Analysis Mathematica, 4 (1978), 159—180

On the central limit theorem for lacunary trigonometric series

I. BERKES

1. Introduction

It is a well known fact that lacunary subsequences of the trigonometric system exhibit certain properties of independent random variables, e.g. the central limit theorem holds for them. A sharp result in this direction is due to ERDŐS [7], who proved that

a) *If the sequence* $\{n_k\}$ *of integers satisfies*

$$n_{k+1}/n_k \geqq 1+c_k/\sqrt{k}, \quad c_k \to \infty, \tag{1.1}$$

then the sequence $\cos 2\pi n_k x$ *obeys the central limit theorem.*

b) *For any constant* $c>0$ *there is a sequence* $\{n_k\}$ *of integers such that*

$$n_{k+1}/n_k \geqq 1+c/\sqrt{k} \tag{1.2}$$

and the sequence $\cos 2\pi n_k x$ *does not satisfy the central limit theorem.*[1]

In other words, (1.1) is the optimal growth condition for the validity of the central limit theorem.

It is natural to ask what causes this sudden change in the behaviour of $\cos 2\pi n_k x$ when we replace (1.1) by (1.2), what is the property it has under (1.1) but not necessarily under (1.2) which causes it to satisfy the central limit theorem. Of course, one can expect information on this question from Erdős' proof. This proof really gives an explanation but it depends on a fairly complicated number-theoretic fact, namely on the asymptotic enumeration of the number of solutions of a certain diophantic equation. In fact, Erdős showed that $\cos 2\pi n_k x$ obeys the central limit theorem if, for any $p\geqq 1$, the number of solutions of the equation

$$\pm n_{k_1} \pm n_{k_2} \pm \dots \pm n_{k_{2p}} = 0 \quad (1 \leqq k_1, k_2, \dots, k_{2p} \leqq N) \tag{1.3}$$

Received December 17, 1977.

[1] Erdős gave the counterexample without proof. Years later, TAKAHASHI [14] gave another counterexample with a proof.

is $\sim \frac{(2p)!}{p!} N^p$ as $N \to \infty$.[2] And he proved that under (1.1) (but not necessarily under (1.2)) this asymptotic formula for the number of solutions of (1.3) really holds. The proof of this result, however, is tedious and gives little insight into the finer details of the behaviour of $\cos 2\pi n_k x$ around the critical gap condition (1.1). One purpose of the present paper is to give a different proof of (the positive half of) Erdős' theorem which really shows what the real role of (1.1) is. This proof depends on the easily verifiable martingale property of block sums X_k of the sequence $\cos 2\pi n_k x$. Once this martingale property is established, one has to find the order of magnitude of $\sum_{k=1}^{N} E(X_k^2 | X_1, \dots, X_{k-1})$ (which is very easy even under a growth condition much weaker than (1.1)) and to verify Lindeberg's (or Ljapunov's) condition. And it will turn out that (1.1) is a sufficient condition (one out of many possible ones) for the Lindeberg condition.

The above martingale approach enables us to extend the study of the behaviour of $\cos 2\pi n_k x$ also for cases when (1.1) is not satisfied. Erdős' theorem does not state that in the abscence of (1.1) the central limit theorem never holds. In fact, we shall see that even if $\{n_k\}$ grows much more slowly than the order dictated by (1.1),[3] *the sequence* $\cos 2\pi n_k x$ *obeys the central limit theorem if* $\{n_k\}$ *satisfies a simple condition of arithmetic nature.* Using this fact, we can easily construct a large class of sequences $\{n_k\}$ growing more slowly than e^{k^ε} for any $\varepsilon > 0$ (or even sequences $n_k = O(e^{(\log k)^3})$ such that $\cos 2\pi n_k x$ satisfies the central limit theorem. (In our forthcoming paper [5] we shall exactly determine what is the "slowest" growth order of $\{n_k\}$ which still permits the validity of the central limit theorem for $\cos 2\pi n_k x$.)

An interesting conjecture of Erdős (see [7]) concerns the particular sequence $n_k = [e^{k^\beta}]$ with $\beta > 0$. It is easy to see that this sequence satisfies (1.1) for $\beta > 1/2$ but not for $\beta \leqq 1/2$. Erdős, however, conjectured that with the above n_k, $\cos 2\pi n_k x$ obeys the central limit theorem even for $0 < \beta \leqq 1/2$. Whether this is true or not we are unable to decide. An attack on this problem would be to show that the above mentioned arithmetical type condition holds for this sequence. This, however, seems to be very difficult to show; in fact, this leads to a complicated number-theoretic problem related to the "degree of transcendence" of e. In view of recent results of Baker [1] and Mahler [10] this is not hopeless to show but one would need some improvement of their results. On the contrary, it is very easy to show that for any $\beta > 0$, "almost all" sequences $n_k \sim e^{k^\beta}$ have the central limit property.

---

[2] Under the number of solutions of (1.3) we mean the number of $4p$-tuples $(k_1, k_2, \dots, k_{2p}, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_{2p})$ such that $1 \leqq k_1, k_2, \dots, k_{2p} \leqq N$, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{2p} = \pm 1$ and $\varepsilon_1 n_{k_1} + \dots + \varepsilon_{2p} n_{k_{2p}} = 0$. A similar definition applies for equations appearing later.

[3] (1.1) implies that $n_k / e^{\sqrt{k}} \to \infty$.

A further advantage of the martingale approach is that it is not confined to the central limit theorem but it yields, at no extra cost, a large class of limit theorems including almost sure invariance principles, iterated logarithm type results etc. Considering the large difficulties in the known proofs of the law of the iterated logarithm under growth conditions like (1.1) (see [15], [16]), the methodological gain is even larger here than in the case of the central limit theorem.

We mention, in conclusion, one more result which is a consequence of the connection between (1.1) and the Lindenberg condition. In fact, we shall show that though, by Erdős' theorem, under (1.2) the central limit theorem generally does not hold, it is "almost" satisfied if (1.2) is valid with a large c. More exactly, if $\{n_k\}$ satisfies (1.2) then we have

$$\overline{\lim_{N\to\infty}} \sup_t \left| P\left(\sum_{k=1}^{N} \cos 2\pi n_k x < t\sqrt{N/2}\right) - \Phi(t)\right| \leqq C_1 c^{-1/5}$$

with an absolute constant C_1. In other words, when we replace (1.1) by (1.2), the central limit theorem breaks down "continuosly".[4]

The idea of using martingale properties of block sums to prove limit theorems for certain classes of random variables is adapted from our previous papers [2], [3], [4]; the same idea was discovered independently by PHILIPP and STOUT [11]. The method seems to have a few more interesting applications to which we shall return elsewhere.

2. Results

All the theorems below are probabilistic statements concerning lacunary trigonometric series $\cos 2\pi n_k x$. The underlying probability space in all of them is the interval [0, 1] with the Borel σ-field and the Lebesgue measure. The symbols $P(...)$ and $E(...)$ will denote Lebesgue measure and integral of the set or function in brackets.

We say that a sequence $\{n_k\}$ of integers satisfies condition B_2 if there is a constant C such that the number of solutions of the equation

$$n_k \pm n_l = \nu$$

does not exceed C for any $\nu > 0$.

Theorem 1. *If the sequence $\{n_k\}$ of integers satisfies condition B_2 and*

$$n_{k+1}/n_k \geqq 1 + c_k(\log k)/k, \quad c_k \to \infty, \tag{2.1}$$

[4] An analogous phenomenon was investigated in [4] for general lacunary series $f(n_k x)$.

then $\cos 2\pi n_k x$ *obeys the central limit theorem, i.e.*

$$P\left(\sum_{k=1}^{N} \cos 2\pi n_k x < t\sqrt{N/2}\right) \to \Phi(t) \quad \textit{for all } t.$$

It is possible that condition (2.1) is superfluous in Theorem 1 and the theorem is valid under the mere assumption that $\{n_k\}$ satisfies condition B_2.

Theorem 2. *Suppose that* $\{n_k\}$ *satisfies*

$$n_{k+1}/n_k \geqq 1 + c/k^\alpha \quad (c > 0) \tag{2.2}$$

with $0<\alpha<1$ *and furthermore, for any* $v>0$, *the number of solutions of*

$$n_k \pm n_l = v \quad (1 \leqq k,\ l \leqq N)$$

is at most $C_1 N^\gamma$ *with constants* $C_1>0$, $\gamma<(1-\alpha)/2$. *Then* $\cos 2\pi n_k x$ *obeys the central limit theorem.*

The case $\gamma=0$ in the above theorem means that the sequence $\{n_k\}$ satisfies condition B_2; in this special case Theorem 2 follows from Theorem 1. Let us also remark that, by Erdős' theorem, (2.2) alone implies the central limit theorem if $\alpha<1/2$, hence our theorem gives new result only if $1/2 \leqq \alpha < 1$.

