Journal of Mathematical Sciences, Vol. 114, No. 4, 2003

SUMMATION OF DIVERGENT SERIES AND ERGODIC THEOREMS

V. V. Kozlov* UDC 519.21

In this article questions on the possibility of sharpening classic ergodic theorems is considered. To sharpen these theorems the author uses methods of summation of divergent sequences and series. The main topic is connected with the individual ergodic Birkhoff–Khinchin theorem. The theorem is studied in connection with the Riesz and Voronoi summation methods. These methods are weaker than those of the Cesaro method of arithmetic means. It is shown that already for the Bernoulli transformation of the unit interval, meaningful problems arise. These problems are interesting in connection with the possibility of extension of the strong law of large numbers. The questions of suitable summation factors and of the solution of homological equations by means of divergent series is also discussed.

1 RIESZ AND VORONOI AVERAGES

Let M be a space with a finite measure μ ($\mu(M) < \infty$), and let $T\colon M \to M$ be a measure preserving (possibly, non-invertible) mapping, $\mu(T^{-1}A) = \mu(A)$, for any measurable domain A. Consider a measurable function $f\colon M \to \mathbb{R}$. Taking into account the *Poincaré recurrence theorem*, it can be claimed that the sequence $f(T^n x)$, $n = 1, 2, \dots$ is divergent, in general. However, according to the *Birkhoff–Khinchin theorem*, if $f \in L_1$, then for almost all $x \in M$, the sequence $f(T^n x)$ converges in the sense of Cesaro to an invariant (with respect to the transformation T) integrable function $\bar f(x)$, so that

$$\int_M f\,d\mu = \int_M \bar f\,d\mu. \tag{1.1}$$

This notable result had been preceded by the *von Neumann statistical ergodic theorem*, which claims convergence in the mean

$$f(T^n(x)) \to \bar f(x) \quad \text{(C)} \tag{1.2}$$

(the label (C) indicates convergence in the sense of Cesaro). Let L_2 be the linear space of square summable functions f on M. To each function f on M we can associate the function $g(x) = f(Tx)$, $x \in M$. It is well known (see, for instance, [1]) that the mapping $f \to g = Uf$ determines an isometric operator U in L_2. The von Neumann theorem claims that

$$U^n f \to Pf \quad \text{(C)},$$

where P is the operator of projection of L_2 to the space of all vectors invariant with respect to U.

Historically, however, the first ergodic theorem should be credited to Bole–Serpinsky–Weyl [2]. Here M is the n-dimensional torus $\mathbb{T}^n = \{x_1, \dots, x_n \bmod 2\pi\}$ with the standard measure, and T is the shift operator

$$x_s \mapsto x_s + \omega_s, \quad 1 \leqslant s \leqslant n, \quad \omega_s = \text{const}. \tag{1.3}$$

It turns out that if the function $f\colon \mathbb{T}^n \to \mathbb{R}$ is integrable in the sense of Riemann, then the limit relation (1.2) holds for all $x \in \mathbb{T}^n$; moreover, the limit function $\bar f$ is also integrable in the sense of Riemann and, of course, (1.1) holds.

Translated from Trudy Seminara imeni I. G. Petrovskogo, No. 22, pp. 142–168, 2002.

*This work was supported by the Russian Foundation for Basic Research (grant 99-01-0196) and grant “Leading Scientific Schools” (00-15-96146).

Usually, one considers ergodic shifts (1.3) with $\omega_1, \dots, \omega_n, 1$ being incommensurable with respect to rational numbers. Then $\bar{f}$ does not depend x and

$$\bar{f} = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{T}^n} f(x)\, d^n x. \tag{1.4}$$

The general case can be easily reduced to the analysis of this special case.

Convergence in the sense of Cesaro is not the only method used for the summation of divergent sequences and series. This is a highly advanced branch of analysis (see, for instance, the classical book by Hardy [3]). Usually, one resorts to summation methods which are *linear, regular*, and *natural*. The most common are the *matrix methods* [3, 4].

Let $p_0 > 0$, $p_n \geqslant 0$, and $\sum p_n = \infty$. By definition, Riesz averages of a sequence s_n $(n = 0, 1, \dots)$ are ratios of the form

$$\frac{p_0 s_0 + \dots + p_n s_n}{p_0 + \dots + p_n}.$$

If $p_n = p_0$ for all n, we obtain the Cesaro average (arithmetic mean). The sequence p_n specifies the Riesz summation method [3]: if $t_n \to s$ for $n \to \infty$, then, by definition,

$$s_n \to s \quad (\mathrm{R}, p_n).$$

This method is regular: if $s_n \to s$, then $t_n \to s$.

The Riesz summation method is also defined for convergent series $\sum p_n$. In this case, however, there is no regularity property.

The possibility of strengthening the Bole–Serpinsky–Weyl theorem by replacing the Cesaro method with a more general Riesz method was indicated by Hermann Weyl in his classical work on the uniform distribution [2]. Weyl proves this theorem for the Riesz method (R, p_n) in the following two cases:

1) p_n monotonically decreases;

2) p_n monotonically increases with n, so that

$$(n+1)p_n = o\biggl(\sum_{s=0}^{n} p_s\biggr). \tag{1.5}$$

In fact, this statement gives us nothing new, since under the above assumptions on the sequence p_n, the following property holds: if $s_n \to s$ (C), then $s_n \to s$ (R, p_n). In other words, in this situation the Riesz method *includes* the Cesaro method. On the other hand, for $p_{n+1} \geqslant p_n$, the Cesaro method includes the Riesz method [3, Theorem 14]. If in addition, condition (1.5) is satisfied, then the two methods are equivalent.

It is interesting to observe that the general theorem about the inclusion of the Riesz summation methods was proved by Cesaro way back in 1888, whereas the work of Weyl was published only about thirty years later, in 1916. This fact remained unnoticed for many authors (see, for instance, the well known book by Pólya and Szegö; Problem 173 in Sec. 2).

Note that if (1.5) holds, then numbers p_n may grow as powers of n. On the other hand, if p_n grow exponentially, then the Riesz method fails and becomes equivalent to the usual convergence [3, Theorem 15]. Therefore, of real interest are the cases with $p_n \sim \exp n^\alpha$, $0 < \alpha < 1$.

Apart from the Riesz convergence, we will use convergence in the sense of Voronoi [3]. Again, let $q_0 > 0$, $q_n \geqslant 0$. For a sequence $\{s_n\}_0^\infty$, let

$$u_n = \frac{q_n s_0 + q_{n-1} s_1 + \dots + q_0 s_n}{q_0 + q_1 + \dots + q_n}.$$

Clearly, if $s_i = s_0$, then $u_n = s_0$. For $u_n \to s$, one writes $s_n \to s$ (W, q_n). A detailed exposition of the theory of summation in the sense of Voronoi is given in [3].

The criterion of regularity of the W-method is as follows:

$$\frac{q_n}{q_0 + \dots + q_n} \to 0.$$

It turns out that any two regular Voronoi methods are *compatible*: if $s_n \to s$ (W) and $s_n \to s'$ (W$'$), then $s = s'$.

If the W-method is regular and the sequence q_n is non-decreasing, then the (W, q_n) includes the Cesaro method. The inverse inclusion is a very interesting problem. Let $q_0 = 1$, $q_n > 0$. If the W-method is regular and q_n is non-decreasing, and

$$\frac{q_{n+1}}{q_n} \geqslant \frac{q_n}{q_{n-1}}$$

for all $n > 0$, then C includes (W, q_n).

