symmetrization Lemma 3.1. Writing $Z_N = (X_1 + \cdots + X_{N/2})/\sqrt{N}$ we have

$$|h(t)|^2 \le \mathbf{E} \exp\left\{ict\delta\left(QG, \tilde{Z}_N\right)\right\} = \mathbf{E} \exp\left\{-ct^2\delta^2\left(BQ\tilde{Z}_N, BQ\tilde{Z}_N\right)\right\}$$
$$\le \mathbf{E} \exp\left\{-ct^2\delta^2 N^{-1/2}\left(BQ\tilde{Z}_N, BQ\tilde{Z}_N\right)\right\},$$

since $1 \ge N^{-1/2}$. Integration and the change of variable $t = c\tau N^{3/5}$ yields

$$J_2 \le \int_{cN^{2/5} \le \tau \le c\sqrt{N}} \sqrt{\mathbf{E} \exp\left\{ -c\tau^2 \left(BQ\tilde{Z}_N, BQ\tilde{Z}_N \right) \right\}} \, \frac{d\tau}{\tau}.$$

However,

$$\mathbf{E} \exp\left\{ -c\tau^2 \left(BQ\tilde{Z}_N, BQ\tilde{Z}_N \right) \right\} = \mathbf{E} \exp\left\{ -ic\tau \left(QG, \tilde{Z}_N \right) \right\}$$

and we may estimate this integral as in the proof of Lemma 2.1, which concludes the proof of the lemma. □

Let

$$p(x) = (2\pi)^{-k/2} \det B^{-1} \exp\{ -(B^{-2}x, x)/2 \}, \qquad x \in \mathbf{R}^k,$$

denote the density of G in $\mathbf{R}^k$. Define

$$\kappa(r) = -\int_{(Q(x-a),\, x-a) \le r} \mathbf{E} D_X^3 p(x)\, dx, \tag{2.4}$$

and for measurable $A \subset \mathbf{R}^k$,

$$\kappa(A) = -\int_A \mathbf{E} D_X^3 p(x)\, dx,$$

where

$$\mathbf{E} D_X^3 p(x) = -p(x) \left(\mathbf{E}(B^{-2}X, x)^3 - 3\mathbf{E}(B^{-2}X, x)(B^{-2}X, X) \right).$$

The proof of Theorem 1.1 is completed by the following lemma combined with Lemma 2.4.

LEMMA 2.5. *Let $k < \infty$ and $\mathbf{E}|X|^4 < \infty$. If $\delta^2 \approx m/N$ with $m = N^{3/10}$, then*

$$\sup_A \left| \mathbf{P}\left\{ \delta G + (X_1 + \cdots + X_{N-m})\sqrt{N} \in A \right\} - \mathbf{P}\{G \in A\} - \frac{\kappa(A)}{6\sqrt{N}} \right| = O\left(\frac{1}{N} \right). \tag{2.5}$$

PROOF. We shall proceed similarly as in Bhattacharya and Rao (1986), where Edgeworth expansions in the multidimensional case were systematically developed. Unfortunately we cannot use the corresponding results of this monograph directly.

The left-hand side of (2.5) is invariant under nondegenerate linear transformations. Therefore, without loss of generality, we may assume that covariances of X and G are identity matrices. Thus $\mathbf{E}(X, x)^2 = |x|^2$,

$$p(x) = (2\pi)^{-k/2} \exp\{ -|x|^2/2 \}$$

and

$$\mathbf{E} D_X^3 p(x) = -p(x) \left(\mathbf{E}(X, x)^3 - 3\mathbf{E}(X, x)(X, X) \right).$$

Using truncation [see Bhattacharya and Rao (1986), Section 14], we may replace $X_1 + \cdots + X_{N-m}$ by $Y_1 + \cdots + Y_{N-m}$, where $Y_j = X_j \mathbf{I}\{|X_j| \le \sqrt{N}\}$. Define the (signed) measure

$$D_N(A) = \mathbf{P}\{\delta G + Z_N \in A\} - \mathbf{P}\{G \in A\} - \frac{\kappa(A)}{6\sqrt{N}}, \qquad A \subset \mathbf{R}^k,$$

where $Z_N = (Y_1 + \cdots + Y_{N-m})/\sqrt{N}$. Note that $\mathbf{E}|Z_N|^p \le C(p)$, for all $p \ge 0$.

Thus we have to prove that

$$\sup_A |D_N(A)| = O(N^{-1}). \tag{2.6}$$

The measure D_N has a density, say $d_N(x)$, $x \in \mathbf{R}^k$, which is an element of the Schwartz space $\mathbf{S}$ of infinitely differentiable functions decreasing faster than any polynomial at infinity. Let $\mathbf{I}_A(x) = \mathbf{I}\{x \in A\}$ denote the indicator function of A. We may write

$$D_N(A) = \int_{\mathbf{R}^k} \mathbf{I}_A(x) d_N(x)\, dx. \tag{2.7}$$

The function $(1 + |x|^2)^{-k} \mathbf{I}_A(x)$ is integrable and the integral is bounded from above by a constant independent of A. In (2.7) we may apply Parseval's theorem and get

$$|D_N(A)| \le C(k) \int_{\mathbf{R}^k} \left| (1 + \Delta)^k \left(f(t) - g(t) - \frac{g_1(t)}{6\sqrt{N}} \right) \right| dt, \tag{2.8}$$

where

$$t = (t_1, \ldots, t_k), \qquad \partial_j = \frac{\partial}{\partial t_j}, \qquad \Delta = \partial_1^2 + \cdots + \partial_k^2$$

is the Laplace operator and the other functions are defined as

$$\begin{aligned} f(t) &= \mathbf{E} \exp\{i(t, \delta G + Z_N)\} = \exp\{-\delta^2 |t|^2/2\} \mathbf{E} \exp\{i(t, Z_N)\}, \\ g(t) &= \mathbf{E} \exp\{i(t, G)\} = \exp\{-|t|^2/2\}, \\ g_1(t) &= \mathbf{E}(it, X)^3 \exp\{i(t, G)\} = \exp\{-|t|^2/2\} \mathbf{E}(it, X)^3. \end{aligned}$$

Due to (2.8), the estimate (2.6) follows from

$$\int_{\mathbf{R}^k} \left| \partial_1^{s_1} \cdots \partial_k^{s_k} \left(f(t) - g(t) - \frac{g_1(t)}{6\sqrt{N}} \right) \right| dt = O\left(\frac{1}{N} \right) \tag{2.9}$$

provided that $s_1 + \cdots + s_k \le 2k$.

