APPROXIMATION AND PROBABILITY
BANACH CENTER PUBLICATIONS, VOLUME 72
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 2006

GEBELEIN'S INEQUALITY AND ITS CONSEQUENCES

M. BEŚKA
Faculty of Applied Mathematics, Gdańsk University of Technology
Narutowicza 11/12, 80-952 Gdańsk, Poland
E-mail: beska@mif.pg.gda.pl

Z. CIESIELSKI
Institute of Mathematics, Polish Academy of Sciences
Abrahama 18, 81-125 Sopot, Poland
E-mail: Z.Ciesielski@impan.gda.pl

Abstract. Let $(X_i, i = 1, 2, \dots)$ be the normalized gaussian system such that $X_i \in N(0,1)$, $i = 1, 2, \dots$ and let the correlation matrix $\rho_{ij} = E(X_i X_j)$ satisfy the following hypothesis:

$$C = \sup_{i\geq 1} \sum_{j=1}^{\infty} |\rho_{i,j}| < \infty.$$

We present Gebelein's inequality and some of its consequences: Borel-Cantelli type lemma, iterated log law, Levy's norm for the gaussian sequence etc. The main result is that

$$\frac{f(X_1) + \dots + f(X_n)}{n} \to 0 \text{ a.s.}$$

for $f \in L^1(\nu)$ with $(f, 1)_\nu = 0$.

1. Mehler's kernel and Gebelein's inequality. Let (X, Y) be a gaussian random vector such that $X, Y \in N(0,1)$ and $E(XY) = \rho$, $(|\rho| < 1)$. Its density is equal then to

$$p(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 + y^2 - 2\rho xy)\right).$$

We denote by ν the normalized one-dimensional gaussian measure i.e.

$$\nu(dx) = p(x)\, dx = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx,$$

2000 *Mathematics Subject Classification*: 47B34, 60F15, 60F20.

The paper is in final form and no version of it will be published elsewhere.

and use L^p (or $L^p(\nu)$) for $L^p(\mathbb{R}, d\nu)$. In L^p we have the norm

$$\|f\|_p = \Big(\int_{\mathbb{R}} |f(x)|^p \,\nu(dx)\Big)^{\frac{1}{p}}, \quad 1 \le p \le \infty,$$

and in L^2 the scalar product

$$(f,g)_\nu = \int_{\mathbb{R}} f(x)\,g(x)\,\nu(dx).$$

For $f \in L^2$ the conditional expectation

$$P_\rho f(y) = E(f(X)|Y = y) \tag{1.1}$$

can be computed. Introducing r.v. $Z \in N(0,1)$ such that Z, Y are independent, we find that the gaussian vectors (X,Y) and (U,Y) with $U = \rho Y + \sqrt{1-\rho^2}Z$ have the same joint distribution. Thus, with $h(y) = Ef(\rho y + \sqrt{1-\rho^2}Z)$, we have

$$E(f(X)g(Y)) = E(f(U)g(Y)) = E(h(Y)g(Y)),$$

whence

$$P_\rho f(y) = E(f(X)|Y = y) = h(y). \tag{1.2}$$

This implies that

$$P_\rho f(x) = \int_{\mathbb{R}} K(x,y;\rho)\,f(y)\,\nu(dy), \tag{1.3}$$

where

$$K(x,y;\rho) = \frac{p(x,y;\rho)}{p(x)p(y)} = \frac{1}{\sqrt{1-\rho^2}} \exp\Big(-\frac{1}{2(1-\rho^2)}(\rho^2(x^2+y^2) - 2\rho xy)\Big), \tag{1.4}$$

is the Mehler kernel (see [S]). It follows immediately by (1.2) that

$$\int_{\mathbb{R}} K(x,y;\rho)\,\nu(dy) = 1. \tag{1.5}$$

Since the kernel is symmetric and positive we obtain by Hölder's inequality

PROPOSITION 1.1. *Given the gaussian vector (X,Y) and $f \in L^p$, $1 \le p \le \infty$, we have*

$$\|P_\rho f\|_p \le \|f\|_p. \tag{1.6}$$

We now substitute $\rho = e^{-t}$ and set $Q_t = P_\rho$ and $K_t(x,y) = K(x,y;\rho)$. In this notation we have

PROPOSITION 1.2. *For $f \in L^1$ and $t, s > 0$ the semigroup property takes place i.e.*

$$Q_{s+t}f = Q_s(Q_t f) = Q_t(Q_s f), \tag{1.7}$$

and

$$K_{s+t}(x,y) = \int_{\mathbb{R}} K_s(x,z)K_t(z,y)\,\nu(dz). \tag{1.8}$$

Proof. Use formulas (1.1) and (1.3). ■

The Mehler kernel has its representation in terms of orthogonal Hermite polynomials $\{H_n; n = 0, 1, \ldots\}$ which are uniquely determined by the following properties: H_n is of degree n and

$$\int_{\mathbb{R}} H_n(x)\, H_m(x)\, \exp(-x^2)\, dx = 2^n n! \sqrt{\pi} \delta_{n,m}, \quad \text{for} \quad n, m = 0, 1, \ldots$$