Given a concrete sequence $\{n_k\}$, the second condition of Theorem 2 is generally not easy to verify (even in the case $\gamma=0$ i.e. condition B_2). This is quite apparent in the case of Erdős' conjecture i.e. $n_k=[e^{k^\beta}]$, $0<\beta\leqq 1/2$. Let us restrict our attention to the simplest case $\beta=1/2$. To show that the sequence $n_k=[e^{\sqrt{k}}]$ satisfies the second condition of Theorem 2 it would be sufficient to prove that every interval of length 2 contains at most $C_1 N^\gamma$ numbers of the form $e^{\sqrt{k}} \pm e^{\sqrt{l}}$ $(1 \leqq l < k \leqq N)$. This, in turn, would follow if we could show that the difference of any two such sums (differences) is at least $C_2 N^{-\gamma}$, or equivalently,

$$|e^{\sqrt{k}} \pm e^{\sqrt{l}} \pm e^{\sqrt{m}} \pm e^{\sqrt{n}}| \geqq C_2 N^{-\gamma} \quad (1 \leqq k, l, m, n \leqq N) \tag{2.3}$$

except the trivial case when the left side of (2.3) is zero because its terms are pairwise equal with opposite signs. By Lindemann's classical transcendence theorem (see e.g. [12, pp. 119—120]) the sum

$$|k_1 e^{\alpha_1} + \ldots + k_r e^{\alpha_r}| \tag{2.4}$$

is never zero if $k_1, \ldots, k_r$ are nonzero integers and $\alpha_1, \ldots, \alpha_r$ are different algebraic numbers.[5] Also, there are lower estimates (see [12, p. 124]) for (2.4) if we make the restriction that the polynomials defining $\alpha_1, \ldots, \alpha_r$ are of degree $\leqq L$, their coefficients are $\leqq M$ in absolute value, and $k_1, \ldots, k_r$ are $\leqq P$ in absolute value.

[5] This trivially implies the transcendence of e as well as that of π.

Evidently, (2.3) requires a specific estimate of this sort which does not seem be known. (The best known estimate concerns the case $L=1$ and is due to MAHLER [10]. We would need an analogous result for $L=2$ with a little better lower bound.)

On the contrary, it is quite easy to construct sequences $n_k \sim e^{k^\beta}$ for any given $\beta>0$ which satisfy both conditions of Theorem 2 (in fact, they are B_2 sequences). Indeed, an easy calculation shows that if $n_k=e^{k^\beta}+O(k^r)$ with a constant $r>0$ then $\{n_k\}$ satisfies (2.2) (with $\alpha=1-\beta$). It is sufficient, therefore, to construct a B_2 sequence with $n_k=e^{k^\beta}+O(k^3)$. Now, such a sequence can be obtained by induction in the following way (see [8, p. 97]). Let $n_1=1$. Suppose that $n_1, \dots, n_{k-1}$ are already constructed, they satisfy $|n_j-e^{j^\beta}|\le 5j^3$ $(j=1, \dots, k-1)$ and all the numbers $\pm n_{j_1}\pm n_{j_2}\pm n_{j_3}\pm n_{j_4}$ $(1\le j_1, j_2, j_3, j_4\le k-1)$ are different from zero (except in the trivial case). Since the interval $[e^{k^\beta}-5k^3, e^{k^\beta}+5k^3]$ contains at least $10k^3$ integers and the number of different sums $\pm n_{j_1}\pm n_{j_2}\pm n_{j_3}$ $(1\le j_1, j_2, j_3\le k-1)$ is at most $8k^3$, we can choose n_k in the above interval such that it differs from all the numbers $\pm n_{j_1}\pm n_{j_2}\pm n_{j_3}$ $(1\le j_1, j_2, j_3\le k-1)$. Hence $|n_j-e^{j^\beta}|\le 5j^3$ and $\pm n_{j_1}\pm n_{j_2}\pm n_{j_3}\pm$ $\pm n_{j_4}\ne 0$ will continue to hold also for $j=k$ and $1\le j_1, j_2, j_3, j_4\le k$. The obtained sequence $\{n_k\}$ is therefore a B_2 sequence (in fact, all the numbers $n_k\pm n_l$ are different) and $n_k=e^{k^\beta}+O(k^3)$.

The above argument shows that if in the induction step we choose n_k from the interval $[e^{k^\beta}-5k^5, e^{k^\beta}+5k^5]$ "at random", choosing each number of the interval with equal probability, then the probability of a "wrong" choice (i.e. a choice which makes n_k equal to one of the sums $\pm n_{j_1}\pm n_{j_2}\pm n_{j_3}$ $(1\le j_1, j_2, j_3\le k-1)$ is $\le 8k^3/10k^5\le 1/k^2$. Hence, by the Borel—Cantelli lemma, there are only finitely many wrong choices with probability one. Consequently, "almost all" sequences $n_k=$ $=e^{k^\beta}+O(k^5)$ satisfy condition B_2.

It is also easy to see that if $n_k=e^{c_k(\log k)^2}+O(k^r)$ with a constant $r>0$ and $c_k\uparrow\infty$ then $\{n_k\}$ satisfies (2.1). The induction argument above shows that there is a B_2 sequence $\{n_k\}$ such that $n_k=e^{c_k(\log k)^2}+O(k^3)$. Hence, by Theorem 1, for any $c_k\uparrow\infty$ there is a sequence $n_k\sim e^{c_k(\log k)^2}$ such that $\cos 2\pi n_k x$ satisfies the central limit theorem. It is interesting to compare this growth order with condition (1.1) of Erdős' theorem which implies that n_k grows faster than $e^{\sqrt{k}}$.

Theorem 3. *Let us suppose that $\{n_k\}$ satisfies either the conditions of Theorem 2 or relation (2.2) with $\alpha<1/2$. Then the sequence $\cos 2\pi n_k x$ obeys the almost sure invariance principle, i.e. on a suitable new probability space we can find a sequence X_k of random variables having the same joint distributions as $\cos 2\pi n_k x$ and*

$$X_1+\dots+X_N = \zeta(N/2)+o(N^{1/2-\lambda}) \quad a.s.$$

with a constant $\lambda>0$ where ζ is a Wiener process on the new space.

Theorem 3 is much stronger than Theorem 2, in fact, it implies not only the central limit theorem but also the law of the iterated logarithm, its functional form, the so-called Lévy—Erdős—Feller upper-lower class test etc. for $\cos 2\pi n_k x$. We formulated (and will prove) Theorem 2 separately to show the general pattern in the proof of the central limit theorem (the proofs of Theorems 1, 2 and also of Theorem 4 below differ only in the choice of certain parameters). It is also instructive to see that after proving the martingale property of the block sums, the proof of an a.s. invariance principle does not require much more effort than the proof of the central limit theorem.

Theorem 4. *Suppose that*

$$n_{k+1}/n_k \geqq 1+c/\sqrt{k} \tag{2.5}$$

for sufficiently large k where $c>0$ is a constant. Then

$$\overline{\lim_{N\to\infty}} \sup_t \left| P\left(\sum_{k=1}^{N} \cos 2\pi n_k x < t\sqrt{N/2}\right) - \Phi(t)\right| \leqq C_1 c^{-1/5}$$

where C_1 is an absolute constant.

Theorem 4 shows that if $\{n_k\}$ satisfies (2.5) with a large c then $\cos 2\pi n_k x$ "almost" satisfies the central limit theorem.[6]

3. Martingale tools

In this section we formulate two well-known martingale results which we shall use in the proof of our theorems.

Theorem A (Heyde—Brown [9]). *Let $Y_1, Y_2, \dots$ be a martingale difference sequence with finite fourth moments, let $V_n = \sum_{i=1}^{n} E(Y_i^2 | Y_1, \dots, Y_{i-1})$ and let a_n be any sequence of positive numbers. Then*

$$\sup_t \left| P((Y_1+\dots+Y_n)/\sqrt{a_n} < t) - \Phi(t)\right| \leqq C \left\{ \frac{\sum_{i=1}^{n} EY_i^4 + E((V_n-a_n)^2)}{a_n^2} \right\}^{1/5} \tag{3.1}$$

with an absolute constant C.

[6] Of course this implies that if $\{n_k\}$ satisfies (1.1) then $\cos 2\pi n_k x$ obeys the central limit theorem (Erdős' theorem).

In [9] this theorem is stated only for $a_n = EV_n = \sum_{i=1}^{n} EY_i^2$; the proof, however, is valid for general a_n without any change. (Of course, it is customary to choose $a_n = EV_n$ to make the right-hand side of (3.1) small. But the above general version of the theorem will be useful for us because in our proofs we shall choose a_n not equal, only to be close to EV_n.)