The function $f(t)$ is a product of $\exp\{-\delta^2 |t|^2\}$ and a differentiable function with bounded derivatives. Therefore,

$$\int_{|t| \ge C\delta^{-1} \ln N} \left| \partial_1^{s_1} \cdots \partial_k^{s_k} \left(f(t) - g(t) - \frac{g_1(t)}{6\sqrt{N}} \right) \right| dt = O\left(\frac{1}{N} \right)$$

provided the constant $C = C(\mu, k)$ is sufficiently large. Using $C\delta^{-1} \ln N \le N^{2/5}$, for sufficiently large N, we see that (2.9) is a consequence of

$$\int_{|t| \le N^{2/5}} \left| \partial_1^{s_1} \cdots \partial_k^{s_k} \left(f(t) - g(t) - \frac{g_1(t)}{6\sqrt{N}} \right) \right| dt = O\left(\frac{1}{N}\right) \tag{2.10}$$

provided that $s_1 + \cdots + s_k \le 2k$.

The relation (2.10) in turn follows from the estimate

$$\left| \partial_1^{s_1} \cdots \partial_k^{s_k} \left(f(t) - g(t) - \frac{g_1(t)}{6\sqrt{N}} \right) \right| \le CN^{-1} \exp\{-c|t|^2\}, \tag{2.11}$$

which holds for $|t| \le c\sqrt{N}$ with some sufficiently large constant $C = C(\mu, k)$ [resp. sufficiently small $c = c(\mu, k)$]. The proof of (2.11) is similar to the proofs in Chapter 2 in the book of Bhattacharya and Rao (1986). This completes the proof of the lemma. □

3. Estimates of the characteristic functions. Recall that for a random variable X, $\bar{X}$ denotes an independent copy and $\tilde{X}$ denotes a symmetrization, say $\tilde{X} = X - \bar{X}$.

By $\mathbf{E}_{X,Y,\dots}$ we shall denote the conditional expectation given all random variables but $X, Y, \dots$. For example, $\mathbf{E}_X u(X,Z) = \mathbf{E}(u(X,Z)|Z)$.

By $\varepsilon_1, \varepsilon_2, \dots$ we shall denote a sequence of i.i.d. symmetric Bernoulli random variables such that $\mathbf{P}\{|\varepsilon_1| = 1\} = 1$. Finally write

$$B_N = \varepsilon_1 + \cdots + \varepsilon_N.$$

The following symmetrization inequality improves slightly the well-known inequality due to Götze (1979). Similar improvements of the inequality were used by Nagaev (1988) and Sazonov, Ulyanov and Zalesskii (1989).

LEMMA 3.1. *Let $L \in \mathbf{R}^k$ and $C \in \mathbf{R}$. Let Y_1, Y_2, Y_3, W denote independent random variables taking values in $\mathbf{R}^k$. Denote by*

$$P(x) = (Qx, x) + (L, x) + C \quad \text{for } x \in \mathbf{R}^k,$$

a real-valued polynomial of second order. Then

$$2|\mathbf{E}\exp\{itP(Y_1 + Y_2 + Y_3 + W)\}|^2 \le \mathbf{E}\exp\left\{2it\left(Q\tilde{Y}_1, \tilde{Y}_2\right)\right\} + \mathbf{E}\exp\left\{2it\left(Q\tilde{Y}_1, \tilde{Y}_3\right)\right\}.$$

PROOF. Write

$$A = (QY_1, Y_1) - \left(Q\bar{Y}_1, \bar{Y}_1\right) \quad \text{and} \quad B = (L, Y_1) - \left(L, \bar{Y}_1\right).$$

We obtain by conditioning arguments

$$
\begin{aligned}
&\left|\mathbf{E}\exp\{itP(Y_1 + Y_2 + Y_3 + W)\}\right|^2 \\
&\quad \le \mathbf{E}\left|\mathbf{E}_{Y_1}\exp\{it(QY_1, Y_1) + 2it(QY_1, Y_2 + Y_3 + W) + it(L, Y_1)\}\right|^2 \\
&\quad = \mathbf{E}\exp\left\{itA + itB + 2it(QY_1, Y_2 + Y_3 + W) - 2it\left(Q\overline{Y}_1, Y_2 + Y_3 + W\right)\right\} \\
&\quad \le \mathbf{E}\left|\mathbf{E}_{Y_2, Y_3}\exp\left\{2it\left(Q\tilde{Y}_1, Y_2 + Y_3\right)\right\}\right| \\
&\quad = \mathbf{E}\left|\mathbf{E}_{Y_2}\exp\left\{2it\left(Q\tilde{Y}_1, Y_2\right)\right\}\right|\left|\mathbf{E}_{Y_3}\exp\left\{2it\left(Q\tilde{Y}_1, Y_3\right)\right\}\right| \\
&\quad \le \tfrac{1}{2}\mathbf{E}\exp\left\{2it\left(Q\tilde{Y}_1, \tilde{Y}_2\right)\right\} + \tfrac{1}{2}\mathbf{E}\exp\left\{2it\left(Q\tilde{Y}_1, \tilde{Y}_3\right)\right\},
\end{aligned}
$$

where the product of characteristic functions has been bounded using the elementary inequality $2ab \le a^2 + b^2$. □

The next lemma reduces the estimation of the characteristic function $|f(t)|$ to the estimation of characteristic functions with independent coordinates. Recall that U_j, $1 \le j \le 5$, denote the coordinates of the random variable U in the independence condition (1.2). Furthermore, let U_{js}, $s \ge 1$, denote independent copies of U_j.

LEMMA 3.2. *Assume the independence condition* (1.2). *Let M denote the largest integer such that $M \le N/3$. Then*

$$
|f(t)| \le \sum_{j=1}^{5}\left(\mathbf{E}\exp\left\{2it\lambda_j\tilde{Y}_{j1}\tilde{Y}_{j2}\right\}\right)^{5/2},
$$

where λ_j are defined in (1.3) *and*

$$
\tilde{Y}_{j1} = \sum_{1 \le s \le M}\tilde{U}_{js}/\sqrt{N} \quad \text{and} \quad \tilde{Y}_{j2} = \sum_{M+1 \le s \le 2M}\tilde{U}_{js}/\sqrt{N}.
$$

PROOF. According to the independence condition (1.2), $X = U + V$, where U and V are independent. We may write the sum S_N as the sum of four independent sums Y_1, Y_2, Y_3, Y_4,

$$
S_N = Y_1 + Y_2 + Y_3 + Y_4,
$$

such that each of Y_1, Y_2, Y_3 is a sum of M independent copies of U. Applying the symmetrization Lemma 3.1 we get

$$
|f(t)|^2 \le \mathbf{E}\exp\left\{2it\left(Q\tilde{Y}_1, \tilde{Y}_2\right)\right\}
$$

because Y_2 and Y_3 are identically distributed. The random variable u has independent coordinates. Therefore, the coordinates $\tilde{Y}_{j1}$, $1 \le j \le 5$, of $\tilde{Y}_1$ (resp. of $\tilde{Y}_2$) are independent as well and

$$
\mathbf{E}\exp\left\{2it\left(Q\tilde{Y}_1, \tilde{Y}_2\right)\right\} = \prod_{j=1}^{5}\mathbf{E}\exp\left\{2it\lambda_j\tilde{Y}_{j1}\tilde{Y}_{j2}\right\}.
$$

Now the geometric–arithmetic mean inequality

$$a_1 \cdots a_m \le \frac{1}{m}\left(a_1^m + \cdots + a_m^m\right) \quad \text{for } a_1, \dots, a_m \ge 0,$$

completes the proof of the lemma. □

The following lemma allows us to replace arbitrary random variables by Bernoulli random variables in the estimation of integrals of characteristic functions.