Defining

$$h_n(x) = \frac{H_n(x/\sqrt{2})}{\sqrt{2^n n!}}$$

we obtain that

$$(h_n, h_m)_\nu = \delta_{n,m} \quad \text{for} \quad n, m = 0, 1, \ldots.$$

The orthonormal system $\{h_n, n = 0, 1, \ldots\}$ is complete in L^2 (see Natanson C.T.F, completeness is due to Steklov) and

$$K(x, y; \rho) = \sum_0^\infty \rho^n h_n(x) h_n(y), \quad |\rho| < 1, \tag{1.9}$$

whence in particular

$$P_\rho f = \sum_0^\infty \rho^n (f, h_n)_\nu h_n \quad \text{for} \quad f \in L^2. \tag{1.10}$$

Now, the Parseval identity gives

$$\|P_\rho f\|_2^2 = \sum_0^\infty \rho^{2n} |(f, h_n)_\nu|^2. \tag{1.11}$$

As a consequence from (1.11) we obtain Gebelein's inequality (1.12) (see [G] and [DK])

PROPOSITION 1.3. *If $f \in L^2$ and $(f, 1)_\nu = 0$, then*

$$\|P_\rho f\|_2 \leq |\rho| \cdot \|f\|_2, \tag{1.12}$$

or equivalently for any $g \in L^2$ and f as above

$$|(P_\rho f, g)_\nu| \leq |\rho| \cdot \|f\|_2 \cdot \|g\|_2. \tag{1.13}$$

In both inequalities we have equality if and only if $f(x) = c \cdot x$.

2. Applications of Gebelein's inequality. The normalized gaussian sequence (X_i, $i = 1, 2, \ldots$) of random variables is given. In particular $X_i \in N(0, 1)$ for each i. It is assumed that the correlation matrix $\rho_{i,j} = E(X_i X_j)$ satisfies the following *hypothesis*

$$C = \sup_i \sum_j |\rho_{i,j}| < \infty. \tag{R}$$

Related formulation of the following lemma for the first time appears in [R] and the proof is presented here for completeness.

LEMMA 2.1. *Under hypothesis (R) for arbitrary Borel subsets ($A_i, i = 1, 2, \ldots$) of $\mathbb{R}$ we have*

$$E\left(\frac{\sum_{i=1}^n I_{A_i}(X_i)}{\sum_{i=1}^n P\{X_i \in A_i\}} - 1\right)^2 \leq \frac{C}{\sum_{i=1}^n P\{X_i \in A_i\}}. \tag{2.1}$$

Proof. For any two dimensional normalized gaussian vector (X, Y) and for any $f, g \in L^2$ with the property that $Ef(X) = Eg(Y) = 0$ we have by (1.13)

$$|E(f(X)g(Y))| = |(P_\rho f, g)_\nu| \le |\rho| \|f\|_2 \|g\|_2 = |\rho| \sqrt{E(f(X)^2)} \sqrt{E(g(Y)^2)},$$

with $\rho = E(XY)$. This inequality applied to the functions $f_i(x) = I_{A_i}(x) - P\{X_i \in A_i\}$ and $g_j(x) = I_{A_j}(x) - P\{X_j \in A_j\}$, where I_A is the indicator of the set A, gives

$$\begin{aligned} |P\{X_i \in A_i, X_j \in A_j\} - P\{X_i \in A_i\}P\{X_j \in A_j\}| \\ &\le |\rho_{i,j}| \sqrt{P\{X_i \in A_i\}P\{X_i \notin A_i\}P\{X_j \in A_j\}P\{X_j \notin A_j\}} \\ &\le |\rho_{i,j}| \sqrt{P\{X_i \in A_i\}P\{X_j \in A_j\}} \\ &\le |\rho_{i,j}| \frac{P\{X_i \in A_i\} + P\{X_j \in A_j\}}{2}. \end{aligned}$$