Theorem 1. *Let $f\colon \mathbb{T}^n \to \mathbb{R}$ be a Riemann integrable function and let T be the shift transformation defined by (1.3) with rationally incommensurable numbers $\omega_1, \dots, \omega_n, 1$.*

(a) *If $q_{n+1} \leqslant q_n$ and $\sum q_n = \infty$, then*

$$f(T^n x) \to \bar{f} \quad (\mathrm{W}, q_n).$$

(b) *If $p_{n+1} \geqslant p_n$ and $p_n \Big/ \sum\limits_0^n p_s \to 0$, then*

$$f(T^n x) \to \bar{f} \quad (\mathrm{R}, p_n).$$

Here $\bar{f}$ is the average (1.4). These two statements are actually equivalent and can be easily proved by the Weyl method.

In fact, Theorem 1 has been known in the theory of uniform distribution of sequences mod 1 (see [6,7]). It is presumed [8] that an extension of the Weyl theory by means of the Riesz summation methods was initially obtained by Tsuji [9]. In connection with Theorem 1, there is an open problem: *describe all matrix methods of summation for which the Bole–Serpinsky–Weyl theorem holds*. In [10], some new conditions of uniform distribution are given for convergence in the sense of Riesz and Voronoi.

Under the assumptions (a) and (b) of Theorem 1, a statistical ergodic theorem can be established for the Riesz and the Voronoi summation methods.

Theorem 2. *Let U be an isometric operator in a complex Hilbert space, and let P be the operator of projection of that space onto the subspace of all vectors invariant with respect to U.*

(a) *If $q_{n+1} \leqslant q_n$ and $\sum q_n = \infty$,then for all f, we have*

$$U^n f \to Pf \quad (\mathrm{W}, q_n). \tag{1.6}$$

(b) *If $p_{n+1} \geqslant p_n$ and $p_n \Big/ \sum\limits_0^n \to 0$, then for all f, we have*

$$U^n f \to Pf \quad (\mathrm{R}, p_n).$$

For $q_n = 1$ or $p_n = 1$, we obtain the classical von Neumann theorem. Theorem 2 claims that for $n \to \infty$, the functions $f(T^n x)$ converge on the average to an invariant function of the transformation T in a stronger sense than in the von Neumann theorem.

Theorem 2 can be obtained from the results of paper [11] dealing with more general linear methods of summation specified in terms of an infinite matrix with non-negative elements $a_{n,k}$ $(n, k = 0, 1, 2, \dots)$ such that

$$\sum_{k=0}^{\infty} a_{n,k} = 1$$

for all n. In [11], the following criterion is given for the mean values

$$\sum_{k=0}^{\infty} a_{n,k} U^k f$$

to be convergent in the mean to Pf:

(1) $\lim\limits_{n\to\infty}\sum\limits_{k=0}^{\infty} a_{n,k\alpha+j} = 1/\alpha$ for all $0 \leqslant j \leqslant \alpha$ and $\alpha = 2, 3, 4, \dots,$

(2) for any irrational γ, the sequence $\{n\gamma\}$ is uniformly distributed with respect to this summation method.

Note that the second condition is nonconstructive: as mentioned above, a description of all matrix methods for which the sequences $\{n\gamma\}$ are uniformly distributed on the interval $[0, 1]$ is as yet an open problem (see [8]).

In our case, property (1) follows from the monotonicity of weight coefficients q_n (respectively, p_n) under the additional condition that $\sum q_n = \infty$ (respectively, $p_n \Big/ \sum\limits_{0}^{n} p_s \to 0$). Condition (2) follows from Theorem 1. Paper [11] also contains sufficient conditions on the coefficients $a_{n,k}$ ensuring the statistical ergodic theorem for transformations with weak or strong mixing. We give a simple proof of Theorem 2 by the Riesz method [1]. This proof makes clear the meaning of the conditions (a) and (b). For the sake of definiteness, consider the Voronoi average.

For $Uf = f$, the statement (1.6) is obvious. Let $f = g - Ug$ for some G. Then

$$\|q_0U^n f + \ldots + q_{n-1}Uf + q_n f\| = \| - q_0U^{n+1}g + (q_0 - q_1)U^n g + \ldots + (q_{n-1} - q_n)Ug + q_n g\| \leqslant 2q_0\|g\|,$$

since U is isometric and therefore, $\|U^k g\| = \|g\|$. Since $\sum q_n = \infty$, the Voronoi averages under consideration tend to zero.

It turns out [1] that each element of a Hilbert space can be represented as a sum $f_1 + f_2$, where $Uf_1 = f_1$ and f_2 belongs to the closure of the linear manifold that consists of elements of the form $g - Ug$. It remains to show that $W_n f \to 0$ for each f from the closure of the said manifold. Here,

$$W_n = \frac{q_0U^n + \ldots + q_nU^0}{q_0 + \ldots + q_n}$$

is the Voronoi average. First, we note that $\|W_n f\| \leqslant \|f\|$. Further, let $f_k \to f$, and if $W_n f_k \to 0$ as $n \to \infty$ for all k, then $W_n f \to 0$. Indeed,

$$\|W_n f\| \leqslant \|W_n(f - f_k)\| + \|W_n f_k\| \leqslant \|f - f_k\| + \|W_n f_k\|.$$

Let us take k such that $\|f - f_k\|$ is small. Then we choose n so as to make $\|W_n f_k\|$ small, and our statement is proved.

Theorem 2 remains meaningful if the Hilbert space is replaced by a complex space. In this case, the isometric operator is represented by a complex number z with $|z| = 1$. If $z = 1$, then the Voronoi averages are obviously equal to 1. For $z \neq 1$, Theorem 2 claims that

$$z^n \to 0 \quad (\mathrm{W}, q_n).$$

As an example, consider the case of $q_n = (n + 1)^{-1}$. Clearly, we have

$$W_n(z) = \frac{z^{n+1}\left[z^{-1} + \frac{z^{-2}}{2} + \ldots + \frac{z^{-n-1}}{(n+1)}\right]}{\ln n + O(1)}.$$

Since $z^{-1} \neq 1$, it follows that for $n \to \infty$ the expression in square brackets tends to $\ln(1-z^{-1})$. This is a single-valued continuous function on the circle with a deleted point, $\{z\colon\ |z| = 1 \text{ and } z \neq 1\}$. Consequently, $W_n(z) \to 0$ as $1/\ln n$.

Results of another type pertaining to generalization of the von Neumann theorem for weighted means can be found in [12].

2 STRONG LAW OF LARGE NUMBERS

Is it possible to strengthen the Birkhoff–Khinchin pointwise ergodic theorem by replacing summation in the sense of Cesaro with a weaker summation method of Riesz or Voronoi? As shown in Sec. 1, the answer is positive in the case of convergence in the mean (Theorem 2). However, the answer may happen to be negative for pointwise convergence in a typical situation. Nevertheless, in the case of shifts of a torus, the Cesaro method can be replaced by a weaker method of Riesz or Voronoi (Theorem 1).