Theorem B (STRASSEN [13]). *Let $Y_1, Y_2, \ldots$ be a martingale difference sequence with finite fourth moments, let $V_n = \sum_{i=1}^{n} E(Y_i^2 | Y_1, \ldots, Y_{i-1})$ and suppose $V_n \sim a_n$ a.s.*[7] *with some positive sequence a_n and*

$$\sum_{n=1}^{\infty} \frac{EY_n^4}{a_n^{2\vartheta}} < \infty \tag{3.2}$$

with $0 < \vartheta < 1$. Then the sequence $Y_1, Y_2, \ldots$ can be redefined on a new probability space such that its finite dimensional distributions remain the same and on the new space there is a Wiener process $\zeta(t)$ such that

$$Y_1 + \ldots + Y_n = \zeta(V_n) + o(V_n^{(1+\vartheta)/4} \log V_n) \quad a.s.$$

Observe that (3.2) and the Beppo Levi theorem imply the a.s. convergence of $\sum_{n=1}^{\infty} a_n^{-2\vartheta} E(Y_n^4 | Y_1, \ldots, Y_{n-1})$ and hence by $V_n \sim a_n$ the series $\sum_{n=1}^{\infty} V_n^{-2\vartheta} E(Y_n^4 | Y_1, \ldots, Y_{n-1})$ is also a.s. convergent. Now

$$\sum_{n=1}^{\infty} \frac{1}{V_n^{\vartheta}} \int_{x^2 > V_n^{\vartheta}} x^2 \, dP(Y_n < x | Y_1, \ldots, Y_{n-1}) \leqq \sum_{n=1}^{\infty} \frac{1}{V_n^{2\vartheta}} \int_{-\infty}^{+\infty} x^4 \, dP(Y_n < x | Y_1, \ldots, Y_{n-1}) =$$

$$= \sum_{n=1}^{\infty} \frac{1}{V_n^{2\vartheta}} E(Y_n^4 | Y_1, \ldots, Y_{n-1}) < \infty \quad \text{a.s.}$$

and thus Theorem B follows from Theorem (4.4) of [13].

The conditions of the two theorems above are very similar. To prove asymptotic normality of $(Y_1 + \ldots + Y_n)/\sqrt{a_n}$ via Theorem A, we have to verify the Ljapunov condition

$$a_n^{-2} \sum_{i=1}^{n} EY_i^4 \to 0. \tag{3.3}$$

On the other hand, the applicability of Theorem B requires proving the slightly stronger condition (3.2).[8] (In applications there will be no problem with verifying $E((V_n - a_n)^2)/a_n^2 \to 0$ or $V_n \sim a_n$ a.s.) This explains why the proof of a central limit theorem and an a.s. invariance principle for $\cos 2\pi n_k x$ are very similar.

---

[7] For two sequences c_n, d_n the symbol $c_n \sim d_n$ means $\lim_{n \to \infty} c_n/d_n = 1$.

[8] (3.2) implies (3.3) by the Kronecker lemma (if a_n is increasing).

4. The martingale property of the block sums

To handle conditional expectations, it will be convenient to work not directly with the trigonometric functions $\cos 2\pi n_k x$ but first to approximate them by step-functions $\varphi_k(x)$ as follows. Let $2^l \leqq n_k < 2^{l+1}$, put $m = [l + 8 \log k]$ and let $\varphi_k(x)$ denote the function in $[0, 1)$ which is constant in the intervals $[i2^{-m}, (i+1)2^{-m})$ $(0 \leqq i \leqq 2^m - 1)$ and these constant values coincide with the respective values of $\cos 2\pi n_k x$ at the points $i2^{-m}$ $(0 \leqq i \leqq 2^m - 1)$. Evidently $|\varphi_k(x)| \leqq 1$,

$$|\cos 2\pi n_k x - \varphi_k(x)| \leqq 2\pi n_k 2^{-m} \leqq 8\pi 2^{-8 \log k} \leqq 32 k^{-4}, \tag{4.1}$$

and thus

$$\sum_{k=1}^{\infty} |\cos 2\pi n_k x - \varphi_k(x)| < \infty \quad x \in [0, 1).$$

Hence the sequence $\cos 2\pi n_k x$ obeys the central limit theorem (or the a.s. invariance principle) if $\varphi_k(x)$ does and conversely.

Let us divide the set of positive integers into consecutive blocks $\Delta_1, \Delta_1', \Delta_2, \Delta_2', \ldots$. The length of these blocks will be specified later (in different ways in the proofs of different theorems). In all cases, however, the lengths of $\Delta_1, \Delta_2, \Delta_3, \ldots$ will tend to infinity, the same holds for $\Delta_1', \Delta_2', \ldots$ and Δ_k' will be shorter than Δ_k $(k = 1, 2, \ldots)$. For these reasons, we call the Δ_k "long blocks" and the Δ_k' "short blocks". Let $p = p(k)$ be the largest integer of Δ_{k-1}, let further $r = r(k)$ and $t = t(k)$ be the smallest and the largest integer of Δ_k, respectively. Put

$$T_k = \sum_{\nu \in \Delta_k} \cos 2\pi n_\nu x, \quad T_k' = \sum_{\nu \in \Delta_k'} \cos 2\pi n_\nu x, \quad D_k = \sum_{\nu \in \Delta_k} \varphi_\nu(x), \quad D_k' = \sum_{\nu \in \Delta_k'} \varphi_\nu(x).$$

Now we formulate our basic

Lemma (4.1). *Let us assume that for some increasing function $f(k) \leqq k$ we have*

$$n_{k+1}/n_k \geqq 1 + 1/f(k) \qquad k = 1, 2, \ldots \tag{4.2}$$

and also

$$2^{-(r-p)/f(r)} t^{11} \leqq k^{-2}, \tag{4.3}$$

$$t \leqq 2r. \tag{4.4}$$

Then we have

$$E(D_k | \mathscr{F}_{k-1}) = O(k^{-2}) \quad (0 \leqq x < 1), \tag{4.5}$$

$$E(D_k^2 | \mathscr{F}_{k-1}) = \frac{1}{2} |\Delta_k| + O(k^{-2}) \quad (0 \leqq x < 1), \tag{4.6}$$

where $\mathscr{F}_{k-1}$ denotes the σ-field generated by $D_1, \ldots, D_{k-1}$ and $|\Delta_k|$ is the number of integers contained in Δ_k. The constants in O are absolute.

Proof. Let l be an integer such that $2^l \leqq n_p < 2^{l+1}$ and put $w=[l+8\log p]$. From the definition of φ_k it is evident that every φ_ν, $1\leqq \nu\leqq p$, takes a constant value on each interval A of the form

$$A = [i2^{-w}, (i+1)2^{-w}) \quad (0 \leqq i \leqq 2^w-1) \tag{4.7}$$

and thus every set $\{D_1=a_1, \dots, D_{k-1}=a_{k-1}\}$ where $a_1, \dots, a_{k-1}$ are constants, can be obtained as a union of intervals of the form (4.7). In other words, the σ-field $\mathscr{F}_{k-1}$ is purely atomic and each of its atoms is a union of intervals of the form (4.7). Hence to prove (4.5), (4.6) it is sufficient to show that

$$|A|^{-1} \int_A D_k\, dx = O(k^{-2}), \tag{4.8}$$

$$|A|^{-1} \int_A D_k^2\, dx = \frac{1}{2}|\Delta_k| + O(k^{-2}) \tag{4.9}$$

hold (with absolute constants in O) for any A in (4.7) ($|A|$ denotes the length of A). By (4.1), $|\varphi_k|\leqq 1$, $|\Delta_k|\leqq t$, $r\geqq k$, and (4.4) we have [9]

$$|D_k-T_k| \leqq 32 \sum_{\nu\in\Delta_k} \nu^{-4} \leqq 32 \sum_{\nu=r}^{\infty} \nu^{-4} \leqq Cr^{-3} \leqq Ck^{-2}, \tag{4.10}$$

$$|D_k^2-T_k^2| \leqq \big||D_k-T_k|(|D_k|+|T_k|)\big| \leqq Cr^{-3}2|\Delta_k| \leqq Cr^{-3}2t \leqq Cr^{-2} \leqq Ck^{-2}. \tag{4.11}$$

Hence it suffices to show (4.8), (4.9) with T_k in place of D_k.