Using this we obtain

$$\begin{aligned} E\left(\frac{\sum_{i=1}^n I_{A_i}(X_i)}{\sum_{i=1}^n P\{X_i \in A_i\}} - 1 \right)^2 &\le \frac{\sum_{i=1}^n \sum_{j=1}^n |\rho_{i,j}|(P\{X_i \in A_i\} + P\{X_j \in A_j\})}{2(\sum_{i=1}^n P\{X_i \in A_i\})^2} \\ &\le \frac{C}{\sum_{i=1}^n P\{X_i \in A_i\}}, \end{aligned}$$

and the proof is complete. ∎

COROLLARY 2.1 (Borel-Cantelli Lemma). *Let the normalized gaussian sequence* $(X_i, i = 1, 2, \ldots)$ *satisfy hypothesis* (R) *and let* $(A_i, i = 1, 2, \ldots)$ *be a sequence of Borel sets in* $\mathbb{R}$ *such that*

$$\sum_{i=1}^{\infty} P\{X_i \in A_i\} = \infty. \tag{2.2}$$

then

$$P\{X_i \in A_i \ i.o.\} = 1. \tag{2.3}$$

Moreover, if

$$\sum_{i=1}^{\infty} P\{X_i \in A_i\} < \infty, \tag{2.4}$$

then

$$P\{X_i \in A_i \ i.o.\} = 0. \tag{2.5}$$

COROLLARY 2.2 (Iterated log law). *Let the normalized gaussian sequence* (X_i) *satisfy hypothesis* (R). *Then*

$$P\left\{ \limsup_n \frac{X_n^2 - 2\log n}{\log\log n} = 1 \right\} = 1. \tag{2.6}$$

Proof. Using for large a the asymptotic expansion (see [H])

$$\int_a^\infty \exp\left(-\frac{x^2}{2}\right) dx = \frac{\exp(-\frac{a^2}{2})}{a} \left(1 + \sum_{k=1}^{\infty} (-1)^k \frac{(2k-1)!!}{a^{2k}}\right) \tag{2.7}$$

we find that for the choice

$$A_n = A_n(\gamma) := (\sqrt{2\log n + \gamma \log\log n}, \infty) \cup (-\infty, -\sqrt{2\log n + \gamma \log\log n})$$

with $\gamma > 0$ the following two series are equiconvergent:

$$\sum_{n\geq 10} P\{X_n \in A_n(\gamma)\} \quad \text{and} \quad \sum_{n\geq 10} \frac{1}{n(\log n)^{\frac{1+\gamma}{2}}}. \ \blacksquare$$

Corollary 2.3 (Levy's norm). *Let the normalized gaussian sequence* (X_i) *satisfy hypothesis* (R). *Then*

$$P\Big\{\limsup_j \frac{1}{\sqrt{j}} \sup_{1\leq k\leq 2^j} |X_{2^j+k}| = \sqrt{2\log 2}\Big\} = 1. \tag{2.8}$$

Proof. For $\eta \geq 0$ and $1 \leq k \leq 2^j$, $j \geq 0$ define

$$A_{2^j+k}(\eta) = \mathbb{R}\backslash(-\sqrt{(2j\log 2)(1+\eta)}, \sqrt{(2j\log 2)(1+\eta)}).$$

Let $X \in N(0,1)$. Then

$$\begin{aligned}\sum_n P\{X_n \in A_n(\eta)\} &= \sum_j \sum_{1\leq k\leq 2^j} P\{|X| \geq \sqrt{(2j\log 2)(1+\eta)}\} \\ &= \sum_j 2^j P\{|X| \geq \sqrt{(2j\log 2)(1+\eta)}\}.\end{aligned}$$

However, the last series equiconverges with

$$\sum_{j=1}^{\infty} \frac{1}{\sqrt{j}2^{\eta j}}.$$

Thus, in case of $\eta > 0$ this implies the easy part of the statement. In case $\eta = 0$ the series diverges and therefore by Lemma 2.1

$$P\Big\{\sum_{n=1}^{\infty} I_{A_n(0)}(X_n) = \infty\Big\} = 1,$$

whence

$$P\Big\{\sum_{1\leq k\leq 2^j} I_{|X_{2^j+k}|\geq\sqrt{2j\log 2}} \geq 1 \text{ i.o.}\Big\} = 1$$

and consequently

$$P\Big\{\limsup_j \frac{1}{\sqrt{j}} \sup_{1\leq k\leq 2^j} |X_{2^j+k}| \geq \sqrt{2\log 2}\Big\} = 1. \ \blacksquare$$

Corollary 2.4. *Let* $(X_i,\ i = 1, 2, \ldots)$ *be a centered gaussian sequence with correlation matrix* (ρ_{ij}) *satisfying hypothesis* (R). *Then*

$$\bigvee_{r>0} \quad \sum_{i=1}^{\infty} P\{|X_i| > r\} < \infty \iff P\{\sup_i |X_i| < \infty\} = 1, \tag{a}$$

$$\bigwedge_{r>0} \quad \sum_{i=1}^{\infty} P\{|X_i| > r\} < \infty \iff P\{\lim_i X_i = 0\} = 1. \tag{b}$$