In order to have a better understanding of the difficulties arising in this situation, consider more closely the case of M being the unit interval $[0,1]$ and T being the *Bernoulli transformation*

$$x \to 2x \bmod 1.$$

This transformation is irreversible, but it preserves the common measure on $[0,1]$: $\operatorname{meas} L = \operatorname{meas}(T^{-1}L)$, where L is an arbitrary measurable set on $[0,1]$ and $T^{-1}L$ is the total inverse image of L. Since the Bernoulli transformation is ergodic, it follows that for any integrable function f and almost all x, we have

$$f(T^n x) \to \int_0^1 f(x)\,dx \quad \text{(C)}. \tag{2.1}$$

Let us define a piecewise constant function f as follows: $f(x) = 1$ for $0 \leqslant x < 1/2$, and $f(x) = -1$ for $1/2 \leqslant x \leqslant 1$. The functions $r_{n+1}(x) = f(T^n x)$, $n \geqslant 0$, are called *Rademacher functions*. These form an orthonormal system and, in particular, have zero mean value and unit variance. Relation (2.1) turns into the *strong law of large numbers* (in the sense of Borel [13]):

$$r_n(x) \to 0 \quad \text{(C)} \tag{2.2}$$

for almost all x.

Now, let (W, q_n) be a Voronoi summation method. If

$$\frac{q_0^2 + \ldots + q_n^2}{(q_0 + \ldots + q_n)^2} \to 0, \tag{2.3}$$

then

$$\lim_{n\to\infty} \operatorname{meas}\left\{\frac{q_n r_1(x) + \ldots + q_0 r_{n+1}(x)}{q_0 + \ldots + q_n} > \varepsilon\right\} = 0. \tag{2.4}$$

This is a generalization of the *weak law of large numbers* for Bernoulli trials. Relation (2.4) is obtained directly from (2.3) with the help of the Chebyshev inequality.

Using the Cauchy–Stolz theorem, one can show that if the sequence q_n is non-increasing and $\sum q_n = \infty$ (condition (a) of Theorems 1 and 2), then (2.3) holds. This result can be expected since convergence in the mean implies convergence with respect to the measure (2.4).

Theorem 3. *Suppose that for some integer $k > 0$ we have*

$$\sum_{n+1}^{\infty} \left[\frac{q_0^2 + \ldots + q_n^2}{(q_0 + \ldots + q_n)^2}\right]^k < \infty. \tag{2.5}$$

Then

$$r_n(x) \to 0 \quad (\mathrm{W}, q_n) \tag{2.6}$$

for almost all x.

The proof of this statement is based on the Lewy theorem: for a sequence of non-negative integrable functions $\xi_n(x)$, the convergence of

$$\sum \int_0^1 \xi_n(x)\,dx$$

implies the convergence of the series

$$\sum \xi_n(x)$$

almost everywhere. In particular, $\xi_n(x) \to 0$ for almost all x.

Let us introduce the random variables

$$\eta_n(x) = \frac{q_n r_1(x) + \ldots + q_0 r_{n+1}(x)}{q_0 + \ldots + q_n}$$

and first obtain formulas for their moments

$$m_k(n) = \int_0^1 \eta_n^k(x)\,dx.$$

Clearly, all m_k with odd k are equal to zero. Further, letting

$$\sigma_s = q_0^s + \ldots + q_n^s, \tag{2.7}$$

we consecutively obtain

$$\begin{aligned} m_2 &= \frac{\sigma_2}{\sigma_1^2}, \\ m_4 &= \frac{3\sigma_2^2 - 2\sigma_4}{\sigma_1^4}, \\ m_6 &= \frac{15\sigma_2^3 - 30\sigma_4\sigma_2 + 16\sigma_6}{\sigma_1^6}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{2.8}$$

It is easy to see that $\sigma_1^{2k} m_{2k}$ are symmetric polynomials of $q_0^2, \ldots, q_n^2$. According to a known algebraic result, these polynomials can be represented as polynomials of sums of powers of (2.7).

Using the obvious inequalities

$$\sigma_{2k} \leqslant \sigma_2^k, \quad k = 1, 2, \ldots,$$

we can show that the series

$$\sum_{n=1}^{\infty} m_k(n)$$

is convergent, provided that the series (2.5) is convergent. Now, taking $\xi_n = \eta_n^k$, we obtain the statement of Theorem 3.

The statement (2.6) — *the strong law of large numbers* — is much more meaningful than (2.4). Note that condition (2.3) follows from (2.5). The converse, however, is not true, as shown by the following simple example: $q_n = (n+1)^{-1}$. In this case, the ratio (2.3) decays with the rate of $\ln^{-2} n$, but all series (2.5) are divergent. Problem: does (2.6) hold if $q_n = (n+1)^{-1}$?

3 KOLMOGOROV CRITERION FOR RIESZ AVERAGES

With obvious modifications, all statements of Sec. 2 can be formulated for the summation in the sense of Riesz. However, for the Riesz averages, some further results can be proved. In contrast to Sec. 2, the random variables considered here are of a more general type.

Assume that the sequence p_n, $n \geqslant 0$, which determines the Riesz summation method, satisfies the following conditions:

$$p_{n+1} \geqslant p_n, \quad n = 0, 1, 2, \ldots \tag{3.1}$$

and

$$\frac{p_n}{p_0 + \ldots + p_n} \to 0 \tag{3.2}$$

as $n \to \infty$ (condition (b) of Theorems 1 and 2). Recall that condition (3.1) means that the method (R, p_n) is not stronger than that of Cesaro.

Proposition 1. *Let* $\xi_0, \xi_1, \xi_2 \dots$ *be identically distributed independent random variables with zero mean values. Then, under the conditions* (3.1) *and* (3.2), *the weak law of large numbers holds: for any* $\varepsilon > 0$,

$$\mathrm{P}\left\{\left|\sum_0^n p_s\xi_s \Big/ \sum_0^n p_s\right| < \varepsilon\right\} \to 0,$$

as $n \to \infty$.

This statement is proved by the standard argument involving characteristic functions. It should be stressed that there is no assumption about the existence of variance of the random variables ξ_s (in contrast to Sec. 2).

For matrix summation methods, some authors estimate the decay rate of the probability from Proposition 1 with the help of properties of the distribution function for the random variable ξ (see, for instance, [14, 15]).

Theorem 4. *Let* $\xi_0, \xi_1, \dots$ *be identically distributed independent random variables with zero mean values and finite moments of orders* $\leqslant 2k$, *and let* $p_0, p_1, \dots$ *satisfy the inequality*

$$\sum_{n=1}^{\infty}\left[\frac{p_0^2 + \ldots + p_n^2}{(p_0 + \ldots + p_n)^2}\right]^k < \infty. \tag{3.3}$$

Then

$$\mathrm{P}\{\xi_n \to 0\ (\mathrm{R}, p_n)\} = 1.$$

Corollary 1. *Suppose that* ξ_n *have finite moments of all orders. Then*

$$\xi_n \to 0 \quad (\mathrm{R}, \exp(n+1)^\alpha) \quad a.\ s.$$

for all $0 \leqslant \alpha < 1$.

In particular, this result holds for the Rademacher functions from Sec. 2.

In order to prove the above corollary, let $p_n = \exp(n+1)^\alpha$. For $k = 2$, condition (3.3) holds for $0 \leqslant \alpha < 1/2$. For $k = 3$, the series (3.3) is convergent on a larger interval, $0 \leqslant \alpha < 2/3$, etc.