LEMMA 3.3. *Assume that W is a real mean zero random variable with positive variance and let $W_1, W_2, \dots$ be a sequence of independent copies of W. Write*

$$Z_N = (W_1 + \cdots + W_N)/\sqrt{N}$$

and let $\overline{Z}_N$ denote an independent copy of Z_N. Then there exist positive numbers c_1, c_2, δ and θ depending only on the distribution of W such that the integrals

$$I = \int_{T \le |t| \le H} \left(\mathbf{E}\exp\{itZ_N\overline{Z}_N\}\right)^{5/2} dt/|t|,$$

$$J = \int_{T \le |t| \le H} \left(\mathbf{E}\exp\{itZ_N\overline{Z}_N\}\right)^{5/2} dt$$

can be bounded from above as

$$I \le c_1 N^{-5/2}\ln(H/T) + c_1\int_{\delta T}^{\theta H}\left(\mathbf{E}\exp\{itN^{-1}B_{2m}\overline{B}_{2m}\}\right)^{5/2} dt/t,$$

$$J \le c_1 N^{-5/2}(H - T) + c_1\int_{\delta T}^{\theta H}\left(\mathbf{E}\exp\{itN^{-1}B_{2m}\overline{B}_{2m}\}\right)^{5/2} dt,$$

for all $0 \le T \le H$ and for some $m \ge c_2 N$.

PROOF. In the proof of the lemma we shall assume that $N \ge C(L(W))$ is sufficiently large. Otherwise the result is obvious. We shall restrict ourselves to proving the upper bound for I since the estimation of J is similar.

Let $\alpha_1, \alpha_2, \dots$ denote a sequence of i.i.d. random variables independent of all other random variables such that

$$\mathbf{P}\{\alpha_1 = 0\} = \mathbf{P}\{\alpha_1 = 1\} = 1/2.$$

Conditioning on $\alpha_1, \dots, \alpha_n$, we have

$$\mathbf{E}\{itZ_N\overline{Z}_N\} = \mathbf{E}\exp\{itT_N\overline{T}_N\},$$

where

$$T_N = \left(\alpha_1 W_1 + (1 - \alpha_1)\overline{W}_1 + \cdots + \alpha_N W_N + (1 - \alpha_N)\overline{W}_N\right)/\sqrt{N}.$$

Write $\varepsilon_j = 2\alpha_j - 1$ and $2Y_j = W_j - \overline{W}_j$. The random variables ε_j, $1 \le j \le N$, are symmetric Bernoulli random variables which are independent of all other random variables and such that $\mathbf{P}\{|\varepsilon_j| = 1\} = 1$. We may write

$$T_N = Q_N + R,$$

where

$$Q_N = (\varepsilon_1 Y_1 + \cdots + \varepsilon_N Y_N)/\sqrt{N},$$
$$R = \big(W_1 + \overline{W}_1 + \cdots + W_N + \overline{W}_N\big)/(2\sqrt{N}).$$

Conditioning we have

$$\mathbf{E}\exp\{itT_N\overline{T}_N\} \le \mathbf{E}\big|\mathbf{E}_\varepsilon \exp\{it(Q_N + R)\overline{T}_N\}\big| = \mathbf{E}\big|\mathbf{E}_\varepsilon \exp\{itQ_N\overline{T}_N\}\big|, \tag{3.1}$$

where $\mathbf{E}_\varepsilon$ denotes the conditional expectation given all random variables but $\varepsilon_1, \dots, \varepsilon_N$.

Let M denote the largest integer such that $M \le N/2$. We may split the sum Q_N into three (conditionally, given all random variables but $\varepsilon_1, \dots, \varepsilon_N$) independent parts,

$$Q_n = U + V + R,$$

where

$$U = \sum_{j=1}^{M} \varepsilon_j Y_j/\sqrt{N} \quad \text{and} \quad V = \sum_{j=M+1}^{2M} \varepsilon_j Y_j/\sqrt{N}.$$

Therefore, applying the inequality $2ab \le a^2 + b^2$, we get

$$\begin{aligned} 2\big|\mathbf{E}_\varepsilon \exp\{itQ_N\overline{T}_N\}\big| &\le \big|\mathbf{E}_\varepsilon \exp\{itU\overline{T}_N\}\big|^2 + \big|\mathbf{E}_\varepsilon \exp\{itV\overline{T}_N\}\big|^2 \\ &\le \mathbf{E}_\varepsilon \exp\big\{it\tilde{U}\overline{T}_N\big\} + \mathbf{E}_\varepsilon \exp\big\{it\tilde{V}\overline{T}_N\big\}, \end{aligned}$$

where

$$\tilde{U} = \sum_{j=1}^{M} \tilde{\varepsilon}_j Y_j/\sqrt{N} \quad \text{and} \quad \tilde{V} = \sum_{j=M+1}^{2M} \tilde{\varepsilon}_j Y_j/\sqrt{N}. \tag{3.2}$$

Substitution of this estimate in (3.1) together with equality of the distributions of U and V yields

$$\mathbf{E}\exp\{itT_N\overline{T}_N\} \le \mathbf{E}\exp\big\{it\tilde{U}\overline{T}_N\big\}. \tag{3.3}$$

Repeating this procedure with $\overline{T}_N$ instead of T_N, we arrive at

$$\mathbf{E}\exp\{itT_N\overline{T}_N\} \le \mathbf{E}\exp\{it\tilde{U}\tilde{V}\},$$

where $\tilde{V}$ denotes an independent copy of $\tilde{U}$, for example, the one given by (3.2).

The random variable W, as well as Y_1, has positive variance. Therefore, there exist numbers $\delta > 0$ and $\theta < \infty$ depending only on the distribution of W such that the probability $p = \mathbf{P}\{\delta \le |Y_1| \le \theta\}$ satisfies $0 < p < 1$. Consider the random variables

$$\xi_j = \mathbf{I}\big\{\delta \le |Y_j| \le \theta\big\}, \qquad 1 \le j \le M.$$

Then $\mathbf{E}\xi_j = p$ and

$$\begin{aligned}(3.4)\quad \mathbf{P}\{\xi_1 + \cdots + \xi_M \le Mp/2\} &\le c(Mp)^{-2}\mathbf{E}(\xi_1 - p + \cdots + \xi_M - p)^2 \\ &\le c/(Mp) \le c/(Np).\end{aligned}$$

Let m denote the largest integer such that $m \le Mp/2$. Due to (3.4) the complement of the event

$$A = \{\text{the number of r.v. } |Y_1|, \dots, |Y_M| \text{ contained in } [\delta, \theta] \text{ is at least } m\}$$

occurs with probability less than $c/(Np)$. There exists a (random) set, say $I_A \subset \{1, \dots, M\}$, of cardinality m such that

$$\{\delta \le |Y_i| \le \theta, \text{ for } i \in I_A\} \subset A.$$

Let $\mathbf{I}_A$ denote the indicator function of event A.