Proof. For $i = 1, 2, \ldots$ define $Y_i = X_i/\sigma_i$, where σ_i denotes the standard deviation of X_i. It is clear that (Y_i) forms normalized gaussian sequence satisfying hypothesis (R). Applying now Corollary 2.1 to the gaussian sequence (Y_i) and to the sets $A_{i,r} = \{y : |y| > r/\sigma_i\}$ we obtain a proof of our statement. $\blacksquare$

Corollary 2.5. *Let $X = (X_i, i = 1, 2, \ldots)$ be as in Corollary 2.4. Then the probability distribution of X is concentrated on the Banach space c_0 if and only if*

$$\bigwedge_{r>0} \quad \sum_{i=1}^{\infty} \exp\Big\{-\frac{r}{\sigma_i^2}\Big\} < \infty, \quad \text{where} \quad \sigma_i^2 = EX_i^2, \quad i = 1, 2, \ldots. \tag{2.9}$$

Proof. It is well known that condition (2.9) is sufficient in the more general situation, without the hypothesis (R) (see [VTC]). The necessity of condition (2.9) follows (using similar methods as in the independent case) from Corollary 2.4 (b) and from the asymptotic expansion (2.7). ■

3. The laws of large numbers. Let us now consider the average

$$\frac{f(X_1) + \cdots + f(X_n)}{n}, \tag{3.1}$$

where f is a Borel function. The question is: For which functions f is the average (3.1) convergent to $Ef(X_1)$? In [BC] it was proved that the average (3.1) converges in $L^1(P)$ for $f \in L^1(\nu)$ and for f being algebraic polynomials we get a.s. convergence. It was also conjectured that for $f \in L^1(\nu)$ (3.1) converges a.s. In what follows we prove this conjecture.

In sequel we need the following result (see for instance [B]):

Theorem 3.1. *Let the distribution of the random variable Y be determined by its moments and let the random variables $(Y_n, n \geq 1)$ have moments of all orders. Moreover, let*

$$\lim_{n\to\infty} E(Y_n^r) = E(Y^r), \qquad r = 1, 2, \ldots.$$

Then $Y_n \Rightarrow Y$ in distribution, as $n \to \infty$. ■

Now, we can state

Theorem 3.2. *Let the normalized gaussian sequence $(X_i, i = 1, 2, \ldots)$ satisfy the hypothesis (R). Moreover, let f be a bounded function and let its set of points of discontinuity be of Lebesgue measure zero. Then*

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \xrightarrow[n\to\infty]{} Ef(X_1), \qquad a.s.$$

Proof. By Theorem 2.3 [BC] it follows that we can find a measurable set $\Omega_0 \subset \Omega$, $P(\Omega_0) = 1$, such that

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k(\omega) \xrightarrow[n\to\infty]{} EX_1^k, \qquad \omega \in \Omega_0, \quad k = 1, 2, \ldots. \tag{3.2}$$

Next, define the empirical distribution functional

$$\Omega_0 \ni \omega \mapsto F_n(\,\cdot\,, \omega) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(\omega)}(\cdot); \tag{3.3}$$

and observe, that by (3.2) for any $k \geq 1$

$$\int_{\mathbb{R}} x^k \, dF_n(x,\omega) = \frac{1}{n}\sum_{i=1}^{n} X_i^k(\omega) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} x^k \, d\nu(x), \qquad \omega \in \Omega_0. \tag{3.4}$$

Hence and from Theorem 3.1 (the gaussian distribution is determined by moments) for every $\omega \in \Omega_0$

$$F_n(\,\cdot\,,\omega) \Longrightarrow \nu, \qquad n \to \infty. \tag{3.5}$$

Therefore for the function f satisfying the assumptions of our theorem we have:

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i(\omega)) = \int_{\mathbb{R}} f(x)\, dF_n(x,\omega) \to \int_{\mathbb{R}} f(x)\, d\nu(x) = Ef(X_1), \qquad \omega \in \Omega_0,$$

and the proof is complete. ∎

By $\mathbb{R}_0^\infty$ we denote the set of all real sequences with a finite number of nonzero terms, i.e.

$$\mathbb{R}_0^\infty = \{(x_i) \in \mathbb{R}^\infty : x_j = 0 \text{ for } j > n, \text{ for some } n\},$$

Let us define a linear operator $R : \mathbb{R}_0^\infty \to \mathbb{R}$ by the formula

$$R(x) = \Big(\sum_{j=1}^{\infty} |\rho_{ij}| x_j\Big), \quad x = (x_j) \in \mathbb{R}_0^\infty.$$

It is well known that R can be extended to a continuous linear operator over the spaces l^p, $p \geq 1$. The proof below is given here just for the sake of completeness.