Theorem 4 is proved in the same way as Theorem 3, with formulas (2.8) replaced by more general formulas

$$\begin{aligned}
&\sigma_1^2 m_2 = \mu_2\sigma_2,\\
&\sigma_1^4 m_4 = 3\mu_2^2\sigma_2^2 + (\mu_4 - 3\mu_2^2)\sigma_4,\\
&\sigma_1^6 m_6 = 15\mu_2^3\sigma_2^3 + 15(\mu_2\mu_4 - 3\mu_2^3)\sigma_2\sigma_4 + 10\mu_3^2\sigma_3^2 + (\mu_6 + 10\mu_3^2 + 30\mu_2^3 - 15\mu_2\mu_4)\sigma_6,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}$$

Here, μ_k is the moment of kth-order of the random variable ξ_k, $\sigma_s = p_0^s + \ldots + p_n^s$.

Remark 1. Now, let

$$p_{n-1} = \exp\frac{n}{\ln^\alpha n}, \quad \alpha > 0. \tag{3.4}$$

It is easy to check that conditions (3.1) and (3.2) are satisfied. The expressions

$$\sum p_k^2 \Big/ \Big(\sum p_k\Big)^2$$

asymptotically decay with $n \to \infty$, and the decay rate is like that of $\ln^{-\alpha} n$. Consequently, the series (3.3) are divergent for all k, and therefore Theorem 4 cannot be applied, although the weak law of large numbers takes place (cf. Sec. 2).

The strong law of large numbers can be approached in another way if one resorts to the ideas used in the proof of the classical *Kolmogorov criterion* [16, 17]: for independent random variables $\xi_0, \xi_1, \dots$ with zero mean values and variances $\sigma_0^2, \sigma_1^2, \dots$ such that

$$\sum \frac{\sigma_s^2}{s^2} < \infty, \tag{3.5}$$

we have

$$\mathrm{P}\{\xi_n \to 0\ (\mathrm{C})\} = 1.$$

This result can be extended to the Riesz averages.

Kronecker lemma. *Suppose that the series*

$$\sum_{j=0}^{\infty} \frac{p_j x_j}{p_0 + \ldots + p_j}$$

is convergent, $p_j > 0$*, and* $\sum p_j = \infty$*. Then*

$$x_n \to 0 \quad (\mathrm{R}, p_n)$$

as $n \to \infty$*.*

Condition $\sum p_j = \infty$ *means that the Riesz summation method* (R, p_n) *is regular. Note that usually the Kronecker lemma is formulated in a different way (see, for instance,* [17, Chap. VII]*).*

Theorem 5. *Let* $\xi_0, \xi_1, \ldots$ *be independent random variables with zero mean values and variances* $\sigma_0^2, \sigma_1^2, \ldots$*. If*

$$\sum \frac{p_j^2 \sigma_j^2}{(p_0 + \ldots + p_j)^2} < \infty, \tag{3.6}$$

then

$$\xi_n \to 0 \quad (\mathrm{R}, p_n) \quad \text{a. s.}$$

Indeed, according to (3.6), the sum of independent random variables

$$\frac{p_j \xi_j}{p_0 + \ldots + p_j}$$

is bounded. Therefore, by the Khinchin–Kolmogorov theorem, the series

$$\sum_{j=0}^{\infty} \frac{p_j \xi_j}{(p_0 + \ldots + p_j)}$$

is convergent almost surely. Applying the Kronecker lemma to this series, we obtain the desired result.

Clearly, for $p_0 = p_1 = \ldots$, condition (3.6) turns into the Kolmogorov condition (3.5).

Let us apply Theorem 5 to the case of identically distributed random variables. Then condition (3.6) becomes

$$\sum \frac{p_n^2}{(p_0 + \ldots + p_n)^2} < \infty. \tag{3.7}$$

For instance, let $p_n = \exp(n+1)^\alpha$, $0 \leqslant \alpha$. In this case, the series (3.7) is convergent only for $\alpha < 1/2$. This result is weaker than that of Theorem 4. However, it should be kept in mind that Theorem 5 has been established under weaker assumptions on the random variables ξ_n.

To conclude this section, we give an example demonstrating the advantages of an extension of summability in the sense of Cesaro. In fact, this example was indicated by Kolmogorov [16] (see also [18, Chap. IX]) in order to demonstrate that condition (3.5) cannot be weakened. Consider a sequence of independent random variables $\xi_1, \xi_2, \ldots$ with zero expectations and variances $\sigma_n^2 = n$. These variables can be chosen such that

$$\mathrm{P} = \{\xi_n = n\} = \mathrm{P}\{\xi_n = -n\} = \frac{1}{2n}, \quad \mathrm{P}\{\xi_n = 0\} = 1 - \frac{1}{n}.$$

Clearly, the series (3.5) is divergent for this sequence and the strong law of large numbers does not hold in this case.

On the other hand, since the series

$$\sum \frac{\sigma_n^2}{n^2 \ln^2 n}$$

is convergent, it follows from Theorem 5 that

$$\mathrm{P}\left\{\xi_n \to 0 \ \left(\mathrm{R}, \frac{1}{n}\right)\right\} = 1.$$

Note that the Riesz method $(\mathrm{R}, 1/n)$ is certainly stronger than that of Cesaro.

4 APPLICATION OF THE ITERATED LOGARITHM LAW

In order to prove the strong law of large numbers, one can use the *Khinchin–Kolmogorov iterated logarithm law*. To that end, we introduce new random variables $\eta_n = p_n\xi_n$. Clearly, $\eta_1, \eta_2, \dots$ are independent random variables with zero mean values.

Let σ_1^2, σ_2^2 be the variances of $\xi_1, \xi_2, \dots$. Then the variances of η_k will be equal to $p_k^2\sigma_k^2, \dots$. Set

$$B_n = \sum_1^n p_k^2\sigma_k^2, \quad S_n = \sum_1^n \eta_k = \sum_1^n p_k\xi_k.$$

According to [19], if

$$B_n \to \infty, \tag{4.1}$$

$$|\eta_n| \leqslant M_n = o\left(\sqrt{\frac{B_n}{\ln\ln B_n}}\right), \tag{4.2}$$

then

$$\lim_{n\to\infty} \sup \frac{|S_n|}{\sqrt{2B_n \ln\ln B_n}} = 1 \quad \text{a. s.} \tag{4.3}$$

From (4.3), it follows, in particular, that

$$S_n = (\sqrt{B_n \ln\ln B_n}) \quad \text{a. s.}$$

Therefore, if

$$\frac{B_n}{\left(\sum\limits_1^n p_k\right)^2} = \frac{\varphi(n)}{\ln\ln B_n}, \tag{4.4}$$

where $\varphi(n) \to 0$ as $n \to \infty$, then

$$\xi_n \to 0 \quad (\mathrm{R}, p_n) \quad \text{a. s.}$$

However, in this case, condition (4.4) should be compatible with conditions (4.1) and (4.2). Condition (4.4) is stronger than the convergence

$$\frac{\sum p_k^2\sigma_k^2}{\left(\sum p_k\right)^2} \to 0,$$

which ensures the weak law of large numbers. This can be expected, since convergence almost everywhere implies convergence with respect to the measure.