Similarly, the probability of the event

$$B = \{\text{the number of r.v. } |Y_{M+1}|, \dots, |Y_{2M}| \text{ contained in } [\delta, \theta] \text{ is at least } m\}$$

is bounded from below by $1 - c/(Np)$. There exists a (random) set, say $I_B \subset \{M+1, \dots, 2M\}$, of cardinality m such that

$$\{\delta \le |Y_i| \le \theta, \text{ for } i \in I_B\} \subset B.$$

Let $\mathbf{I}_B$ denote the indicator function of event B.

Due to the definitions of the events A and B we have

$$(3.5)\qquad \mathbf{E}\exp\{it\tilde{U}\tilde{V}\} = \mathbf{E}\mathbf{I}_A\mathbf{I}_B\exp\{it\tilde{U}\tilde{V}\} + R,$$

where $|R| \le c_1/N$ and the constant c_1 depends on the distribution of W only. Write $\mathbf{E}_\varepsilon$ for the conditional expectation given all random variables but $\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_M$. Then

$$\mathbf{E}\mathbf{I}_A\mathbf{I}_B\exp\{it\tilde{U}\tilde{V}\} = \mathbf{E}\mathbf{I}_A\mathbf{I}_B\mathbf{E}_\varepsilon\exp\{it\tilde{U}\tilde{V}\}.$$

Applying the geometric–arithmetic mean inequality, we have

$$\begin{aligned}\mathbf{E}_\varepsilon\exp\{it\tilde{U}\tilde{V}\} &\le \prod_{j\in I_A}\mathbf{E}_\varepsilon\exp\Big\{itN^{-1/2}Y_j\tilde{\varepsilon}_1\tilde{V}\Big\} \\ &\le \frac{1}{m}\sum_{j\in I_A}\Big(\mathbf{E}_\varepsilon\exp\Big\{itN^{-1/2}Y_j\tilde{\varepsilon}_1\tilde{V}\Big\}\Big)^m \\ &= \frac{1}{m}\sum_{j\in I_A}\mathbf{E}_\varepsilon\exp\Big\{itN^{-1/2}Y_jD_m\tilde{V}\Big\},\end{aligned}$$

where $D_m = \tilde{\varepsilon}_1 + \cdots + \tilde{\varepsilon}_m$. Thus

$$\mathbf{E}\mathbf{I}_A\mathbf{I}_B\exp\{it\tilde{U}\tilde{V}\} \le \frac{1}{m}\mathbf{E}\mathbf{I}_A\mathbf{I}_B\sum_{j\in I_A}\mathbf{E}_\varepsilon\exp\Big\{itY_jN^{-1/2}D_m\tilde{V}\Big\}.$$

Repeating this procedure with $\tilde{V}$ instead of $\tilde{U}$, we arrive at

$$\mathbf{E}\mathbf{I}_A\mathbf{I}_B \exp\{it\tilde{U}\tilde{V}\} \le \frac{1}{m^2}\mathbf{E}\mathbf{I}_A\mathbf{I}_B \sum_{j\in I_A}\sum_{l\in I_B} \exp\left\{itN^{-1}Y_jY_lD_m\overline{D}_m\right\}. \tag{3.6}$$

The elementary inequality $(a+b)^{5/2} \le 8a^{5/2} + 8b^{5/2}$, the Hölder inequality and (3.3), (3.5), and (3.6) together imply

$$\begin{aligned}&\left(\mathbf{E}\exp\left\{itT_N\overline{T}_N\right\}\right)^{5/2}\\ &\quad\le c_1N^{-5/2} + c_2\mathbf{E}\mathbf{I}_A\mathbf{I}_B\left(\sum_{j\in I_A}\sum_{l\in I_B}\frac{1}{m^2}\mathbf{E}_D\exp\left\{itN^{-1}Y_jY_lD_m\overline{D}_m\right\}\right)^{5/2},\end{aligned}$$

where $\mathbf{E}_D$ denotes the conditional expectation given all random variables but D_m and $\overline{D}_m$. The Hölder inequality implies

$$\left(\frac{1}{m^2}\sum_{j=1}^m\sum_{l=1}^m a_{jl}\right)^{5/2} \le \frac{1}{m^2}\sum_{j=1}^m\sum_{l=1}^m a_{jl}^{5/2}$$

whenever $a_{jl} \ge 0$. Thus

$$\left(\mathbf{E}\exp\left\{itT_N\overline{T}_N\right\}\right)^{5/2} \le c_1N^{-5/2} + c_2\mathbf{E}\mathbf{I}_A\mathbf{I}_B\sum_{j\in I_A}\sum_{l\in I_B}\frac{1}{m^2}I_{jl}, \tag{3.7}$$

where

$$I_{jl} = \mathbf{I}_A\mathbf{I}_B\left(\mathbf{E}_D\exp\left\{itN^{-1}Y_jY_lD_m\overline{D}_m\right\}\right)^{5/2}.$$

If $j \in I_A$ and $l \in I_B$, then $\delta \le |Y_j| \le \theta$ and $\delta \le |Y_l| \le \theta$. Thus

$$\begin{aligned}\int_{T\le|t|\le H} I_{jl}\frac{dt}{|t|} &= 2\mathbf{I}_A\mathbf{I}_B\int_T^H\left(\mathbf{E}_D\exp\left\{itN^{-1}|Y_jY_l|D_m\overline{D}_m\right\}\right)^{5/2}\frac{dt}{t}\\ &= 2\mathbf{I}_A\mathbf{I}_B\int_{|Y_jY_l|T}^{|Y_jY_l|H}\left(\mathbf{E}\exp\left\{itN^{-1}D_m\overline{D}_m\right\}\right)^{5/2}\frac{dt}{t}\\ &\le 2\mathbf{I}_A\mathbf{I}_B\int_{\delta^2T}^{\theta^2H}\left(\mathbf{E}\exp\left\{itN^{-1}D_m\overline{D}_m\right\}\right)^{5/2}\frac{dt}{t}\\ &\le 2\int_{\delta^2T}^{\theta^2H}\left(\mathbf{E}\exp\left\{itN^{-1}D_m\overline{D}_m\right\}\right)^{5/2}\frac{dt}{t}.\end{aligned} \tag{3.8}$$

Finally, integration of the inequality (3.7) over t together with (3.8), summation over m and the fact that D_m has the same distribution as B_{2m} imply the desired estimate of I. □

In the following lemma we consider integrals of characteristic functions for Fourier frequencies of magnitude up to $O(\sqrt{N/\ln N})$. The proof is based on the well-known techniques developed earlier for the estimation of the convergence rate in the CLT in Hilbert spaces, and we include this proof only for the sake of completeness.