Observe that for the set A in (4.7) we have

$$|A|^{-1} \int_A T_k^\varrho\, dx = 2^w \int_{i2^{-w}}^{(i+1)2^{-w}} \Big(\sum_{\nu\in\Delta_k} \cos 2\pi n_\nu x\Big)^\varrho dx = \int_i^{i+1} \Big(\sum_{\nu\in\Delta_k} \cos m_\nu s\Big)^\varrho ds$$

for any integer $\varrho\geqq 1$; here $m_\nu = 2^{-w}2\pi n_\nu$. Also, by (4.2)

$$\frac{1}{m_r} = \frac{2^w}{2\pi n_r} \leqq \frac{2^l p^8}{n_r} \leqq \frac{n_p}{n_r} p^8, \quad m_{\nu+1}/m_\nu \geqq 1+1/f(\nu).$$

Hence, using the relations

$$\cos\gamma_1 s \cos\gamma_2 s = \frac{1}{2}[\cos(\gamma_1+\gamma_2)s + \cos(\gamma_1-\gamma_2)s], \quad \Big|\int_i^{i+1} \cos\gamma s\, ds\Big| \leqq \frac{2}{|\gamma|} \tag{4.12}$$

[9] In the sequel, C will denote positive constants, not always the same.

and $|\Delta_k| \leqq t$ we get

$$|A|^{-1} \int_A T_k\, dx = \left| \int_i^{i+1} \sum_{\nu \in \Delta_k} \cos m_\nu s\, ds \right| \leqq 2 \sum_{\nu \in \Delta_k} \frac{1}{m_\nu} \leqq$$

$$\leqq 2 \sum_{\nu \in \Delta_k} \frac{1}{m_r} = \frac{2}{m_r} |\Delta_k| \leqq 2 \frac{n_p}{n_r} p^8 t \leqq 2 \frac{n_p}{n_r} t^9,$$

$$\left| |A|^{-1} \int_A T_k^2\, dx - \frac{1}{2} |\Delta_k| \right| = \left| \int_i^{i+1} \Big(\sum_{\nu \in \Delta_k} \cos m_\nu s \Big)^2 ds - \frac{1}{2} |\Delta_k| \right| =$$

$$= \left| \frac{1}{2} \sum_{\nu \in \Delta_k} \int_i^{i+1} \cos 2m_\nu s\, ds + \sum_{\substack{\mu, \nu \in \Delta_k \\ \mu < \nu}} \int_i^{i+1} [\cos (m_\nu + m_\mu) s + \cos (m_\nu - m_\mu) s]\, ds \right| \leqq$$

$$\leqq \frac{1}{2} \sum_{\nu \in \Delta_k} \frac{2}{2m_\nu} + 2 \sum_{\substack{\mu, \nu \in \Delta_k \\ \mu < \nu}} \left(\frac{1}{m_\nu + m_\mu} + \frac{1}{m_\nu - m_\mu} \right) \leqq \frac{1}{2} \sum_{\nu \in \Delta_k} \frac{1}{m_\nu} +$$

$$+ 4 \sum_{\substack{\mu, \nu \in \Delta_k \\ \mu < \nu}} \frac{1}{m_\nu - m_\mu} \leqq \frac{1}{2} \sum_{\nu \in \Delta_k} \frac{1}{m_\nu} + 4|\Delta_k| \sum_{\nu \in \Delta_k} \frac{1}{m_{\nu+1} - m_\nu} \leqq$$

$$\leqq \frac{1}{2} |\Delta_k| \frac{1}{m_r} + 4|\Delta_k| \sum_{\nu \in \Delta_k} \frac{f(\nu)}{m_\nu} \leqq \frac{1}{2} |\Delta_k| \frac{1}{m_r} + 4|\Delta_k| f(t) \sum_{\nu \in \Delta_k} \frac{1}{m_\nu} \leqq$$

$$\leqq \frac{1}{2} |\Delta_k| \frac{1}{m_r} + 4|\Delta_k| t |\Delta_k| \frac{1}{m_r} \leqq 5|\Delta_k|^2 t \frac{1}{m_r} \leqq 5|\Delta_k|^2 t \frac{n_p}{n_r} p^8 \leqq 5t^3 \frac{n_p}{n_r} p^8 \leqq 5 \frac{n_p}{n_r} t^{11}.$$

In other terms, we have

$$|A|^{-1} \int_A T_k\, dx = O(\lambda_k) \quad \text{and} \quad |A|^{-1} \int_A T_k^2\, dx = \frac{1}{2} |\Delta_k| + O(\lambda_k)$$

where

$$\lambda_k = \frac{n_p}{n_r} t^{11}$$

and the constants in the O are absolute. To conclude the proof, it is sufficient now to remark that by (4.2) we have

$$\frac{n_r}{n_p} \geqq \prod_{\nu=p}^{r-1} \left(1 + \frac{1}{f(\nu)}\right) \geqq \left(1 + \frac{1}{f(r)}\right)^{r-p} = \left[\left(1 + \frac{1}{f(r)}\right)^{f(r)}\right]^{(r-p)/f(r)} \geqq 2^{(r-p)/f(r)},$$

and thus (4.3) implies $\lambda_k \leqq k^{-2}$.

Let

$$\bar{D}_k = D_k - E(D_k | \mathcal{F}_{k-1}).$$

Then $\bar{D}_k$ is a martingale difference sequence and

$$|\bar{D}_k - D_k| = O(k^{-2}) \tag{4.13}$$

provided that the conditions of Lemma (4.1) are satisfied. Furthermore, Lemma (4.1) trivially implies the following

Lemma (4.2). *Under the conditions of Lemma* (4.1) *we have*

$$E(\bar{D}_k^2|\bar{\mathcal{F}}_{k-1}) = \frac{1}{2}|\Delta_k| + O(k^{-2}) \quad (0 \leqq x < 1), \tag{4.14}$$

where $\bar{\mathcal{F}}_{k-1}$ denotes the σ-field generated by $\bar{D}_1, \ldots, \bar{D}_{k-1}$ and the constant in O is absolute.

Proof. Put $U_k = E(D_k|\mathcal{F}_{k-1})$, then

$$E(\bar{D}_k^2|\mathcal{F}_{k-1}) = E\big((D_k - U_k)^2|\mathcal{F}_{k-1}\big) = E(D_k^2|\mathcal{F}_{k-1}) - 2U_kE(D_k|\mathcal{F}_{k-1}) + U_k^2 =$$
$$= E(D_k^2|\mathcal{F}_{k-1}) - U_k^2,$$

since U_k is $\mathcal{F}_{k-1}$ measurable. Hence, using Lemma (4.1) we get

$$E(\bar{D}_k^2|\mathcal{F}_{k-1}) = \frac{1}{2}|\Delta_k| + O(k^{-2}).$$

Taking expected values with respect to $\bar{\mathcal{F}}_{k-1}$ and noting that $\bar{\mathcal{F}}_{k-1} \subset \mathcal{F}_{k-1}$, we get (4.14).

Remark. Lemma (4.1) expresses the almost martingale property of the long block sums D_k under the assumptions (4.2)—(4.4). The completely analogous statement holds for the short block sums D_k'. In other words, if p' denotes the largest integer of Δ_{k-1}', r' and t' denote the smallest and the largest integer of Δ_k', respectively, furthermore $\mathcal{F}_{k-1}'$ denotes the σ-field generated by $D_1', D_2', \ldots, D_{k-1}'$, then Lemma (4.1) remains formally valid if we replace p, r, t, D_k, $\mathcal{F}_{k-1}$, $|\Delta_k|$ by their "primed" versions p', r', t', D_k', $\mathcal{F}_{k-1}'$, $|\Delta_k'|$. The same holds for Lemma (4.2).

5. Estimates for fourth moments

Let $n_1 < n_2 < \ldots$ be a sequence of integers and put

$$S_{N,M} = \sum_{j=M+1}^{M+N} \cos 2\pi n_j x. \tag{5.1}$$

The purpose of the present section is to prove Lemmas (5.2) and (5.3) below which give estimates for $\int_0^1 S_{N,M}^4\,dx$ under different assumptions on the sequence $\{n_k\}$.

The first lemma is almost trivial.

Lemma (5.1). *We have (without any assumption on $\{n_k\}$)*

$$\int_0^1 S_{N,M}^4\,dx \leqq \frac{1}{2}N^3. \tag{5.2}$$

Proof. Writing $S_{N,M}^4$ as a sum of four-term products $\prod_{l=1}^{4}\cos 2\pi n_{j_l}x$ (but not collecting the terms of equal type), rewriting these products by successive applications of the identity in (4.12) and using the fact that $\int_0^1 \cos 2\pi nx\,dx=0$ or 1 according as $n\neq 0$ or $n=0$ (n is integer) we see that the left-hand side of (5.2) is equal to $1/8$ times the number of solutions of the equation

$$n_{j_1}\pm n_{j_2}\pm n_{j_3}\pm n_{j_4}=0 \quad (M+1\leqq j_1,j_2,j_3,j_4\leqq M+N). \tag{5.3}$$

Evidently there are $4N^3$ possibilities to choose n_{j_1}, n_{j_2}, n_{j_3} together with the signs of n_{j_2}, n_{j_3}. After this choice is made, there is at most one possibility for n_{j_4} and its sign. Hence the number of solutions of (5.3) is at most $4N^3$ and thus (5.2) is valid.