LEMMA 3.1. *For every $1 \leq p \leq \infty$ we can extend the operator R to a continuous operator $R : l^p \to l^p$ with $\|R\| \leq C$.*

Proof. Let $x = (x_i) \in \mathbb{R}_0^\infty$ and denote $r_i = \sum_{j=1}^\infty |\rho_{ij}|$, $i = 1, 2, \ldots$. Then by Jensen's inequality and by symmetry of the matrix $(|\rho_{ij}|)$ we have

$$\begin{aligned}
\|R(x)\|_{l^p}^p &= \sum_{i=1}^{\infty}\Big|\sum_{j=1}^{\infty}|\rho_{ij}|\,x_j\Big|^p \leq \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty}|\rho_{ij}|\,|x_j|\Big)^p \\
&= \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty}\frac{|\rho_{ij}|}{r_i}\,|x_j|\Big)^p r_i^p \leq \sum_{i=1}^{\infty} r_i^p \sum_{j=1}^{\infty}\frac{|\rho_{ij}|}{r_i}\,|x_j|^p \leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} r_i^{p-1}|\rho_{ij}|\,|x_j|^p \\
&\leq C^{p-1}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|\rho_{ij}|\,|x_j|^p \leq C^{p-1}\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}|\rho_{ji}|\,|x_j|^p \leq C^p\sum_{j=1}^{\infty}|x_j|^p = C^p\,\|x\|_{l^p}^p.
\end{aligned}$$ ∎

LEMMA 3.2. *Let the normalized gaussian sequence $(X_i, i = 1, 2, \ldots)$ satisfy the hypothesis (R) and let $(f_i, i = 1, 2, \ldots) \subset L^2(\nu)$. Then for each $n \geq 1$ we have*

$$Var\Big(\sum_{i=1}^{n} f_i(X_i)\Big) \leq C\sum_{i=1}^{n} Var(f_i(X_i)).$$

Proof. This follows immediately by Gebelein's inequality and by Lemma 3.1. ∎

Using the last two lemmas and the method adapted from [E1] or [E2] we can prove the a.s. convergence of (3.1) for $f \in L^1(\nu)$. Namely, THEOREM 3.3. *Let the normalized gaussian sequence* $(X_i, i = 1, 2, \ldots)$ *satisfy the hypothesis* (R) *and* $f \in L^1(\nu)$. *Then*

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \xrightarrow[n\to\infty]{} Ef(X_1), \qquad a.s.$$

Proof. We start with the observation that it suffices to prove the theorem for $f \in L^1(\mu)$ and $f \geq 0$. For each $\alpha > 1$ let us define a sequence $(k_n, n = 0, 1, 2, \ldots)$ of integers as follows:

$$k_0 = 1, \quad k_n = [\alpha^n], \quad n \geq 1,$$

where $[x]$ is the greatest integer less than or equal to x. It is clear that

$$\lim_{n\to\infty} \frac{k_n}{k_{n+1}} = \frac{1}{\alpha}.$$

Moreover

$$\bigwedge_{m\geq 1} \bigvee_{n(m)\geq 1} k_{n(m)-1} \leq m \leq k_{n(m)}. \tag{3.6}$$

It now follows that for $f \geq 0$ that

$$\frac{k_{n(m)-1}}{k_{n(m)}}\,\frac{S_{k_{n(m)-1}}}{k_{n(m)-1}} = \frac{S_{k_{n(m)-1}}}{k_{n(m)}} \leq \frac{S_m}{m} \leq \frac{S_{k_{n(m)}}}{k_{n(m)-1}} = \frac{k_{n(m)}}{k_{n(m)-1}}\,\frac{S_{k_{n(m)}}}{k_{n(m)}}, \tag{3.7}$$

where $S_m = \sum_{i=1}^m f(X_i)$. Suppose that holds

$$\bigwedge_{\alpha>1} \frac{S_{k_n}}{k_n} \xrightarrow[n\to\infty]{} Ef(X_1), \qquad \text{a.s.} \tag{3.8}$$

By this assumption and by (3.7) for a fixed $\alpha > 1$ the inequalities

$$\frac{1}{\alpha}\,Ef(X_1) \leq \frac{1}{\alpha}\liminf_{m\to\infty} \frac{S_{k_{n(m)}}}{k_{n(m)}} \leq \liminf_{m\to\infty}\frac{S_m}{m} \leq \limsup_{m\to\infty}\frac{S_m}{m} \leq \alpha \limsup_{m\to\infty}\frac{S_{k_{n(m)}}}{k_{n(m)}}$$
$$= \alpha\, Ef(X_1)$$

hold on some Ω_α with $P(\Omega_\alpha) = 1$. Therefore

$$\lim_{m\to\infty} \frac{S_m}{m} = Ef(X_1), \text{ a.s.}$$

Thus, it is sufficient to check (3.8). To start the proof of (3.8) note that

$$Ef(X_1) < \infty \iff \sum_{i=1}^{\infty} P\{f(X_1) \geq i\} < \infty \iff P\{f(X_i) \geq i \text{ i.o.}\} = 0.$$