The compatibility of (4.4) with the assumptions (4.1) and (4.2) imposes additional restrictions on the random variables ξ_n. The following result holds.

Theorem 6. *Let $\xi_1, \xi_2, \dots$ be independent random variables with zero expectations and identical variances. Suppose that $|\xi_n| \leqslant M$ (a. s.), $M = \mathrm{const}$, and the numbers p_n are such that*

$$\frac{p_n^2}{\sum\limits_1^n p_k^2} = \frac{\psi(n)}{\ln\ln \sum\limits_1^n p_k^2}, \tag{4.5}$$

$\psi(n) \to 0$ as $n \to \infty$. Then

$$\xi_n \to 0 \quad (\mathrm{R}, p_n) \quad \text{a. s.} \tag{4.6}$$

Proof. Since the variances of ξ_n are the same, we have $B_n \to \infty$, according to (3.1). Thus, condition (4.1) ensuring the applicability of the iterated logarithm law is satisfied. Further, since $\eta_n = p_n \xi_n$ and $|\xi_n| \leqslant M$, we have

$$|\eta_n| \leqslant MP_n \leqslant M\left(\frac{B_n \psi(n)}{\ln\ln B_n}\right)^{\frac{1}{2}}, \tag{4.7}$$

in view of (4.5). Since $\psi(n) \to 0$ as $n \to \infty$, it obviously follows that the right-hand side of (4.7) is

$$o\left(\sqrt{\frac{B_n}{\ln\ln B_n}}\right).$$

Therefore, condition (4.2) holds.

Set

$$\frac{\sum_1^n p_k^2}{\left(\sum_1^n p_k\right)^2} = \frac{\varphi(n)}{\ln\ln \sum_1^n p_k^2}. \tag{4.8}$$

By (3.1) we have

$$\frac{p_n \sum_1^n p_k}{\sum_1^n p_k^2} \geqslant 1.$$

This, together with (4.5) and (4.8), implies

$$\varphi(n) \leqslant \psi(n). \tag{4.9}$$

In particular, $\varphi(n) \to 0$ as $n \to \infty$. As a result, we obtain the strong law of large numbers (4.6), which completes the proof.

At first sight, conditions (4.5) and (4.4) (if the variances are equal) seem different. Moreover, for $p_{n+1} \geqslant p_n$, relation (4.5) always implies (4.4). However, in many practically important cases, these conditions are actually equivalent.

For example, let $p_n = f(n)$, where $f(x)$ monotonically increases to ∞ and is a smooth function of the Hardy class with respect to the exponent (see, for instance, [20, Chap. V]). In particular, one can take as $f(x)$ the function

$$\exp x^\gamma \ (0 < \gamma < 1) \quad \text{or} \quad \exp \frac{x}{\ln^\alpha x} \ (\alpha > 0).$$

It can be shown that under the condition (3.2), we have

$$\lim_{n\to\infty} \frac{\varphi(n)}{\psi(n)} = \frac{1}{4}$$

(cf. (4.9)). In particular, if $p_n = \exp n^\gamma$, $\gamma < 1$, then the functions φ and ψ are decreasing as

$$\frac{\ln n}{n^{1-\gamma}}.$$

Hence, we again obtain Corollary 1 from Theorem 4.

Now, set

$$p_n = \exp \frac{n}{\ln^\alpha n}, \quad \alpha > 0. \tag{4.10}$$

Using the Euler–Maclaurin summation formula, one can easily check that

$$\frac{\sum p_k^2}{\left(\sum p_k^2\right)^2}$$

asymptotically decays as n and the decay rate is like that of $1/\ln^{\alpha} n$. Consequently, the series (3.3) are divergent for all k, and therefore Theorem 4 is inapplicable.

However, in this case, we have

$$\psi(n) \sim \ln^{1-\alpha} n.$$

Thus, if $\alpha > 1$, then (by Theorem 6) the strong law of large numbers holds for coefficients p_n of the form (4.10).

It should be mentioned that Theorem 6 imposes stronger restrictions on the random variables ξ_n than Theorem 4.

In connection with this result, there is a question whether the limit relation

$$r_n(x) \to 0 \quad \left(\mathrm{R}, \exp \frac{n}{\ln n}\right) \quad \text{a. s.}$$

holds for the Rademacher functions. It should be emphasized that the sequence $p_n = \exp(n/\ln n)$, of course, satisfies conditions (3.1) and (3.2): the ratios

$$\frac{p_n}{\sum\limits_{1}^{n} p_k}$$

tend to zero as $1/\ln n$ with the growth of n.

Remark 2. In order to obtain conditions ensuring the strong law of large numbers, we could have applied the Abel transformation to the sum $p_1\xi_1 + \ldots + p_n\xi_n$ and then utilized the ordinary Khinchin iterated logarithm law, which establishes the asymptotic behavior of sums of identically distributed random variables $\xi_1 + \ldots + \xi_n$. However, this approach results in a weaker statement than that of Theorem 6. Thus, we would get

$$\mathrm{P}\{\xi_n \to 0 \ (\mathrm{R}, \exp n^{\alpha})\} = 1$$

only if $0 \leqslant \alpha < 1/2$.

Thus, the results of Secs. 2–4 demonstrate serious doubts regarding the possibility of strengthening the Birkhoff–Khinchin theorem by replacing the Cesaro convergence with a weaker method of Riesz or Voronoi satisfying conditions (a) or (b) of Theorems 1, 2, respectively. However, there is still hope to prove convergence

$$f(T^n x) \to \bar{f}(x) \quad (\mathrm{R}, \exp n^{\alpha})$$

for almost all x and some $\alpha > 0$. Clarification of this point is an essential problem of the ergodic theory.

A similar problem was considered by Baxter [21, 22] (see also [23]) for the Riesz method (R, p_n), with the weight coefficients p_n determined by the recurrent relation

$$p_n = f_1 p_{n-1} + \ldots + f_n p_0, \quad p_0 = 1,$$

where f_k $(k \geqslant 1)$ is a sequence of non-negative numbers for which $\sum f_k = 1$. Clearly, all p_n belong to the same segment $[0, 1]$ and p_n tend to a finite limit as $n \to \infty$. Under certain additional conditions, a piecewise ergodic theorem has been established in [21, 22] for the (R, p_n)-method. In particular, for $f_1 = 1$ and $f_k = 0$ $(k \geqslant 2)$, we obtain the usual Birkhoff–Khinchin theorem.

For p_n tending to a finite limit, the (R, p_n)-method may happen to be equivalent to the usual Cesaro method (this possibility has not been excluded by Baxter). Therefore, of greatest interest (according to [22]) is the case of $p_n \to 0$. However if, in addition, the decay of p_n is monotone, then the (R, p_n)-method includes the Cesaro method and therefore, the generalized piecewise ergodic theorem becomes a direct consequence of the usual Birkhoff–Khinchin theorem. We should also mention the paper [24], where the convergence of Riesz averages is considered for a sequence of random variables with non-increasing weights p_n.

Reviews of some new approaches and results pertaining to the generalization of ergodic theorems can be found in [25–27].

5 SUMMABILITY COEFFICIENTS

Again, consider a general measure preserving mapping $T\colon M \to M$. Let $f\colon M \to \mathbb{R}$ be an integrable function with *zero* mean value,

$$\int\limits_M f\,d\mu = 0.$$

For $f \not\equiv 0$, the series

$$\sum_{n=0}^{\infty} f(T^n x) \tag{5.1}$$

is divergent, in general. We assume that T is an ergodic transformation.