Lemma (5.2). *Suppose that for some* $c\geqq 1$

$$n_{k+1}/n_k \geqq 1+c/\sqrt{k} \quad k=1,2,\ldots. \tag{5.4}$$

Then if $M+N\geqq c^2$, *then we have*

$$\int_0^1 S_{N,M}^4\,dx \leqq A\left(N^2+\frac{N(M+N)}{c}\right) \tag{5.5}$$

with an absolute constant A.

Proof. We begin by remarking the simple fact that if $\{n_k\}$ satisfies (5.4) then, for any $0<a<b$, the interval $[a,b]$ contains at most $\frac{2\sqrt{L}}{c}\log\frac{b}{a}+1$ members of the sequence $n_1,\ldots,n_L$ (provided that $L\geqq c^2$). Indeed, if n_p and n_q are the smallest and the largest among $n_1,\ldots,n_L$ in the interval $[a,b]$ then $n_q/n_p\leqq b/a$; on the other hand, by (5.4),

$$\frac{n_q}{n_p}\geqq\prod_{j=p}^{q-1}\left(1+\frac{c}{\sqrt{j}}\right)\geqq\left(1+\frac{c}{\sqrt{q}}\right)^{q-p}\geqq\left(1+\frac{c}{\sqrt{L}}\right)^{q-p}\geqq\exp\left(\frac{c}{2\sqrt{L}}(q-p)\right)$$

using the fact that $1+x\geqq e^{x/2}$ for $0\leqq x\leqq 1$. The two estimates for n_q/n_p imply $\exp\left(\frac{c}{2\sqrt{L}}(q-p)\right)\leqq\frac{b}{a}$, i.e. $q-p+1\leqq\frac{2\sqrt{L}}{c}\log\frac{b}{a}+1$, as stated.

We can now easily prove (5.5). By the proof of Lemma (5.1) the left-hand side of (5.5) is equal to $1/8$ times the number of solutions of (5.3). Therefore, it suffices to show that the number of those solutions of (5.3) where $j_1 \geqq j_2 \geqq j_3 \geqq j_4$, is at most $A_1(N^2+N(M+N)/c)$ with an absolute constant A_1 (the same being valid for any prescribed order of j_1, j_2, j_3, j_4). Now, if $j_1=j_2$ then necessarily $j_3=j_4$; the number of such solutions is evidently $\leqq 8N^2$. It remains to enumerate those solutions wh ee $j_1>j_2 \geqq j_3 \geqq j_4$; we show that the number of such solutions is at most $A_2 N(M+N)/c$ with an absolute constant A_2.

Let $k \geqq 0$ be a fixed integer and enumerate those among the solutions above where $2^k \leqq n_{j_2}/n_{j_3} < 2^{k+1}$. Evidently there are N possibilities to choose n_{j_1}. The assumption made on n_{j_2}/n_{j_3} implies that $n_{j_2}+n_{j_3}+n_{j_4} \leqq (1+2\cdot 2^{-k})n_{j_2}$ and thus (5.3) can be valid only if $n_{j_2}(1+2^{-k+1}) \geqq n_{j_1}$; in other words, n_{j_2} has to lie in the interval $[(1+2^{-k+1})^{-1}n_{j_1}, n_{j_1}]$. Since $j_2 \leqq M+N$, by the remark at the beginning of the proof there are at most $\frac{2\sqrt{M+N}}{c}\log(1+2^{-k+1})+1 \leqq \frac{2\sqrt{M+N}}{c}2^{-k+1}+1$ possibilities to choose n_{j_2} from the mentioned interval. In particular, if $\frac{2\sqrt{M+N}}{c}2^{-k+1}<1$ then there is at most one possibility for n_{j_2} i.e. $n_{j_2}=n_{j_1}$ which case is excluded. Hence $\frac{2\sqrt{M+N}}{c}2^{-k+1} \geqq 1$ i.e. $k \leqq 2\log(4\sqrt{M+N}/c) \leqq \leqq \log 16(M+N)$. Using again $2^k \leqq n_{j_2}/n_{j_3} < 2^{k+1}$ we see that n_{j_3} lies in the interval $[2^{-k-1}n_{j_2}, n_{j_2}]$, hence applying the remark at the beginning of the proof once more we get that after choosing n_{j_2}, there are at most $\frac{2\sqrt{M+N}}{c}\log 2^{k+1}+1 \leqq \leqq \frac{2\sqrt{M+N}}{c}(k+1)+1 \leqq \frac{6\sqrt{M+N}}{c}k$ possibilities to choose n_{j_3}. Finally, if n_{j_1}, n_{j_2}, n_{j_3} are already chosen, there are at most four possibilities for n_{j_4}. Hence the number of those solutions of (5.3) where $j_1>j_2 \geqq j_3 \geqq j_4$ and $2^k \leqq n_{j_2}/n_{j_3} < 2^{k+1}$ is at most

(5.6)
$$N\left(\frac{2\sqrt{M+N}}{c}2^{-k+1}+1\right)\frac{6\sqrt{M+N}}{c}4k = \frac{96N(M+N)}{c^2}k2^{-k}+\frac{24N\sqrt{M+N}}{c}k.$$

As we noticed, $k \leqq \log 16(M+N)$ so summing the numbers (5.6) for $0 \leqq k \leqq \leqq \log 16(M+N)$ and using the convergence of $\sum_{k=0}^{\infty} k2^{-k}$, $c \geqq 1$, and $\log 16(M+N) \leqq \leqq 8(M+N)^{1/4}$ we get that the number of solutions of (5.3) where $j_1>j_2 \geqq j_3 \geqq j_4$ is really $\leqq A_2 N(M+N)/c$ with an absolute constant A_2.

Lemma (5.3). *Suppose that, for any integer ν, the number of solutions of the equation*
$$n_k \pm n_l = \nu \quad (M+1 \leqq k,\ l \leqq M+N)$$

is at most B. Then

$$\int_0^1 S_{N,M}^4\,dx \leqq B^2N^2.$$

Proof. By (4.12) we have

$$S_{N,M}^2 = \frac{1}{2}N+\frac{1}{2}\sum_{j=M+1}^{M+N}\cos 4\pi n_j x+\sum_{M+1\leqq l<k\leqq M+N}\cos 2\pi(n_k+n_l)x+\cos 2\pi(n_k-n_l)x.$$

Collecting the terms with equal frequencies and using the assumption of the lemma we get that

$$S_{N,M}^2 = \frac{1}{2}N+\sum_{\nu} c_\nu\cos 2\pi\nu x \tag{5.7}$$

where the sum contains at most N^2 terms and $|c_\nu|\leqq B$ for all ν. Squaring (5.7) and integrating we get

$$\int_0^1 S_{N,M}^4\,dx = \frac{1}{4}N^2+\frac{1}{2}\sum_\nu c_\nu^2 \leqq \frac{1}{4}N^2+\frac{1}{2}N^2B^2 \leqq N^2B^2,$$

which was to be proved.

6. Conclusion of the proofs

Using the preparatory remarks of paragraphs 3—5, it is easy to carry out the proofs of our central limit theorems. All we have to do is to choose the length of the blocks Δ_k and Δ_k' properly and to apply a martingale central limit theorem (e.g. Theorem A of § 3) for the centered block sums $\bar{D}_k$. This gives immediately the central limit theorem for the sequence $\cos 2\pi n_k x$ along the certain subsequence of indices; to get the result for all the indices requires only an application of the Chebisev inequality.