Thus

$$\frac{S_{k_n} - ES_{k_n}}{k_n} \xrightarrow[n\to\infty]{} 0, \text{ a.s.} \iff \frac{S^c_{k_n} - ES_{k_n}}{k_n} \xrightarrow[n\to\infty]{} 0, \text{ a.s.},$$

where $S^c_m = \sum_{i=1}^m f^c(X_i)$ and $f^c(X_i) = f(X_i)I\{f(X_i) < i\}$. Note also that

$$E[f(X_i)I\{f(X_i) \geq i\}] \to 0, \quad i \to \infty,$$

hence

$$\frac{1}{n}\sum_{i=1}^{n} E[f(X_i)I\{f(X_i) \geq i\}] \xrightarrow[n\to\infty]{} 0,$$

and consequently

$$\frac{S_{k_n} - ES_{k_n}}{k_n} \xrightarrow[n\to\infty]{} 0, \text{ a.s.} \iff \frac{S^c_{k_n} - ES^c_{k_n}}{k_n} \xrightarrow[n\to\infty]{} 0, \text{ a.s.} \tag{3.9}$$

The convergence in (3.9) is equivalent to

$$\bigwedge_{\varepsilon>0} P(\limsup_{n\to\infty}\{|S^c_{k_n} - ES^c_{k_n}| > \varepsilon\, k_n\}) = 0$$

and this will follow once we show the convergence of the series

$$\sum_{n=1}^{\infty} P\{\,|S^c_{k_n} - ES^c_{k_n}| > \varepsilon\, k_n\}.$$

By Chebyshev's inequality and by Lemma 3.2 we obtain

$$\begin{aligned}
\sum_{n=1}^{\infty} P\{\,|S^c_{k_n} - ES^c_{k_n}| > \varepsilon\, k_n\} &\le \frac{1}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{Var(S^c_{k_n})}{k_n^2} \\
&\le \frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{1}{k_n^2}\sum_{i=1}^{k_n} Var(f^c(X_i)) = \frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\sum_{i=1}^{\infty}\frac{Var(f^c(X_i))}{k_n^2}\, I_{\{1,2,\dots,k_n\}}(i) \\
&= \frac{C}{\varepsilon^2}\sum_{i=1}^{\infty} Var(f^c(X_i))\sum_{n=1}^{\infty}\frac{1}{k_n^2}\, I_{\{1,2,\dots,k_n\}}(i) = \frac{C}{\varepsilon^2}\sum_{i=1}^{\infty} Var(f^c(X_i)) \sum_{\substack{n=1\\ i\le k_n}}^{\infty}\frac{1}{k_n^2}.
\end{aligned}$$

It follows that

$$\sum_{\substack{n=1\\ i\le k_n}}^{\infty}\frac{1}{k_n^2} \le \frac{C_1}{i^2}, \qquad i = 1, 2, \dots$$

with some constant $C_1 = C_1(\alpha)$. Therefore, we can write

$$\sum_{n=1}^{\infty} P\{\,|S^c_{k_n} - ES^c_{k_n}| > \varepsilon\, k_n\} \le \frac{C_2}{\varepsilon^2}\sum_{i=1}^{\infty}\frac{Var(f^c(X_i))}{i^2},$$

where $C_2 = C\cdot C_1$. However,

$$\begin{aligned}
\sum_{i=1}^{\infty}\frac{Var(f^c(X_i))}{i^2} &\le \sum_{i=1}^{\infty}\frac{E(f^c(X_i))^2}{i^2} = \sum_{i=1}^{\infty}\frac{E[f(X_1)^2 I\{f(X_1) < i\}]}{i^2} \\
&= \sum_{i=1}^{\infty}\frac{1}{i^2}\sum_{j=1}^{i} E[f(X_1)^2 I\{j-1 \le f(X_1) < j\}] \\
&= \sum_{j=1}^{\infty} E[f(X_1)^2 I\{j-1 \le f(X_1) < j\}]\sum_{i=j}^{\infty}\frac{1}{i^2} \\
&\le \sum_{j=1}^{\infty}\frac{2}{j} E[f(X_1)^2 I\{j-1 \le f(X_1) < j\}] \\
&\le 2\sum_{j=1}^{\infty} E[f(X_1) I\{j-1 \le f(X_1) < j\}] = 2Ef(X_1) < \infty,
\end{aligned}$$

and the proof is complete. ■

The above theorem admits a converse:

Proposition 3.1. *Let f be a Borel function on $\mathbb{R}$ and let*

$$\limsup_{n\to\infty}\left|\frac{S_n}{n}\right| < \infty$$

on a set with positive probability. Then $f \in L^1(\nu)$.