LEMMA 3.4. *Let $k > 0$. There exists an absolute constant $c > 0$ such that*

$$I = \int_T^H \left(\mathbf{E}\exp\{itN^{-1}B_{2N}\overline{B}_{2N}\}\right)^{k/2} \frac{dt}{t} \le c_k T^{-k/2} + c_k N^{-4k},$$

for any $1 \le T \le H \le c\sqrt{N/\ln N}$. If $k > 2$, then

$$J = \int_0^H \left(\mathbf{E}\exp\{itN^{-1}B_{2N}\overline{B}_{2N}\}\right)^{k/2} dt \le c_k,$$

for any $0 \le H \le c\sqrt{N/\ln N}$.

PROOF. We shall restrict ourselves to the estimation of I. We have

$$\mathbf{P}\{|B_N| \ge x\sqrt{N}\} \le 2\exp\{-x^2/8\} \quad \text{for } x \le \sqrt{N},$$

by a well-known large deviation estimate [see, e.g., Petrov (1975)]. Hence

$$\mathbf{P}\{|\overline{B}_{2N}| \ge c\sqrt{N\ln N}\} \le 2\exp\{-8\ln N\} = 2N^{-8} \quad \text{for } N \ge c.$$

Let $\mathbf{I} = \mathbf{I}\{|\overline{B}_{2N}| \le c\sqrt{N\ln N}\}$. Then

$$I \le c_k N^{-4k} + c_k \int_T^H \left(\mathbf{E}\mathbf{I}\left(\mathbf{E}_{\varepsilon_1}\exp\{it\varepsilon_1 N^{-1}\overline{B}_{2N}\}\right)^{2N}\right)^{k/2} \frac{dt}{t}.$$

We have

$$\mathbf{I}\left|\mathbf{E}_{\varepsilon_1}\exp\{it\varepsilon_1 N^{-1}\overline{B}_{2N}\}\right| \le \exp\{-t^2\overline{B}_{2N}^2/(2N^2)\}$$

for $|t| \le c\sqrt{N/\ln N}$. Therefore,

$$I \le c_k N^{-4k} + c_k \int_T^H \left(\mathbf{E}\exp\left\{-\frac{t^2\overline{B}_{2N}^2}{N}\right\}\right)^{k/2} \frac{dt}{t}.$$

We may write

$$\mathbf{E}\exp\{-t^2N^{-1}\overline{B}_{2N}^2/2\} = \mathbf{E}\exp\{-it\xi N^{-1/2}\overline{B}_{2N}\},$$

where ξ denotes a standard normal $(0,1)$ random variable. Repeating the procedure with $\overline{B}_N$ instead of B_N we get

$$I \le c_k N^{-4k} + c_k \int_T^H \left(\mathbf{E}\exp\left\{-\frac{t^2\xi^2}{2}\right\}\right)^{k/2} \frac{dt}{t}.$$

Now the lemma follows since

$$\mathbf{E}\exp\{-t^2\xi^2/2\} = 1/\sqrt{1+2t^2}. \qquad \square$$

In the following lemma we estimate integrals of characteristic functions for Fourier frequencies of magnitude up to $O(1/N)$. It may be regarded as the main (technical) result of the paper. Its proof is related to the circle method of

Hardy and Littlewood (1920). However, the version presented here is adapted to the probabilistic setup. This version is even somewhat simpler than the original method because we could avoid using the "singular series" and the classification of subintervals into "minor" and "major" arcs since we have a smooth normal limit distribution function. Technically more involved approximations using "major arcs" would be necessary for proving limit theorems with number theoretic (discontinuous) limit densities.

LEMMA 3.5. *Let $\delta > 0$ denote a constant. Let U and V denote numbers satisfying*

$$\frac{\delta}{\sqrt{N(1+\ln N)}} \le U \le V \le \frac{1}{4}.$$

Then, for $k \ge 5$ and $N = 1, 2, \dots,$

$$I = \int_U^V (S(t))^{k/2} \frac{dt}{t} \le \frac{c_k(\delta)}{N}, \quad \text{where } S(t) = \mathbf{E}\exp\{i\pi t B_{2N}\overline{B}_{2N}\}.$$

Furthermore,

$$I_0 = \int_U^V (S(t))^{k/2}\, dt \le c_k(\delta).$$

PROOF. We shall estimate only the integral I. The estimation of I_0 is similar. Without loss of generality we may assume that

$$N \ge 2, \qquad \delta \le \frac{1}{8}, \qquad U = \frac{\delta}{\sqrt{N \ln N}} \quad \text{and} \quad V = \frac{1}{4}.$$

Thus the interval $[U, V]$ is nonempty.

We shall use a partition of the interval $[0, 1]$ by the so-called Farey sequence [see Hardy and Wright (1960) for all related facts]. Let n be a natural number. The Farey sequence F_n of order $n \ge 2$ is the ascending sequence of irreducible fractions between 0 and 1 with denominators not exceeding n. Thus p/q belongs to F_n if the integer numbers p and q satisfy

$$1 \le p < q \le n \quad \text{and} \quad (p, q) = 1,$$

where (p, q) denotes as usual the largest common divisor of p and q. The interval $(1/(n+1), n/(n+1)]$ may be represented as the union of nonintersecting (semiclosed) intervals

$$I\left(\frac{p}{q}\right) = \left(\frac{p}{q} - \gamma_{p/q}, \frac{p}{q} + \eta_{p/q}\right], \qquad \frac{p}{q} \in F_n,$$

such that $I(p/q)$ contains only one point of the set F_n, namely p/q, and this point divides $I(p/q)$ into two subintervals which have lengths bounded from below by $1/(q(2n-1))$ and from above by $1/(q(n-1))$; that is,

$$1/(q(2n-1)) \le \gamma_{p/q} \le 1/(q(n+1))$$

and

$$1/(q(2n-1)) \le \eta_{p/q} \le 1/(q(n+1)).$$

Let M be the smallest integer such that

$$\frac{1}{M} \le \frac{U}{4} = \frac{\delta}{4\sqrt{N \ln N}} \quad \text{and} \quad \mathbf{P}\{|B_{2N}| \ge M\} \le \frac{c}{N^3}. \tag{3.9}$$