To begin with the last point, let us remark the obvious fact that if for the distribution function $F(x)$ of a random variable ξ we have $|F(x)-\Phi(x)|\leqq\varepsilon$ for all x and η is a random variable such that $P(|\eta|\geqq\delta)\leqq\delta$, then for the distribution function $G(x)$ of $\xi+\eta$ we have $|G(x)-\Phi(x)|\leqq\varepsilon+2\delta$ for all x (we use here the fact $|\Phi(x)-\Phi(y)|\leqq|x-y|$). A special case is the following

Lemma (6.1). *Let* $S_k=\sum_{j=1}^{k}\cos 2\pi n_j x$, $\sigma=1/\sqrt{2}$, *and* $M>N$ *integers,* $M/N=\lambda$. *Suppose* $\lambda\leqq 2$. *Then*

$$\left|P\left(S_N/\sigma\sqrt{N}<x\right)-\Phi(x)\right|\leqq\varepsilon \quad \textit{for all} \quad x$$

implies

$$\left|P\left(S_M/\sigma\sqrt{M}<x\right)-\Phi(x)\right|\leqq\varepsilon+4\sqrt[4]{\lambda-1} \quad \textit{for all} \quad x.$$

Proof. The decomposition $N^{-1/2}S_N - M^{-1/2}S_M = N^{-1/2}(S_N - S_M) + S_M(N^{-1/2} - M^{-1/2})$, the orthogonality of the trigonometric system and Minkowski's inequality show that the L_2 norm of $|S_N/\sigma\sqrt{N} - S_M/\sigma\sqrt{M}|$ is at most $2\sqrt{\lambda-1}$. Hence, by Chebyshev's inequality, this random variable is $\leqq 2\sqrt[4]{\lambda-1}$ except on a set of probability $\leqq 2\sqrt[4]{\lambda-1}$ and thus our lemma follows from the preceding remark.

An immediate consequence of Lemma (6.1) is that if the distribution of $S_N/\sigma\sqrt{N}$ converges to the normal distribution as N runs through the index sequence $N=N_k$ where $N_{k+1}/N_k \to 1$ then the same holds for the whole sequence. Also, if $S_N/\sigma\sqrt{N}$ "almost" converges to the normal distribution along the index sequence $N=N_k$ where N_{k+1}/N_k "almost" tends to 1, then the same holds for the whole sequence.

Proof of Theorem 2. As we mentioned after Theorem 2, in the case $\alpha<1/2$ condition (2.2) alone implies the central limit theorem, no matter what the arithmetic structure of $\{n_k\}$ is. This is a consequence of Erdős' theorem but we shall prove this statement too because it requires no extra effort and it will show clearly why the value $\alpha=1/2$ is critical. So in what follows let us assume that one of the following two conditions is satisfied:

a) $\{n_k\}$ satisfies the conditions of Theorem 2,
b) $\{n_k\}$ satisfies (2.2) with $\alpha<1/2$.

Let us choose

$$|\Delta_k| = [k^s] \quad \text{and} \quad |\Delta_k'| = [k^{s(1-\varepsilon)}],$$

where s is larger than but close to $\alpha/(1-\alpha)$ and $\varepsilon>0$ is small enough. Then in Lemma (4.1) we have $f(k)=k^\alpha/c$, $p\sim r\sim t\sim Ck^{s+1}$ $(C=1/(s+1))$, $r-p\sim k^{s(1-\varepsilon)}$ and thus the left-hand side of (4.3) is at most

$$C_1 2^{-C_2 k^{s(1-\varepsilon)-(s+1)\alpha}} k^{11s+11} \tag{6.1}$$

with positive constants C_1, C_2. By $s>\alpha/(1-\alpha)$ we have $s-(s+1)\alpha>0$, hence the exponent $s(1-\varepsilon)-(s+1)\alpha$ in (6.1) is positive for small ε and thus (4.3) is satisfied for large k. Also, $r\sim t$ shows that (4.4) is valid for large k. Thus Lemmas (4.1), (4.2) apply and consequently (4.5), (4.6), (4.14) are valid. Let us now apply Theorem A of § 3 for the martingale difference sequence $\bar{D}_k$ with $a_n = \frac{1}{2}\sum_{k=1}^{n}([k^s] + [k^{s(1-\varepsilon)}])$; this requires an estimate for the Ljapunov quantity

$$a_n^{-2} \sum_{k=1}^{n} E(\bar{D}_k^4) \tag{6.2}$$

and also for

(6.3) $$a_n^{-2}E(|V_n-a_n|^2),$$

where $V_n=\sum_{k=1}^{n}E(\bar{D}_k^2|\bar{D}_1, \dots, \bar{D}_{k-1})$. If we assume condition a) above, then Lemma (5.3) gives $E(T_k^4)\leqq Ck^{2s}(k^{s+1})^{2\gamma}$ (since $t\leqq 2k^{s+1}$); if condition b) (i.e. $\alpha<1/2$) is assumed then we simply apply the trivial Lemma (5.1) to get $E(T_k^4)\leqq Ck^{3s}$. Since $|\bar{D}_k-T_k|\leqq|\bar{D}_k-D_k|+|D_k-T_k|=O(k^{-2})=O(1)$ by (4.13), (4.10) the same estimates are valid (by Minkowski's inequality) for $E(\bar{D}_k^4)$. Hence in the two cases we get, respectively, that the sum in (6.2) is $\leqq Cn^{2s+(s+1)2\gamma+1}$ or $\leqq Cn^{3s+1}$; on the other hand, $a_n^2\sim Cn^{2s+2}$. In case a) the assumption $\gamma<(1-\alpha)/2$ shows that if s is close enough to $\alpha/(1-\alpha)$ then $(s+1)2\gamma<1$ holds and thus in this case (6.2) tends to 0. If $\alpha<1/2$ then $\alpha/(1-\alpha)<1$ so if s is close enough to $\alpha/(1-\alpha)$ then $s<1$. In this case $3s+1<2s+2$ so the expression (6.2) tends to zero again. Note also that $V_n=\frac{1}{2}\sum_{k=1}^{n}[k^s]+O(1)=a_n+O(n^{s(1-\varepsilon)+1})$ by Lemma (4.2), thus the expression (6.3) also tends to 0. Hence, Theorem A of § 3 implies that the distribution of the first of the following three expressions

(6.4) $$\frac{\bar{D}_1+\dots+\bar{D}_n}{\sqrt{a_n}},\quad \frac{T_1+\dots+T_n}{\sqrt{a_n}},\quad \frac{T_1+T_1'+\dots+T_n+T_n'}{\sqrt{a_n}}$$

tends to the normal distribution if $n\to\infty$. Evidently the difference of the first two and also of the last two expressions in (6.4) tend to 0 in probability (this follows from $|\bar{D}_k-T_k|=O(k^{-2})$ and the fact that, by the orthogonality of the trigonometric system, the square integral of $(T_1'+\dots+T_n')/\sqrt{a_n}$ is $\frac{1}{2}\sum_{k=1}^{n}[k^{s(1-\varepsilon)}]/a_n=o(1)$). Hence the asymptotic normality of the first expression in (6.4) implies that of the third one. (See the remark before Lemma (6.1).) By the definition of T_k, T_k', the asymptotic normality of the third expression in (6.4) simply means that the distribution of $S_N/\sigma\sqrt{N}$ $\left(S_N=\sum_{j=1}^{N}\cos 2\pi n_j x,\ \sigma=1/\sqrt{2}\right)$ converges to the normal distribution as N runs through the index sequence $N=N_k$ where $N_k=\sum_{i=1}^{k}[i^s]+[i^{s(1-\varepsilon)}]$. But $N_{k+1}/N_k\to 1$ and thus the remark after the proof of Lemma (6.1) shows that $S_N/\sigma\sqrt{N}$ is asymptotically normal as $N\to\infty$.

Proof of Theorem 4. We may assume, without loss of generality, that in (2.5) we have $c\geqq 1$. Let us choose

$$|\Delta_k|=k^2 \quad\text{and}\quad |\Delta_k'|=[k^{3/2+\delta}],$$

where $0<\delta<1/2$. Then, in Lemma (4.1), we have $f(k)=\sqrt{k}/c$, $p\sim r\sim t\sim \frac{1}{3}k^3$, $r-p\sim k^{3/2+\delta}$ and thus the left-hand side of (4.3) is at most $C_1 2^{-C_2 k^\delta} k^{33}$ showing that (4.3) is valid for large k. Also, by $r\sim t$, (4.4) is satisfied for $k\geqq k_0$. By Lemma (5.2) we have $E(T_k^4)\leqq A(k^4+c^{-1}k^5)$ for $k\geqq k_1$ and since $|\bar{D}_k-T_k|=O(1)$ by (4.10), (4.13) the same estimate holds for $E(\bar{D}_k^4)$ (by Minkowski's inequality). Putting $a_n=\frac{1}{2}\sum_{k=1}^{n}(k^2+[k^{3/2+\delta}])\sim\frac{1}{6}n^3$ we see that the Ljapunov quantity (6.2) is less than A_1/c for large n with an absolute constant A_1. Observe also that $V_n=$ $=\sum_{k=1}^{n}E(\bar{D}_k^2|\bar{D}_1,\dots,\bar{D}_{k-1})=\frac{1}{2}\sum_{k=1}^{n}k^2+O(1)=a_n+O(n^{3/2+\delta+1})$ by Lemma (4.2) and thus the expression in (6.3) tends to 0. Applying Theorem A of § 3 for the sequence $\bar{D}_k$ with the a_n above we get that the distribution function $F_1(x)$ of the first expression in (6.4) satisfies