Proof. It suffices to show

$$E|f(X_1)| = \infty \quad \Longrightarrow \quad \limsup_{n\to\infty}\left|\frac{S_n}{n}\right| = \infty \text{ a.s.}$$

By assumption, for fixed $\alpha > 0$, we have

$$E\left|\frac{f(X_1)}{\alpha}\right| = \infty,$$

whence

$$\sum_{n=1}^{\infty} P\{|f(X_1)| \ge \alpha n\} = \infty.$$

By the Borel-Cantelli Lemma for gaussian systems (Corollary 2.1) it follows that

$$P(\limsup_{n\to\infty}\{|f(X_n)| \ge \alpha n\}) = 1.$$

Since

$$|f(X_n)| = |S_n - S_{n-1}| \ge \alpha n \quad \Rightarrow \quad |S_n| \ge \frac{\alpha n}{2} \ \vee\ |S_{n-1}| \ge \frac{\alpha n}{2} \ge \frac{\alpha(n-1)}{2},$$

we see that

$$P(\limsup_{n\to\infty}\{|S_n| \ge \alpha n/2\}) = 1$$

Thus, we have established the following

$$\bigwedge_{\alpha>0} \ \bigvee_{\substack{\Omega_\alpha\in\mathcal{F}\\ P(\Omega_\alpha)=1}} \quad \limsup_{n\to\infty}\frac{|S_n(\omega)|}{n} \ge \frac{\alpha}{2}, \qquad \omega\in\Omega_\alpha.$$

If we put $\Omega_0 = \bigcap_{m=1}^{\infty}\Omega_m$, then

$$\limsup_{n\to\infty}\frac{|S_n(\omega)|}{n} \ge \frac{m}{2}, \qquad \omega\in\Omega_0, \quad m\ge 1.$$

From this we conclude that

$$\limsup_{n\to\infty}\frac{|S_n(\omega)|}{n} = \infty \text{ a.s.}$$

and the proposition follows. ∎

Modifying slightly the proof of Theorem 3.3 we obtain the convergence of (3.1) with $f(X_n)$ replaced by $f_n(X_n)$, $f_n \in L^2(\nu)$ (see also [E2]).

Theorem 3.4. *Let the normalized gaussian sequence $(X_i, i = 1, 2, \ldots)$ satisfy the hypothesis (R) and let $f_i \in L^2(\nu)$, $i \ge 1$. Moreover, let*

$$\sup_{i\ge 1} E|f_i(X_i)| < \infty$$

and

$$\sum_{i=1}^{\infty} \frac{Var(f_i(X_i))}{i^2} < \infty.$$

Then

$$\frac{1}{n}\sum_{i=1}^{n} f_i(X_i) - Ef_i(X_i) \xrightarrow[n\to\infty]{} 0, \ a.s.$$

Proof. Since

$$\mathrm{Var}(f_i(X_i)) \geq \mathrm{Var}((f_i(X_i) - Ef_i(X_i))^+) + \mathrm{Var}((f_i(X_i) - Ef_i(X_i))^-), \qquad i \geq 1,$$

it is sufficient to prove the theorem for non-negative random variables $f_i(X_i)$. Let $S_n = \sum_{i=1}^{n} f_i(X_i)$, $\alpha > 1$ and

$$k_0 = 1, \quad k_n = [\alpha^n], \quad n \geq 1,$$

In the same way as in the proof of Theorem 3.3 we can estimate

$$\sum_{n=1}^{\infty} P\{\, |S_{k_n} - ES_{k_n}| > \varepsilon\, k_n\} \leq \frac{C_2}{\varepsilon^2} \sum_{i=1}^{\infty} \frac{Var(f_i(X_i))}{i^2},$$

for every $\varepsilon > 0$. Thus by the Borel-Cantelli lemma

$$\frac{S_{k_n} - ES_{k_n}}{k_n} \xrightarrow[n\to\infty]{} 0, \ \text{a.s.} \tag{3.10}$$

Now, for given m we have $k_{n(m)-1} \leq m \leq k_{n(m)}$, whence

$$\frac{S_m - ES_m}{m} \leq \left|\frac{S_{k_{n(m)}} - ES_{k_{n(m)}}}{k_{n(m)}}\right| \frac{k_{n(m)}}{k_{n(m)-1}} + \frac{ES_{k_{n(m)}} - ES_{k_{n(m)-1}}}{k_{n(m)-1}} \tag{3.11}$$

and

$$\frac{S_m - ES_m}{m} \geq -\left|\frac{S_{k_{n(m)-1}} - ES_{k_{n(m)-1}}}{k_{n(m)-1}}\right| - \frac{ES_{k_{n(m)}} - ES_{k_{n(m)-1}}}{k_{n(m)-1}}. \tag{3.12}$$

Using (3.11) and (3.12) we obtain

$$\limsup_{m\to\infty} \left|\frac{S_m - ES_m}{m}\right| \leq \sup_{i\geq 1} Ef_i(X_i)(\alpha - 1)$$

for every $\alpha > 1$ which concludes the proof. ∎

We will need the following theorem (for the proof see [W]).