For the investigation of divergent series it is convenient to introduce *summability coefficients* (see, for instance, [3]). In our case, this is a non-increasing sequence of positive numbers α_n such that the series

$$\sum_{n=0}^{\infty} \alpha_n f(T^n x) \tag{5.2}$$

is convergent for almost all x. Of special interest are slowly decreasing sequences α_n.

In some cases, one can take an arbitrary non-decreasing sequence $\alpha_n > 0$, $\alpha_n \to 0$, as summability coefficients. Thus, if $f \in C^2$, then for almost all rotations of the circle, partial sums of the series (5.1) are bounded (see, for instance, [28]), and therefore, by the Abel–Dirichlet theorem the series (5.2) is convergent.

This, however, is a rare exception. For instance, in order to ensure the convergence of the series (5.2), where $f(T^n x)$ are the Rademacher functions $r_{n+1}(x)$, $0 \leqslant x \leqslant 1$, it is required that α_n decay faster than $1/\sqrt{n}$ as $n \to \infty$. Indeed, by the Kolmogorov–Khinchin theorem for series, (5.2) is convergent for almost all x if and only if

$$\sum \alpha_n^2 < \infty.$$

Of special interest is the case of $\alpha_n = 1/n$. If the series (5.2) is convergent for almost all $x \in M$, then (by the Kronecker lemma) we obtain the Birkhoff–Khinchin theorem:

$$f(T^n x) \to 0 \quad \text{(C)} \tag{5.3}$$

almost everywhere. Thus, if one could show that the series (5.2) with $\alpha_n = 1/n$ is convergent almost everywhere for integrable functions f, we would have a remarkable strengthened version of the Birkhoff–Khinchin theorem (at least for ergodic transformations).

Taking $\alpha_n = 1/n$, let us apply the Abel transformation to the series (5.2). Then (5.2) turns into the series

$$\sum_{n=1}^{\infty} \frac{S_n}{n(n+1)}, \tag{5.4}$$

where S_n is a partial sum of the series (5.1). According to (5.3), we have $S_n(x) = o(n)$ for almost all x. Therefore, the convergence of the series (5.4) depends on whether the decay rate of the ratio S_n/n is faster or slower than that of $1/\ln n$.

For example, if T is the Bernoulli shift of the unit segment and $f = r_1(x)$ is the Rademacher function, then

$$S_n = O(\sqrt{n \ln\ln n}) \quad \text{a. s.}$$

(the Khinchin iterated logarithm law). Therefore, the series (5.4) converges by a wide margin.

However, examples show that S_n may grow faster than $n/\ln n$, which results in the divergence of the series (5.4). Consider the case of $M = \{x \bmod 2\pi\}$ being a circle with the standard measure, and let T be the shift

$$x \mapsto x + \alpha$$

by an angle α incommensurable with π.

Theorem 7. *Let $\varphi(z)$, $z > 0$ be a continuously differentiable function, $\varphi' < 0$ and $\varphi(z) \to 0$ for $z \to \infty$. There is a continuous 2π-periodic function $f(x)$ with zero mean value such that for some x_0, the following inequality holds:*

$$S_n(x_0) = f(x_0) + \ldots + f(x_0 + \alpha(n-1)) \geqslant cn\varphi(n) - \gamma, \quad c, \gamma = \text{const} > 0. \tag{5.5}$$

This statement is interesting by itself as a supplement to the classical Bole–Serpinsky–Weyl theorem. It shows that the values of the sum S_n may grow as an arbitrary function of $o(n)$. Note that as a rule, the function f is non-differentiable at all points (cf. [28, Chap. VII]). For $f \in C^2$, we have $S_{n_k} \to 0$ for a subsequence $n_k \to \infty$ uniformly in x [28,29]. This statement also holds for absolutely continuous functions, as shown in a later publication by Sidorov [30].

It should also be observed that in fact there are many more points x_0 for which the estimate (5.5) holds. Such points form (at least) a countable dense set. Indeed, instead of x_0 we can take points of the form $x_0 + m\alpha$ with integer m. In this connection, a question arises whether estimates of the form (5.5) hold for x_0 constituting a set of positive measure. If this is so, then the above strengthened version of the Birkhoff–Khinchin theorem would be impossible. In any case, the Bole–Serpinskii–Weyl theorem cannot be strengthened in this way (it suffices to take $\varphi(z) = 1/\ln z$ in Theorem 7). However, it has been shown in [31] that for almost all x, the sequence $S_n(x)$ of partial sums of the series (5.1) changes sign infinitely many times, and therefore, cannot go to infinity as $n \to \infty$.

The proof of Theorem 7 is based on a construction that goes back to Poincaré [32]. The Poincaré example [33] showing that the sum S_n can grow to infinity, together with precise estimates (and some minor errors in [33] corrected), is considered in the monograph [28]. A special case of Theorem 7 is examined in [28] for $n\varphi(n) = cn^\alpha$, $0 < \alpha < 1$.

We will prove Theorem 7 for a conditionally periodic flow of a two-dimensional torus. Using the Poincaré section, it is easy to obtain a relevant example for rotations of a circle (see, for instance, [28, Chap. VIII]).

Thus, on the two-dimensional torus $\mathbb{T}^2 = \{\varphi_1, \varphi_2 \bmod 2\pi\}$, consider the dynamical system

$$\dot\varphi_1 = 1, \quad \dot\varphi_2 = \sqrt{2}.$$

Its solutions are linear functions

$$\varphi_1 = t + \varphi_1^0, \quad \varphi_2 = \sqrt{2t} + \varphi_2^0, \tag{5.6}$$

$\varphi_s^0 = \text{const}$. Since $\sqrt{2}$ is an irrational number, each trajectory is dense in $\mathbb{T}^2$.

Let us introduce a sequence of integers u_n, v_n, determined by the relation

$$(\sqrt{2} - 1)^n = u_n + v_n\sqrt{2}.$$

Thus, $u_1 = -1$, $u_2 = 3, \ldots$, $v_1 = 1$, $v_2 = -2, \ldots$.

Next, we introduce a function $f\colon\ \mathbb{T}^2 \to \mathbb{R}$, defined by the gap trigonometric series

$$\sum_{n=1}^{\infty} \psi(n) \cos(u_n\varphi_1 + v_n\varphi_2), \tag{5.7}$$

where $\psi(n)$ is a decreasing sequence of positive numbers such that

$$\sum \psi(n) < \infty. \tag{5.8}$$

The function f is continuous and (5.7) is its Fourier series. Moreover, the mean value of f over the torus is equal to zero.