$$|F_1(x)-\Phi(x)|\leqq C_3c^{-1/5}\quad\text{for all } x \tag{6.5}$$

if $n\geqq n_0$; C_3 is an absolute constant. As in the proof of Theorem 2, the difference of the first and the last expression in (6.4) tends to 0 in probability, so if $F_3(x)$ denotes the distribution function of the last expression in (6.4) then (6.5) and the remark preceding Lemma (6.1) imply

$$|F_3(x)-\Phi(x)|\leqq 2C_3c^{-1/5}\quad\text{for all } x, \tag{6.6}$$

for $n\geqq n_1$. (6.6) says that

$$\left|P(S_{N_k}/\sigma\sqrt{N_k}<x)-\Phi(x)\right|\leqq 2C_3c^{-1/5}\quad\text{for all } x, \tag{6.7}$$

for $k\geqq k_0$, where $N_k=\sum_{i=1}^{k}(i^2+[i^{3/2+\delta}])$, $\sigma=1/\sqrt{2}$. Now $N_{k+1}/N_k\to 1$, (6.7) and Lemma (6.1) imply

$$\left|P(S_N/\sigma\sqrt{N}<x)-\Phi(x)\right|\leqq 3C_3c^{-1/5}\quad\text{for all } x,$$

for $N\geqq N_0$, and thus Theorem 4 is proved.

Proof of Theorem 1. We can evidently assume, without loss of generality, that $k/c_k\log k$ is monotone increasing. (Indeed, since $c_k\to\infty$ we can find a sequence $c_k'\to\infty$, $c_k'\leqq c_k$ such that $k/c_k'\log k$ is monotone increasing.) Let us choose

$$|\Delta_k|=[\theta^k]\quad\text{and}\quad|\Delta_k'|=[\varepsilon_k\theta^k],$$

where $1<\theta<2$ is a fixed number and $\varepsilon_k\to 0$ sufficiently slowly. In this case for the quantities appearing in Lemma (4.1) we have $f(k)=k/c_k\log k$, $p=\sum_{j=1}^{k-1}[\theta^j]+$ $+\sum_{j=1}^{k-2}[\varepsilon_j\theta^j]\sim\frac{1}{\theta-1}\theta^k$, similarly $r\sim\frac{1}{\theta-1}\theta^k$, $t\sim\frac{1}{\theta-1}\theta^{k+1}$, $r-p\geqq\frac{1}{2}\varepsilon_{k-1}\theta^{k-1}$

and thus the left-hand side of (4.3) is at most

(6.8) $C_4 2^{-\frac{1}{2}\varepsilon_{k-1}\theta^{k-1}c_r \log r/r}\theta^{11k+11} \leqq C_4 2^{-C_5\varepsilon_{k-1}c_r k}\theta^{11k+11} \leqq C_6 2^{-C_5\varepsilon_{k-1}c_r k+C_7 k}$

for sufficiently large k, where C_4, C_5, C_6, C_7 are constants depending on θ. Here $c_r = c_{r(k)} \geqq c_{[\theta^{k-1}]} \to \infty$, thus if $\varepsilon_k \to 0$ slowly enough then $\varepsilon_{k-1}c_r \to \infty$. Hence the last expression in (6.8) is $O(2^{-C_8 k})$, showing that (4.3) is satisfied for $k \geqq k_0$. Also, $t/r \to \theta < 2$ and hence (4.4) is valid for $k \geqq k_1$. Since $\{n_k\}$ satisfies condition B_2, Lemma (5.3) gives $E(T_k^4) \leqq C_0\theta^{2k}$ (C_0 is a constant depending only on $\{n_k\}$) and since $|\bar{D}_k - T_k| = O(1)$ by (4.10), (4.13) the same estimate holds for $E(\bar{D}_k^4)$. Putting $a_n = \frac{1}{2}\sum_{k=1}^{n}([\theta^k]+[\varepsilon_k\theta^k]) \sim \frac{1}{2(\theta-1)}\theta^{n+1}$ we see that the Ljapunov quantity in (6.2) does not exceed

$$5C_0 \frac{\sum_{k=1}^{n}\theta^{2k}}{(\theta-1)^{-2}\theta^{2n+2}} \leqq 5C_0 \frac{(\theta^2-1)^{-1}\theta^{2n+2}}{(\theta-1)^{-2}\theta^{2n+2}} \leqq 5C_0(\theta-1)$$

for $n \geqq n_0(\theta)$. Also, $V_n = \sum_{k=1}^{n} E(\bar{D}_k^2 | \bar{D}_1, \ldots, \bar{D}_{k-1}) = \frac{1}{2}\sum_{k=1}^{n}[\theta^k] + O(1) = a_n(1+o(1))$ uniformly by Lemma (4.2) and thus the expression in (6.3) tends to 0. Applying Theorem A of § 3 for the sequence $\bar{D}_k$ we get that the distribution function $F_1(x)$ of the first expression in (6.4) satisfies

(6.9) $$|F_1(x) - \Phi(x)| \leqq C^*(\theta-1)^{1/5} \quad \text{for all } x,$$

for $n \geqq n_1(\theta)$; C^* depends only on $\{n_k\}$. As in the previous proofs, the first and third expressions in (6.4) differ only by a random variable tending to 0 in probability, so for the distribution function $F_3(x)$ of the third expression in (6.4) we get, using (6.9) and the remark before Lemma (6.1),

(6.10) $$|F_3(x) - \Phi(x)| \leqq 2C^*(\theta-1)^{1/5} \quad \text{for all } x,$$

for $n \geqq n_2(\theta)$. (6.10) says that

(6.11) $$\left|P(S_{N_k}/\sigma\sqrt{N_k} < x) - \Phi(x)\right| \leqq 2C^*(\theta-1)^{1/5} \quad \text{for all } x,$$

for $k \geqq k_0$, where $N_k = \sum_{j=1}^{k}([\theta^j]+[\varepsilon_j\theta^j])$, $\sigma = 1/\sqrt{2}$. Evidently $N_{k+1}/N_k \to \theta$, thus $N_{k+1}/N_k < \theta + \varepsilon$ for $k \geqq k_1$, and hence (6.11) and Lemma (6.1) imply

$$\left|P(S_N/\sigma\sqrt{N} < x) - \Phi(x)\right| \leqq 2C^*(\theta-1)^{1/5} + 4(\theta+\varepsilon-1)^{1/4} \quad \text{for all } x,$$

for $N \geqq N_0(\theta, \varepsilon)$. Since $\theta - 1$ and ε can be chosen arbitrary small, the last relation shows that the distribution of $S_N/\sigma\sqrt{N}$ tends to the normal distribution.

7. Proof of the a.s. invariance principle

The only difference in proving a central limit theorem and an almost sure invariance principle for $\cos 2\pi n_k x$ is that after choosing the length of Δ_k and Δ'_k properly, we use Theorem B of § 3 instead of Theorem A and we have to be a little more careful in comparing the second and third expressions in (6.4) because instead of the convergence in probability to 0 of their difference we need a.s. convergence.

Proof of Theorem 3. Let us choose $|\Delta_k|$, $|\Delta'_k|$ in the same way as in the proof of Theorem 2. The estimates showing the convergence to 0 of the expression (6.2) in the proof of Theorem 2 show actually that

$$\sum_{k=1}^{\infty} \frac{E(\bar{D}_k^4)}{b_k^{2-\delta}} \tag{7.1}$$

is convergent if δ is small enough (the general term being $O(k^{-(1+\lambda)})$ with a small $\lambda>0$). Here $b_k=\frac{1}{2}\sum_{i=1}^{k}[i^s]\sim Ck^{s+1}$. Also, if $V_n=\sum_{k=1}^{n}E(\bar{D}_k^2|\bar{D}_1, \dots, \bar{D}_{k-1})$, then $V_n=b_n+O(1)$ by Lemma (4.2) (the conditions of Lemma (4.2) were shown to be satisfied in the proof of Theorem 2). Hence Theorem B of § 3 implies that there is a Wiener process $\zeta(t)$ such that

$$\bar{D}_1+\dots+\bar{D}_k = \zeta(V_k)+o(V_k^{1/2-\eta}) \quad \text{a.s.} \tag{7.2}$$

with a constant $\eta>0$.[10] Replacing $\bar{D}_1+\dots+\bar{D}_k$ by $T_1+\dots+T_k$ on the left-hand side of (7.1) we commit an error $O(1)$ (since $|\bar{D}_k-T_k|=O(k^{-2})$); also, $V_k=b_k+ +O(1)$, $b_k\sim Ck^{s+1}$, and well known properties of the Wiener process (see e.g. Lemma (3.6) of [3]) show that $\zeta(V_k)=\zeta(b_k)+O(k^{1/3})=\zeta(b_k)+o(b_k^{1/3})$ a.s. Hence (7.2) implies

$$T_1+\dots+T_k = \zeta(b_k)+o(b_k^{1/2-\eta}) \quad \text{a.s.}^{11} \tag{7.3}$$

We also remark that replacing the left-hand side of (7.3) by $T_1+T'_1+\dots+T_k+T'_k$ we only add a term which is $o(b_k^{1/2-\eta})$ so it does not bother the right side of (7.3). [Indeed, (7.3) has the exact analogue

$$T'_1+\dots+T'_k = \zeta(d_k)+o(d_k^{1/2-\eta}) \quad \text{a.s.} \tag{7.4}$$

[10] Strictly speaking, we have to redefine the whole sequence $\bar{D}_k$ on a new probability space to have representation (7.2). But since the formulas are the same on the new and the old space, we shall speak (with a little inaccuracy) as if ζ were defined on the original space.