Due to the large deviation estimate for B_{2N}, such $M = M(N)$ exists and $32 \le M \le c\sqrt{N \ln N}/\delta$. Let us consider the Farey series F_{2M} and the collection F_{2M}^* of $p/q \in F_{2M}$ such that the corresponding intervals $I(p/q)$ have nonempty intersection with the interval $[U, V]$. Due to the choice of M, the interval (U, V) is covered by intervals $I(p/q)$, $p/q \in F_{2M}^*$,

$$(U, V) \subset \{\cup I(p/q) \colon p/q \in F_{2M}^*\}.$$

We shall prove that for $t \in I(p/q)$,

$$S(t) \le \frac{cM}{N} + \frac{c}{q\sqrt{1+\theta^2 N^2}}, \quad \text{where } \theta = t - \frac{p}{q}. \tag{3.10}$$

This estimate implies the result of the lemma. Indeed, applying the elementary inequality $(a+b)^\alpha \le 2^\alpha(a^\alpha + b^\alpha)$ with $\alpha = k/2$, we obtain

$$I \le c_k J + c_k \left(\frac{M}{N}\right)^{k/2} \ln\left(\frac{V}{U}\right), \tag{3.11}$$

where

$$J = \sum_{p/q \in F_{2M}^*} J_{p/q} \quad \text{and} \quad J_{p/q} = \int_{I(p/q)} \frac{dt}{tq^{k/2}(1+\theta^2 N^2)^{k/4}}.$$

However, for $t = \theta + p/q \in I(p/q)$,

$$t \ge \frac{p}{q} - |\theta| \ge \frac{p}{q} - \frac{1}{2qM} \ge \frac{p}{2q} \quad \text{since } M \ge 1.$$

Therefore,

$$J_{p/q} \le \frac{2q}{pq^{k/2}N} \int_{-\infty}^{\infty} \frac{du}{(1+u^2)^{k/4}} \le \frac{c_k q}{pq^{k/2}N}$$

provided $k > 2$. Thus

$$J \le \frac{c_k}{N} \sum_{p/q \in F_{2M}^*} \frac{q}{pq^{k/2}} \le \frac{c_k}{N} \quad \text{for } k > 4, \tag{3.12}$$

since

$$\sum_{p/q \in F_{2M}^*} \frac{q}{pq^{k/2}} \le \sum_{q=1}^{\infty} \sum_{p=1}^{q} \frac{q}{pq^{k/2}} \le c \sum_{q=1}^{\infty} \frac{q(1+\ln q)}{q^{k/2}} < \infty.$$

Noting that $M \le c\sqrt{N}\ln N/\delta$ and collecting the estimates (3.11) and (3.12), we obtain the result of the lemma.

It remains to prove (3.10). Starting with the elementary inequality

$$\cos^2 u \le \exp\{-u^2/2\} \quad \text{for } |u| \le \pi/2,$$

we get

$$\cos^2 u \le \sum_{j=-\infty}^{\infty} \mathbf{I}\left\{-\frac{\pi}{2} < u - \pi j \le \frac{\pi}{2}\right\}\exp\left\{-\frac{1}{2}(u-\pi j)^2\right\} \quad \text{for all } u \in \mathbf{R}.$$

It follows that

$$\begin{aligned} S(t) &= \mathbf{E}\exp\{i\pi t B_{2N}\overline{B}_{2N}\} = \mathbf{E}\cos^{2N}(\pi t B_{2N}) \\ &\le \sum_{|j|\le N} \mathbf{E}\exp\left\{-\frac{N}{2}(\pi t B_{2N} - \pi j)^2\right\}, \end{aligned}$$

since $|tB_{2N}| \le N/2$, for $0 < t < 1/4$.

Write $p_l = \mathbf{P}\{\varepsilon_1 + \cdots + \varepsilon_{2N} = l\}$. Using $\pi^2 \ge 2$ we have

$$S(t) \le \sum_{|j|\le N}\sum_{|l|\le N} \exp\{-N(tl-j)^2\}p_l.$$

Due to (3.9),

$$S(t) \le \frac{c}{\sqrt{N}} + S_1(t), \tag{3.13a}$$

where

$$S_1(t) = \sum_{|j|\le N}\sum_{|l|\le M} \exp\{-N(tl-j)^2\}p_l. \tag{3.13b}$$

Note that $p_l = p_{-l}$. Thus

$$S_1(t) \le 2S_2(t), \quad \text{where } S_2(t) = \sum_{|j|\le N}\sum_{l=0}^{M} \exp\{-N(tl-j)^2\}p_l. \tag{3.14}$$

Let us write

$$t = \frac{p}{q} + \theta \quad \text{for } t \in I\left(\frac{p}{q}\right), \text{ with } |\theta| \le \frac{1}{2qM}$$

and let us decompose l modulo q:

$$l = mq + r, \quad \text{where } 0 \le r \le q-1 \text{ and } 0 \le m \le \frac{M}{q}.$$

Then

$$S_2(t) = \sum_{0\le m\le M/q}\sum_{r=0}^{q-1}\sum_{|j|\le N} \exp\{-N(\theta l + \{pr/q\} + [pr/q] + pm - j)^2\}p_l,$$

where we denote by $[\alpha]$ the integer nearest to a number α and by $\{\alpha\} = \alpha - [\alpha]$ the (signed) distance between α and the nearest integer (the definition of $[\alpha]$ when $|\{\alpha\}| = 1/2$ is inessential for our purposes).

If $r = 0$, then

$$\left|\theta l + \left\{\frac{pr}{q}\right\} + \left[\frac{pr}{q}\right] + pm - j\right| \geq |pm - j| - |\theta l| \geq 1 - \frac{1}{2q} \geq \frac{1}{2}$$

provided $j \neq pm$. If $1 \leq r \leq q - 1$, then

$$\left|\theta l + \left\{\frac{pr}{q}\right\} + \left[\frac{pr}{q}\right] + pm - j\right| \geq \left|\left[\frac{pr}{q}\right] + pm - j\right| - \frac{1}{2} - \frac{1}{2q} \geq 1 - \frac{3}{4} = \frac{1}{4}$$

provided $j \neq [pr/q] + pm$. Therefore,

$$S_2(t) \leq N^2 \exp\left\{-\frac{N}{16}\right\} + S_3(t) \leq \frac{c}{\sqrt{N}} + S_3(t), \tag{3.15}$$

where

$$S_3(t) = \sum_{0 \leq m \leq M/q} \sum_{r=0}^{q-1} \exp\left\{-N(\theta l + \{pr/q\})^2\right\} p_l.$$

Elementary calculations using Stirling's formula show that

$$p_l \leq \frac{C}{\sqrt{N}} \exp\left\{-\frac{l^2}{2N}\right\} \quad \text{for } 0 \leq l \leq \frac{N^{2/3}}{2}.$$

Therefore, for $l = mq + r$ and $l \leq M$,

$$p_l \leq \frac{C}{\sqrt{N}} \exp\left\{-\frac{q^2 m^2}{2N}\right\}. \tag{3.16}$$