Let us take zero initial phases φ_s^0 in (5.6). On the solution (5.6), f is a function of t. The integral of this function has the form

$$I(t) = \int_0^t \sum \psi(n) \cos\frac{t}{\Lambda^n}\, dt, \quad \Lambda = \sqrt{2} + 1. \tag{5.9}$$

Since the series (5.8) is convergent, we can transpose the integration and the sum,

$$I(t) = \sum_{n=1}^{\infty} \Lambda^n \psi(n) \sin\frac{t}{\Lambda^n}.$$

Let us estimate the growth of this function, following [28]. Let

$$\frac{\pi}{2}\Lambda^{n-1} \leqslant t \leqslant \frac{\pi}{2}\Lambda^{n}, \quad n = 1, 2 \ldots.$$

Then

$$0 < \frac{\pi}{2\Lambda^{k+1}} \leqslant \frac{t}{\Lambda^{n+k}} \leqslant \frac{\pi}{2\Lambda^{k}} < \frac{\pi}{2}, \quad k = 0, 1, \ldots.$$

It follows that

$$\sin\frac{t}{\Lambda^{n+k}} \geqslant \frac{2t}{\pi\Lambda^{n+k}} \geqslant \frac{1}{\Lambda^{k+1}},$$

and therefore,

$$I(t) \geqslant \sum_{k=0}^{\infty} \Lambda^{n+k}\psi(n+k)\sin\frac{t}{\Lambda^{n+k}} - \left|\sum_{k=0}^{n-1} \Lambda^{k}\psi(k)\sin\frac{t}{\Lambda^{k}}\right| \geqslant \Lambda^{n-1}\sum_{k=0}^{\infty}\psi(n+k) - \sum_{k=0}^{n-1}\psi(k). \tag{5.9'}$$

In the last estimate, we have used the obvious inequality $\sin x < x$ $(x > 0)$. In view of (5.8), the second term in the estimate (5.9) does not exceed a constant γ.

By assumption, $\psi(x)$ is a monotonically decreasing function. Therefore,

$$\sum_{k=0}^{\infty}\psi(n+k) \geqslant \int_{n}^{\infty}\psi(x)\,dx.$$

Thus, we have obtained the following estimate on the interval $[\pi\Lambda^{n-1}/2,\ \pi\Lambda^{n}/2]$:

$$I(t) \geqslant \Lambda^{n-1}\int_{n}^{\infty}\psi(x)\,dx - \gamma. \tag{5.10}$$

Clearly, on this interval, we have

$$\Lambda^{n-1} \geqslant \frac{2t}{\pi\Lambda}, \quad n \leqslant \frac{1}{\ln\Lambda}\ln\frac{2t}{\pi} + 1.$$

Taking into account these inequalities, we can refine the estimate (5.10):

$$I(t) \geqslant \frac{2t}{\pi\Lambda}\int_{\xi(t)}^{\infty}\psi(x)\,dx - \gamma, \tag{5.11}$$

where

$$\xi(t) = \frac{1}{\ln\Lambda}\ln\frac{2t}{\pi} + 1.$$

In view of (5.8) and the assumption that $\psi(x)$ is monotonically decreasing, the function

$$\int_{\xi(t)}^{\infty}\psi(x)\,dx \tag{5.12}$$

tends to zero as $t \to \infty$. Let $\varphi(t) > 0$ be an arbitrary continuously differentiable function, $\varphi'(t) < 0$ and $\varphi(t) \to 0$ as $t \to \infty$. Let us take $\varphi(t) = \zeta(\ln t)$ and equate ζ to the integral (5.12):

$$\int_{\xi(t)}^{\infty}\psi(x)\,dx = \zeta(\ln t).$$

Differentiating both sides of this relation in t, we obtain

$$\psi(\xi(t)) = -\zeta'\ln\Lambda,$$

or

$$\psi\left(\frac{z}{\ln\Lambda} + \frac{\ln(2/\pi)}{\ln\Lambda} + 1\right) = -\zeta'(z)\ln\Lambda.$$

Since $\zeta' < 0$, this relation determines a function ψ for which condition (5.8) is fulfilled. As a result, the estimate (5.11) takes the desired form

$$I(t) \geqslant ct\varphi(t) - \gamma, \quad c = \frac{2}{\pi\Lambda}.$$

6 HOMOLOGICAL EQUATION

Here, we continue our discussion of questions pertaining to the summation of the divergent series (5.1). Again, we assume that the integrable function f has zero mean value. In contrast to Sec. 5, the transformation T may be non-ergodic.

An important role in the ergodic theory belongs to the *homological equation*

$$g(Tx) - g(x) = f(x). \tag{6.1}$$

As an example, consider a *cylindrical cascade* — a discrete dynamical system on the direct product $M \times \mathbb{R} = \{x, y\}$ determined by the mapping

$$x \mapsto Tx, \quad y \mapsto y + f(x). \tag{6.2}$$

It is easy to understand that each solution g of the homological equation (6.1) determines an invariant set $\{y = g(x)\}$ of the mapping (6.2). The homological equation occurs widely in problems of reduction and conjugation of dynamical systems (see, for instance, [34]).

By definition, a *solution of Eq.* (6.1) is a Lebesgue integrable function $g\colon M \to \mathbb{R}$ satisfying (6.1) for almost all $x \in M$. Integrating both sides of (6.1) over the set M and taking into account that the measure μ is invariant with respect to T, we obtain a necessary condition of solvability of the homological equation:

$$\int\limits_M f\, d\mu = 0. \tag{6.3}$$

Theorem 8. *The homological equation admits a solution if and only if the series* (5.1) *converges almost everywhere in the sense of Cesaro to an integrable function.*

Proof. Let g be an integrable solution of (6.1). Iterating the homological equation, we obtain the chain of equalities

$$g(T^{n+1}x) - g(T^n x) = f(T^n x), \quad n = 0, 1 \ldots.$$

Hence,

$$S_n = g(T^n x) - g(x),$$

where $S_n (n \geqslant 1)$ are partial sums of the series (5.1). Therefore,

$$\frac{S_1 + \ldots + S_n}{n} = \frac{g(Tx) + \ldots + g(T^n x)}{n} - g(x).$$

By the Birkhoff–Khinchin theorem, for almost all $x \in M$, we have

$$S_n(x) \to \bar{g}(x) - g(x) \quad (\mathrm{C}),$$

where $\bar{g}$ is an integrable function invariant with respect to T.

Conversely, let

$$\sum_{n=0}^{\infty} f(T^n x) = G(x) \quad (\mathrm{C}) \tag{6.4}$$

almost everywhere on M, and let G be an integrable function. Therefore,

$$\sum_{n=0}^{\infty} f(T^n x) = G(Tx) \quad (\mathrm{C}). \tag{6.5}$$

It is well known ([3, Chap. I]) that the Cesaro method is *natural* in the following sense: if

$$a_0 + a_1 + a_2 + \ldots = a \quad (\mathrm{C}),$$

then

$$a_1 + a_2 + \ldots = a - a_0 \quad \text{(C)};$$

and the converse statement is true. Using this property, together with (6.4) and (6.5), we obtain the relation

$$f(x) = G(x) - G(Tx).$$

Consequently, the integrable function $-G$ is a solution of the homological equation (6.1). The proof is complete.

As a simple example, consider the homological equation for the Bernoulli shift of the unit interval, taking f equal to the Rademacher function $r_1(x)$. Clearly, the necessary condition of solvability (6.3) is satisfied in this case. If the equation admits a Lebesgue integrable solution, then (by Theorem 8) the series

$$\sum_{n=0}^{\infty} r_n(x), \quad 0 \leqslant x \leqslant 1, \tag{6.6}$$

is convergent in the sense of Cesaro for almost all x. However, as shown by Zygmund [35] (see also [36]), if the series (6.6) is summable by some matrix method (for instance, the Cesaro method) almost everywhere, it must be convergent almost everywhere. However, this is not so, and therefore the corresponding homological equation admits no summable solutions. In fact, this statement can be proved directly.