[11] In what follows, η will denote positive constants, not necessarily the same

for the short block sums (the proof is the same);[12] here $d_k=\frac{1}{2}\sum_{i=1}^{k}[i^{s(1-\varepsilon)}]$. Now it is sufficient to use the estimate $\zeta(t)=o(t^{1/2}\log t)$ a.s. on the right side of (7.4) and remark that $d_k\leqq b_k^{1-\theta}$ with a constant $\theta>0$.] Hence (7.3) implies

(7.5) $$T_1+T_1'+\ldots+T_k+T_k'=\zeta(b_k)+o(b_k^{1/2-\eta})\quad\text{a.s.}$$

If $c_k=b_k+d_k$ then $c_k-b_k=O(b_k^{1-\theta})$ with a constant $\theta>0$ and thus $\zeta(b_k)=\zeta(c_k)+ +o(c_k^{1/2-\eta})$ a.s. by Lemma (3.6) of [3]. Hence (7.5) implies

$$T_1+T_1'+\ldots+T_k+T_k'=\zeta(c_k)+o(c_k^{1/2-\eta})\quad\text{a.s.}$$

This latter relation simply says that

(7.6) $$S_N=\zeta(N/2)+o(N^{1/2-\eta})\quad\text{a.s.}$$

is valid if $N=N_k$, where $N_k=\sum_{i=1}^{k}([i^s]+[i^{s(1-\varepsilon)}])$. To get (7.6) for all N we only have to show that

(7.7) $$\max_{N_k\leqq N\leqq N_{k+1}}|S_N-S_{N_k}|=o(N_k^{1/2-\eta})\quad\text{a.s.,}$$

(7.8) $$\max_{N_k\leqq N\leqq N_{k+1}}|\zeta(N/2)-\zeta(N_k/2)|=o(N_k^{1/2-\eta})\quad\text{a.s.}$$

For any $N_k\leqq M<N\leqq N_{k+1}$ we have, with a constant C,

(7.9) $$E(|S_N-S_M|^4)\leqq\begin{cases}CN_{k+1}^{2\gamma}(N-M)^2,\\ \frac{1}{2}(N-M)^3,\end{cases}$$

according as we assume the conditions of Theorem 2 or (2.2) with $\alpha<1/2$ (see Lemmas (5.3) and (5.1)). Using the Markov inequality and Theorem 12.2 of [6] we get that in the two cases, respectively, the probability

(7.10) $$P(\max_{N_k\leqq N\leqq N_{k+1}}|S_N-S_{N_k}|\geqq t)$$

(t is an arbitrary positive number) is at most $C_1t^{-4}N_{k+1}^{2\gamma}(N_{k+1}-N_k)^2$ or $C_1t^{-4}(N_{k+1}-N_k)^3$ with an other constant C_1. Since $N_{k+1}\sim C_2k^{s+1}$, $N_{k+1}-N_k\sim k^s$, the expressions standing after C_1t^{-4} above are exactly of the same order of magnitude as were the upper estimates for $E(\bar{D}_k^4)$ in the proof of Theorem 2 (which upper estimates were used to show the convergence of (7.1)). Hence choosing $t=b_k^{(2-\delta)/4}$ with a small $\delta>0$, the proof of the convergence of (7.1) yields automatically that the expression (7.10) (with the t above) is the general term of a convergent series. By the Borel—Cantelli lemma this implies that, with probability one, the left side

[12] See the remark at the end of § 4.

of (7.7) is $O(b_k^{(2-\delta)/4})=O(N_k^{(2-\delta)/4})=O(N_k^{1/2-\delta/4})$ $\left(\text{since } b_k \sim \frac{1}{2} N_k\right)$ and thus (7.7) is valid. Relation (7.8) can be proved in the same way, using

$$E(|\zeta(N/2)-\zeta(M/2)|^4) \leqq C_3(N-M)^2$$

instead of (7.9).

References

[1] A. BAKER, On some diophantine inequalities involving the exponential function, *Canad. J. Math.*, **17** (1965), 616—626.

[2] I. BERKES, Approximation of lacunary Walsh series with Brownian motion, *Studia Sci. Math. Hungar.*, **9** (1974), 111—122.

[3] I. BERKES, On the asymptotic behaviour of $\Sigma f(n_k x)$. I. Main theorems, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, **34** (1976), 319—345.

[4] I. BERKES, On the asymptotic behaviour of $\Sigma f(n_k x)$. II. Applications, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, **34** (1976), 347—365.

[5] I. BERKES, A central limit theorem for trigonometric series with small gaps, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* (to appear).

[6] P. BILLINGSLEY, *Convergence of probability measures*, Wiley (New York, 1968).

[7] P. ERDŐS, On trigonometric sums with gaps, *Magyar Tud. Akad. Mat. Kut. Int. Közl.*, **7** (1962), 37—42.

[8] H. HALBERSTAM and K. F. ROTH, *Sequences*. I, Clarendon Press (Oxford, 1966).

[9] C. C. HEYDE and B. M. BROWN, On the departure from normality of a certain class of martingales, *Ann. Math. Statist* **41** (1970), 2161—2165.

[10] K. MAHLER, On a paper by A. Baker on the approximation of rational powers of e, *Acta Arith.*, **27** (1975), 61—87.

[11] W. PHILIPP and W. F. STOUT, Almost sure invariance principles for sums of weakly dependent random variables, *Mem. Amer. Math. Soc.*, No. **161**.

[12] T. SCHNEIDER, *Einführung in die Transcendenten Zahlen*, Springer (Berlin, 1957).

[13] V. STRASSEN, Almost sure behaviour of sums of independent random variables and martingales, *Proc. 5th Berkeley Sympos. Math. Statist. Probab.*, Vol. II, Part 1; 315—343 (Univ. of California Press, 1967).

[14] S. TAKAHASHI, On lacunary trigonometric series. II, *Proc. Japan Acad.*, **44** (1968), 766—770.

[15] S. TAKAHASHI, On the law of the iterated logarithm for lacunary trigonometric series, *Tôhoku Math. J.*, **24** (1972), 319—329.

[16] S. TAKAHASHI, On the law of the iterated logarithm for lacunary trigonometric series. II, *Tôhoku Math. J.*, **27** (1975), 391—403.

О центральной предельной теореме для лакунарных тригонометрических рядов

И. БЕРКЕШ

Хорошо известно, что вероятностное поведение лакунарного тригонометрического ряда $\{\cos 2\pi n_k x\}$ тесно связано с «критическим» условием лакунарности

$$(*) \qquad \frac{n_{k+1}}{n_k} \geqq 1+\frac{c_k}{\sqrt{k}}, \quad c_k \to \infty.$$

Например, если выполнено условие $(*)$, то последовательность $\{\cos 2\pi n_k x\}$ удовлетворяет центральной предельной теореме, и при этом условие $(*)$ не может быть ослаблено. Для последовательностей, удовлетворяющих $(*)$, известны и другие результаты подобного рода, в то время как для более медленно растущих последовательностей $\{n_k\}$ не известно, по-видимому, ничего. В статье развит метод, который при помощи мартингальной техники позволяет проводить исследование систем $\{\cos 2\pi n_k x\}$ для последовательностей, не удовлетворяющих условию $(*)$. Получено простое объяснение условия $(*)$, изучено, как «пропадает» центральная предельная теорема при постепенном ослаблении условия $(*)$ и доказаны некоторые центральные предельные теоремы в отсутствие этого условия. Получены другие предельные теоремы для $\{\cos 2\pi n_k x\}$, например, закон повторного логарифма и принципы инвариантности.

I. BERKES
MATHEMATICAL INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES
REÁLTANODA U. 13—15
1053 BUDAPEST, HUNGARY