Let us split the sum $S_3(t)$ according to whether $r = 0$ or $r > 0$. Using (3.16) we obtain

$$S_3(t) = S_4(t) + S_5(t), \tag{3.17}$$

where

$$S_4(t) = \frac{C}{\sqrt{N}} \sum_{0 \leq m \leq M/q} \exp\left\{-\frac{q^2 m^2 (1 + N^2\theta^2)}{2N}\right\},$$

$$S_5(t) = \frac{C}{\sqrt{N}} \sum_{0 \leq m \leq M/q} \exp\left\{-\frac{q^2 m^2}{2N}\right\} \sum_{r=1}^{q-1} \exp\left\{-N\left(\theta l + \left\{\frac{pr}{q}\right\}\right)^2\right\}.$$

For $S_4(t)$ we have

$$
\begin{aligned}
S_4(t) &\le \frac{C}{\sqrt{N}} + \frac{C}{\sqrt{N}} \sum_{1 \le m \le \infty} \exp\left\{ -\frac{q^2 m^2 (1 + N^2\theta^2)}{2N} \right\} \\
&\le \frac{C}{\sqrt{N}} + \frac{C}{\sqrt{N}} \int_{0 \le m \le \infty} \exp\left\{ -\frac{q^2 m^2 (1 + N^2\theta^2)}{2N} \right\} dm \qquad (3.18) \\
&\le \frac{C}{\sqrt{N}} + \frac{C}{q\sqrt{1 + \theta^2 N^2}}.
\end{aligned}
$$

The estimation of $S_5(t)$ is somewhat involved. Since numbers p and q are relative primes and $r \neq 0$, we have $|\{pr/q\}| \ge 1/q$ and

$$
\left| \theta l + \left\{ \frac{pr}{q} \right\} \right| \ge \left| \left\{ \frac{pr}{q} \right\} \right| - |\theta l| \ge \left| \left\{ \frac{pr}{q} \right\} \right| - \frac{1}{2q} \ge \frac{1}{2} \left| \left\{ \frac{pr}{q} \right\} \right|.
$$

Thus

$$
S_5(t) \le \frac{C}{\sqrt{N}} \sum_{0 \le m \le M/q} \exp\left\{ -\frac{q^2 m^2}{2N} \right\} \sum_{r=1}^{q-1} \exp\left\{ -\frac{N}{4} \left\{ \frac{pr}{q} \right\}^2 \right\}.
$$

The sets

$$
\{\{pr/q\} : 1 \le r \le q - 1\} \quad \text{and} \quad \{\{r/q\} : 1 \le r \le q - 1\}
$$

are equal since $(p, q) = 1$, and we have

$$
\sum_{r=1}^{q-1} \exp\left\{ -\frac{N}{4} \left\{ \frac{pr}{q} \right\}^2 \right\} \le 2 \sum_{r=1}^{\infty} \exp\left\{ -\frac{Nr^2}{4q^2} \right\} \le 2 \int_1^{\infty} \exp\left\{ -\frac{Nr^2}{8q^2} \right\} dr \le \frac{cq}{\sqrt{N}}.
$$

Collecting these estimates we get

$$
\begin{aligned}
S_5(t) &\le \frac{Cq}{N} \sum_{m=0}^{\infty} \exp\left\{ -\frac{q^2 m^2}{2N} \right\} \\
&\le \frac{Cq}{N} + \frac{Cq}{N} \int_0^{\infty} \exp\left\{ -\frac{q^2 m^2}{2N} \right\} dm \qquad (3.19) \\
&\le \frac{Cq}{N} + \frac{C}{\sqrt{N}} \le \frac{CM}{N}
\end{aligned}
$$

since $q \le 2M$ and $\sqrt{N} \le cM$.

Now (3.10) follows from (3.13a)–(3.19).

4. A smoothing lemma.

LEMMA 4.1. *Let F be a distribution function with the characteristic function f. Then, for any $H > 0$,*

$$
F(x) = \frac{1}{2} + \frac{i}{2\pi} \mathrm{P.V.} \int_{-H}^{H} \exp(-ixt) f(t) \frac{dt}{t} + R,
$$

where

$$|R| \le \frac{1}{H}\int_{-H}^{H}|f(t)|\,dt,$$

and where P.V. *denotes Cauchy's principal value,*

$$\text{P.V.}\int_{\mathbf{R}} u(t)\,dt = \lim_{h\downarrow 0}\int_{|t|\ge h} u(t)\,dt.$$

Let G denote a distribution function with the characteristic function g. Lemma 4.1 obviously implies

$$|F(x) - G(x)| \le \frac{1}{2\pi}\int_{-H}^{H}|f(t) - g(t)|\frac{dt}{|t|} + \frac{1}{H}\int_{-H}^{H}(|f(t)| + |g(t)|)\,dt,$$

which recalls the classical Esseen inequality for characteristic functions.

Lemma 4.1 is a consequence of the following smoothing inequalities of Prawitz (1972). Define the function $2K(s) = K_1(s) + iK_2(s)/(\pi s)$ by

$$K_1(s) = 1 - |s|, \qquad K_2(s) = \pi s(1 - |s|)\cot \pi s + |s| \quad \text{for } |s| \le 1,$$

and $K_1(s) = K_2(s) = 0$, for $|s| \ge 1$. Then (notice that all integrals are real and that the expressions under the signs of integrals vanish unless $|t| \le H$)

$$F(x+) \le \frac{1}{2} + \text{P.V.}\int_{\mathbf{R}} \exp(-ixt)\frac{1}{H}K\left(\frac{t}{H}\right)f(t)\,dt, \tag{4.1}$$

$$F(x-) \ge \frac{1}{2} - \text{P.V.}\int_{\mathbf{R}} \exp(-ixt)\frac{1}{H}K\left(-\frac{t}{H}\right)f(t)\,dt, \tag{4.2}$$

where $F(x+) = \lim_{z\downarrow x} F(z)$ and $F(x-) = \lim_{z\uparrow x} F(z)$.

The following lemma is elementary.

LEMMA 4.2. *For* $0 \le s \le 1$ *we have*

$$K_2(0) = 1, \qquad K_2(1) = 0, \qquad K_2(\tfrac{1}{2}) = \tfrac{1}{2},$$
$$K_2'(s) \le 0, \qquad K_2(s) + K_2(1-s) = 1.$$

Furthermore,

$$1 - 2(1-s)\sin^2\frac{\pi s}{2} \le K_2(s) \le 1 \quad \text{for } 0 \le s \le \frac{1}{2},$$

and

$$0 \le K_2(s) \le 2s\sin^2\frac{\pi(1-s)}{2} \quad \text{for } \frac{1}{2} \le s \le 1.$$

It follows from Lemma 4.2 that

$$|1 - K_2(s)| \le 2|s| \quad \text{for all } s \in \mathbf{R}.$$

Therefore, (4.1), (4.2) and the definition of the function K imply Lemma 4.1.