Theorem 3.5 (Orno Theorem). *Let $\sum_{n=1}^{\infty} Y_n$ be a series of random variables (Y_n) unconditionally convergent in probability. Then $\sum_{n=1}^{\infty} \frac{Y_n}{\ln(n+1)}$ converges a.s.* ∎

Application of Orno's result gives the following version of the Strong Law of Large Numbers.

Theorem 3.6. *Let the normalized gaussian sequence $(X_i, i = 1, 2, \ldots)$ satisfy the hypothesis (R) and let $f_i \in L^2(\nu)$ for $i \geq 1$. Moreover, let*

$$\sum_{n=1}^{\infty} \frac{Var(f_n(X_n))}{n^2} \ln^2(n+1) < \infty. \tag{3.13}$$

Then

$$\frac{1}{n}\sum_{i=1}^{n} f_i(X_i) - Ef_i(X_i) \xrightarrow[n\to\infty]{} 0, \qquad a.s. \tag{3.14}$$

Proof. We see at once from (3.13) that the series

$$\sum_{n=1}^{\infty} \frac{f_n(X_n) - Ef_n(X_n)}{n} \ \ln(n+1)$$

is unconditionally convergent in probability. By Orno's theorem it follows that the series

$$\sum_{n=1}^{\infty} \frac{f_n(X_n) - Ef_n(X_n)}{n} \tag{3.15}$$

converges a.s. Applying Kronecker's lemma to (3.15) we obtain (3.14) and the proof is complete. ∎

Notice that a slight change in the proof of the classical Menchoff inequality (see [SW]) shows that for normalized gaussian sequence $(X_i, i = 1, 2, \ldots)$ satisfying the hypothesis (R) and for $f_i \in L^2(\nu),\ i \geq 1$ $(Ef_i(X_i) = 0,\ i \geq 1)$ we have

$$E(\max_{1\leq i\leq n} S_i^2) \leq C\left[\frac{\ln(4n)}{\ln 2}\right]^2 \sum_{i=1}^{n} E[f_i(X_i)]^2, \qquad n \geq 1$$

where $S_i = \sum_{j=1}^{i} f_j(X_j)$. From this (in a standard way) we obtain

Theorem 3.7 (Theorem of Rademacher-Menchoff type). *Suppose additionally that*

$$\sum_{n=1}^{\infty} (\ln n)^2 E[f_n(X_n)]^2 < \infty.$$

Then S_n converges a.s. ∎

It is easy to see that applying Theorem 3.7 and Kronecker's Lemma we obtain another proof of Theorem 3.6.

Acknowledgements. The authors are indebted to Stanisław Kwapień for his constructive remarks. In particular, he pointed out that Theorem 3.6 for $f \in L^2(\nu)$ follows from Orno's Theorem and this motivated us to prove Theorem 3.3.

References

[B] P. Billingsley, *Probability and Measure*, John Wiley & Sons, New York, 1979.

[BC] M. Beśka and Z. Ciesielski, *Ergodic type theorem for gaussian systems*, Proc. Steklov Inst. Math. 248 (2005), 40–45.

[DK] H. Dym and H. P. McKean, *Gaussian Processes, Function Theory and Inverse Spectral Problem*, Academic Press, 1976.

[E1] N. Etemadi, *An elementary proof of the strong law of large numbers.* Z. Wahrsch. Verw. Gebiete 55 (1981), 119–122.

[E2] N. Etemadi, *On the laws of large numbers for nonnegative random variables*, J. Multivar. Anal. 13 (1983) 187–193.

[G] H. Gebelein, *Das statistische Problem der Korrelation als Variations und Eigenwertproblem und sein Zusammenhang mit der Ausgleichsrechnung*, Z. Angew. Math. Mech. 21 (1941), 364–379.

[H] P. I. Hadgi, *Functions of Probability*, A. N. Moldav. SSR, Kishinev, 1971.

[R] A. Rényi, *Warscheinlichkeitsrechnung*, Deutscher Verlag Wiss., Berlin, 1962.

[S] D. W. Stroock, *Probability Theory*, Cambridge Univ. Press, 2003.

[SW] W. F. Stout, *Almost Sure Convergence*, Academic Press, New York, 1974.

[VTC] N. N. Vakhania, V. I. Tarieladze and S. A. Chobanyan, *Probability Theory on Banach Spaces*, Reidel, Dordrecht, 1987.

[W] P. Wojtaszczyk, *Banach Spaces for Analysts*, Cambridge Studies in Adv. Math. 25, Cambridge Univ. Press, 1991.