From the proof of Theorem 8, we see that the Cesaro method can be replaced by any stronger method which is natural in the above sense, in particular, the Abel method.

As an illustration, let us calculate the solution of Eq. (6.1) for the rotation of the circle $x \mapsto x + \alpha$. Let f be represented by the Fourier series

$$\sum{}' c_m e^{imx},$$

with

$$\sum{}' |c_m| < \infty. \tag{6.7}$$

The prime indicates that $m \neq 0$. From (6.7), it follows that f is continuous.

Let us apply the Abel method and associate the series (5.1) with the power series

$$\sum_n \Big[\sum_m{}' c_m e^{imx} e^{imn\alpha} \Big] z^n.$$

In view of (6.7), the summs over m and n can be transposed,

$$\sum_m{}' c_m e^{imx} \sum_n (e^{im\alpha} z)^n.$$

For $|z| < 1$, the last sum is equal to

$$\sum_m{}' \frac{c_m}{1 - z e^{im\alpha}} e^{imx}. \tag{6.8}$$

Assuming that

$$\sum{}' \left| \frac{c_m}{1 - e^{im\alpha}} \right| < \infty$$

and passing to the limit in (6.8) as $z \to 1$, we obtain the function

$$\sum{}' \frac{c_m}{1 - e^{im\alpha}} e^{imx}, \tag{6.9}$$

which, multiplied by -1, is a continuous solution of the homological equation

$$g(x + \alpha) - g(x) = f(x).$$

Of course, formula (6.9) can be obtained directly if this equation is solved by the Fourier method.

REFERENCES

1. Halmos P., *Lectures on the Ergodic Theory* [Russian translation], Inostr. Lit., Moscow (1959).
2. Weyl H., *Selected Works* [Russian translation], Nauka, Moscow (1984).
3. Hardy G., *Divergent Series* [Russian translation], Inostr. Lit., Moscow (1951).
4. Cook R., *Infinite Matrices and Spaces of Sequences* [Russian translation], Inostr. Lit., Moscow (1960).
5. Polya G. and Szegö G., *Problems and Theorems of Analysis*, Vol. 1 [Russian translation], Nauka, Moscow (1978).
6. Cigler J., "Methods of summability and uniform distribution mod 1," *Composito Math.*, **16**, 44–51 (1964).
7. Doroidar A. F., "A note on the generalized uniform distribution (mod 1)," *J. Natur. Sci. und Math.*, **11**, 185–189 (1971).
8. Keipers L. and Niederreuter G., *Uniform Distribution of Sequences* [Russian translation], Nauka, Moscow (1985).
9. Tsuji M., "On the uniform distribution of numbers mod 1," *J. Math. Soc. Japan*, **4**, 313–322 (1952).
10. Kozlov V. V., "On uniform distribution," *Izv. Vuzov, Sev.-Kavk. Region. Estestv. Nauki*, Special issue, 96–99 (2001).
11. Hanson D. L. and Pledger G., "On the mean ergodic theorem for weighted averages," *Z. Wahrscheinlichkeits-theor. und verw. Geb.*, **13**, N 13, 141–149 (1969).
12. Cohen L. W., "On the mean ergodic theorem," *Ann. Math. II Ser.*, **41**, 505–509 (1940).
13. Katz M., *Statistical Independence in Probability, Analysis, and Theory of Numbers* [Russian translation], Inostr. Lit., Moscow (1963).
14. Frank W. E. and Hanson D. L., "Some results giving rates of convergence in the law of large numbers for weighted sums of independent random variables," *Trans. Amer. Math. Soc.*, **124**, No. 2, 347–359 (1966).
15. Hanson D. L. and Wright F. T., "Some convergence results for weighted sums of independent random variables," *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **19**, 81–89 (1971).
16. Kolmogorov A. N., "On the strong law of large numbers," in: A. N. Kolmogorov, *Theory of Probability and Mathematical Statistics* [in Russian], Nauka, Moscow (1986), 59–60.
17. Feller W., *Introduction to the Theory of Probability and Its Applications*, Vol. 2 [Russian translation], Mir, Moscow (1984).
18. Halmos P., *Theory of Measure* [Russian translation], Inostr. Lit., Moscow (1953).
19. Kolmogorov A. N., "On the iterated logarithm law," in: A. N. Kolmogorov, *Theory of Probability and Mathematical Statistics*, Nauka, Moscow (1986), 34–44.
20. Bourbaki N., *Functions of Real Variable* [Russian translation], Nauka, Moscow (1965).
21. Baxter G., "An ergodic theorem with weighted averages," *J. Math. and Mech.*, **13**, No. 3, 481–488 (1964).
22. Baxter G., "A general ergodic theorem with weighted averages," *J. Math. and Mech.*, **14**, No. 2, 277–288 (1965).
23. Chacon R. V., "Ordinary means imply recurrent means," *Bull. Amer. Math. Soc.*, **70**, 796–797 (1964).
24. Gaposhkin V. F., "On summation of stationary sequences by the Riesz methods," *Mat. Zametki*, **57**, No. 1, 653–662 (1995).
25. Katok A. B., Sinai Ya. G., and Stepin A. M., "Theory of Dynamical Systems and General Groups of Transformations with an Invariant Measure," *Mathematical Analysis* (Itogi Nauki i Tekhniki, Vol. 13) [in Russian], VINITI, Moscow (1975), 129–262.
26. Krengel U., "Recent progress in ergodic theorems," *Astérisque*, No. 50, 151–192 (1977).
27. Vershik A. M., Kornfeld I. P., and Sinai Ya. G., "General ergodic theory of groups of transformations with an invariant measure," *Modern Problems of Mathematics. Fundamental Trends*, (Itogi Nauki i Tekhniki, Vol. 2) [in Russian], VINITI, Moscow (1985), 5–111.

28. Kozlov V. V., *Methods of Qualitative Analysis in the Dynamics of Solid Bodies* [in Russian], Izd-vo Mosk. Univ., Moscow (1980).

29. Kozlov V. V., "On a problem of Poincaré," *Prikl. Mat. Mekh.*, **40**, No. 2, 325–355 (1976).

30. Sidorov E. A., "On the conditions of uniform stability in the sense of Poisson for cylindrical systems," *Uspekhi Mat. Nauk*, **34**, No. 6, 184–188 (1979).

31. Halász G., "Remarks on the remainder in Birkhoff's ergodic theorem," *Acta Math. Acad. Sci. Hungaricae*, **28**, No. 3–4, 389–395 (1976).

32. Poincaré H., "Sur les séries trigonométriques," *Comp. Rend. Acad. Sci.*, Paris, **101**, No. 2, 1131–1134 (1885).

33. Poincaré H., *On Curves Determined by Differential Equations* [Russian translation], Gostekhizdat, Moscow, Leningrad (1974).

34. Arnold V. I. and Avets A., *Ergodic Problems of Classical Mechanics* [in Russian], Ed. Zhurn. Regular. i Khaotich. Dinam., Izhevsk (1999).

35. Zygmund A., "On the convergence of lacunary trigonometric series," *Fund. Mathematicae*, **16**, 90–107 (1930).

36. Kahane J.-P., *Random Functional Series* [Russian translation], Mir, Moscow (1973).