It is known [see, for instance, Chung (1974)], that if we redefine a distribution function G at discontinuity points (say x) as $2G(x) = G(x+) + G(x-)$, then

$$G(x) = \frac{1}{2} + \frac{i}{2\pi} \lim_{M\to\infty} \text{P.V.}\int_{|t|\le M} \exp(-itz) g(t) \frac{dt}{t}, \tag{4.3}$$

where g denotes the characteristic function of G. One can generalize (4.3) to functions of the bounded variation.

REFERENCES

BENTKUS, V. (1984a). Asymptotic expansions in the CLT in Hilbert space. *Lithuanian Math. J.* **24** 210–225.

BENTKUS, V. (1984b). Asymptotic expansions for distributions of sums of independent random elements of a Hilbert space. *Lithuanian Math. J.* **24** 305–319.

BENTKUS, V. (1985). Asymptotic expansions in local limit theorem in a Hilbert space. *Lithuanian Math. J.* **25** 1–10.

BENTKUS, V., GÖTZE, F., PAULAUSKAS, V. and RAČKAUSKAS, A. (1991). The accuracy of Gaussian approximation in Banach spaces. In *Encyclopedia of Mathematical Sciences*. Springer, Berlin. To appear. [*Itogi Nauki i Tehniki* **81** 39–139 (1991) (in Russian).]

BENTKUS, V., GÖTZE, F. and ZITIKIS, R. (1993). Asymptotic expansions and local limit theorems in Banach spaces with applications to ω-statistics. *J. Theoret. Probab.* **6** 727–780.

BENTKUS, V. and ZALESSKII, B. (1985). Asymptotic expansions with nonuniform remainders in the CLT in Hilbert space. *Lithuanian Math. J.* **25** 199–208.

BHATTACHARYA, R. N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*. Wiley, New York.

BLECHER, P. M. and LEBOWITZ, J. L. (1993). Energy-level statistics of model quantum systems: universality and scaling in a lattice-point problem. Unpublished manuscript.

CHUNG, K. L. (1974). *A Course in Probability Theory*. Academic Press, New York.

ESSEEN, C. G. (1945). Fourier analysis of distribution functions. *Acta Math.* **77** 1–125.

FRICKER, F. (1982). *Einführung in die Gitterpunktlehre*. Birkhäuser, Basel.

GÖTZE, F. (1979). Asymptotic expansions for bivariate von Mises functionals. *Z. Wahrsch. Verw. Gebiete* **50** 333–355.

GÖTZE, F. (1984). Expansions for von Mises functionals. *Z. Wahrsch. Verw. Gebiete* **65** 599–625.

GÖTZE, F., PROHOROV, YU. V. and ULYANOV, V. V. (1994). Bounds for characteristic functions of polynomials in asymptotically normal variables. Preprint 006, Univ. Bielefeld.

GRAHAM, S. W. and KOLESNIK, G. (1991). *Van der Corput's Method of Exponential Sums*. Cambridge Univ. Press.

HARDY, G. H. and LITTLEWOOD, J. E. (1920). A new solution of Waring's problem. *Quart. J. Math. Oxford* **48** 272–293.

HARDY, G. H. and WRIGHT, E. M. (1960). *An Introduction to the Theory of Numbers*. Clarendon, Oxford.

JARNIK, C. G. (1928). Über Gitterpunkte in mehrdimensionalen Ellipsoiden. *Math. Ann.* **100** 699–721.

KENDALL, D. G. (1948). On the number of lattice points inside a random oval. *Quart. J. Math. Oxford* **19** 1–26.

KRÄTZEL, E. (1988). *Lattice Points*. Kluwer, Dordrecht.

LANDAU, E. (1915). Zur analytischen Zahlentheorie der definiten quadratischen Formen (Über die Gitterpunkte in einem mehrdimensionalen Ellipsoid). *Sitzungsber. Preuss. Akad. Wiss.* **31** 458–476.

LANDAU, E. (1924). Über Gitterpunkte in mehrdimensionalen Ellipsoiden. *Math. Z.* **21** 126–132.

LANDAU, E. (1962). *Ausgewählte Abhandlungen zur Gitterpunktlehre* (A. Walfisz, ed.). Deutscher Verlag der Wissenschaften, Berlin.

MAJOR, P. (1992). Poisson law for the number of lattice points in a random strip with finite area. *Probab. Theory Related Fields* **92** 423–464.

MATTHES, T. K. (1970). The multivariate central limit theorem for regular convex sets. *Ann. Probab.* **3** 503–515.

NAGAEV, S. V. (1988). A Berry–Esseen type estimate for sums of Hilbert space values r.v. *Siberian Math. J.* **30** 413–423.

PETROV, V. V. (1975). *Sums of Independent Random Variables.* Springer, New York.

PRAWITZ, H. (1972). Limits for a distribution, if the characteristic function is given in a finite domain. *Skand. Aktuartidskr.* **55** 138–154.

SAZONOV, V. V. (1981). *Normal Approximation—Some Recent Advances. Lecture Notes in Math.* **879**. Springer, Berlin.

SAZONOV, V. V., UL'YANOV, V. V. and ZALESSKII, B. A. (1989). A precise estimate of the rate of convergence in CLT in Hilbert space, *Mat. Sb.* **180** 1587–1613.

SENATOV, V. V. (1989). *On the Estimate of the Rate of Convergence in the Central Limit Theorem in Hilbert Space. Lecture Notes in Math.* **1412** 309–327. Springer, Berlin.

SINAI, YA. G. (1991). *Mathematical Problems in the Theory of Quantum Chaos. Lecture Notes in Math.* **1469** 41–59.

VINOGRADOV, I. M. (1934). A new evaluation of $G(n)$ in Waring's problem. *Dokl. Acad. Sci. USSR* **5** 251–253.

WALFISZ, A. (1927). Über Gitterpunkte in mehrdimensionalen Ellipsoiden. *Math. Z.* **27** 245–268.

WALFISZ, A. (1957). *Gitterpunkte in Mehrdimensionalen Kugeln.* Panstwowe Wydawn. Naukowe, Warsaw.

WEYL, H. (1915/1916). Über die Gleichverteilung der Zahlen mod-Eins. *Ann. Phys.* **77** 313–352.

YARNOLD, J. (1972). Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set. *Ann. Math. Statist.* **43** 1566–1580.

YURINSKII, V. V. (1982). On the accuracy of normal approximation of the probability of hitting a ball. *Theory Probab. Appl.* **27** 280–289.

ZALESSKII, B. A. (1982). Estimates of the accuracy of normal approximations in a Hilbert space. *Theory Probab. Appl.* **27** 290–298.

FAKULTÄT FÜR MATHEMATIK
UNIVERSITÄT BIELEFELD
POSTFACH 100131
33501 BIELEFELD 1
GERMANY
E-mail: bentkus@mathematik.uni-bielefeld.de
goetze@mathematik.uni-bielefeld.de