ON THE STRONG LAW OF LARGE NUMBERS

BY

P. ERDÖS

In the present note $f(x)$, $-\infty < x < \infty$, will denote a function satisfying the following conditions: (1) $f(x+1)=f(x)$, (2) $\int_0^1 f(x)=0$, $\int_0^1 f(x)^2=1$. By $n_1<n_2<\cdots$ we shall denote an arbitrary sequence satisfying $n_{k+1}/n_k>c>1$, and by $S_n(f)$ the nth partial sum of the Fourier series of $f(x)$.

In a recent paper Kac, Salem, and Zygmund(1) prove (among others) that if for some $\epsilon>0$

$$\int_0^1 (f(x) - \phi_n(f))^2 = O\left(\frac{1}{(\log n)^\epsilon}\right), \tag{1}$$

then for almost all x

$$\lim_{N\to\infty} \frac{1}{N}\left(\sum_{k=1}^{N} f(n_k x)\right) = 0, \tag{2}$$

or roughly speaking the strong law of large numbers holds for $f(n_k x)$ (in fact the authors prove that $\sum f(n_k x)/k$ converges almost everywhere).

The question was raised whether (2) holds for any $f(x)$. This was known for the case $n_k=2^k$(2). In the present paper it is shown that this is not the case. In fact we prove the following theorem.

THEOREM 1. *There exists an $f(x)$ and a sequence n_k so that for almost all x*

$$\limsup_{N\to\infty} \frac{1}{N}\left(\sum_{k=1}^{N} f(n_k x)\right) = \infty. \tag{3}$$

Further we prove the following sharpening of the result of Kac-Salem-Zygmund:

THEOREM 2. *Assume that for some $\epsilon>0$*

$$\int_0^1 (f(x) - \phi_n(f))^2 = O\left(\frac{1}{(\log\log n)^{2+\epsilon}}\right), \tag{4}$$

then (2) *holds.*

By a slight modification of the construction of the $f(x)$ of Theorem 1 it is easy to construct an $f(x)$ and a sequence n_k for which (3) holds and for which

Presented to the Society, February 28, 1948; received by the editors January 28, 1948.

(1) Trans. Amer. Math. Soc. vol. 63 (1948) pp. 235–243.

(2) This result is due to Raikov. See F. Riesz, Comment. Math. Helv. vol. (17) (1944) p. 223.

$$\int_0^1 (f(x) - \phi_n(f))^2 < \frac{1}{(\log\log\log n)^c}. \tag{5}$$

There is clearly a gap between (4) and (5). It seems probable that, in Theorem 2, (4) can be replaced by $1/(\log\log\log n)^{c_1}$, but much sharper methods would be needed than used here.

The following problem also seems of some interest: By an easy modification in the construction of the $f(x)$ of Theorem 1 we can show the existence of an $f(x)$ and a sequence n_k, so that for almost all x

$$\limsup_{N\to\infty} \frac{1}{N(\log\log N)^{1/2-\epsilon}}\left(\sum_{k=1}^{N} f(n_k x)\right) = \infty. \tag{6}$$

On the other hand we can show that for almost all x

$$\lim_{N\to\infty} \frac{1}{N(\log N)^{1/2+\epsilon}}\left(\sum_{k=1}^{N} f(n_k x)\right) = 0. \tag{7}$$

Again there is a gap between (6) and (7). (6) seems to give the right order of magnitude, but I can not prove this.

One final remark. The $f(x)$ of Theorem 1 is unbounded. The possibility that (2) holds for all bounded functions $f(x)$ remains open.

Proof of Theorem 1. Let u_k, v_k, and A_k tend to infinity sufficiently fast (their growth will be specified later). $r_m(x)$ denotes the mth Rademacher function(³). Put

$$f(x) = \sum_{k=1}^{\infty} \sum_{m=u_k+1}^{v_k m} \frac{v_m(x)}{(A_k(v_k - u_k))^{1/2}}, \qquad \sum_{k=1}^{\infty} \frac{1}{A_k} = 1. \tag{8}$$

Clearly the series for $f(x)$ converges almost everywhere and $\int_0^1 f(x) = 0$, $\int_0^1 f(x)^2 = 1$. Now we define the n_k. Put $j_k = [e^{A_k^3}]$. Denote by $I_t^{(k)}$ the interval

$$((2t-1)v_k,\ (2t-1)v_k + l_t^{(k)}), \qquad t = 1, 2, \cdots, j_k,$$

where $l_t^{(k)} = 2l_{t-1}^{(k)}$ and $l_1^{(k)}$ is very large compared to v_{k-1}, A_{k-1}, $l_{j_{k-1}}^{(k-1)}$, and will be specified later. If we choose $v_k > l_{j_k}^{(k)}$ then the $I_t^{(k)}$ don't overlap. The n_k are the integers of the form 2^m where $m \subset I_t^{(k)}$, $k = 1, 2, \cdots; t = 1, 2, \cdots, j_k$.

Order the l's according to their size. Clearly each l is greater than the sum of all previous l's. Thus a simple argument shows that to prove (3) it will be sufficient to show that for every fixed c and almost all x

$$\limsup \frac{1}{l_t^{(k)}}\left(\sum_{m\subset I_t^{(k)}} f(2^m x)\right) > c, \qquad k = 1, 2, \cdots; t = 1, 2, \cdots, j_k. \tag{9}$$

(³) Instead of $r_m(x)$ I originally used $\cos 2^m x$. The advantage of using Rademacher functions was pointed out to me by Kac.

(Since if $m_{r+1} > 2m_r$, and for every c lim sup $(1/(m_{r+1}-m_r)) \sum_{m_r}^{m_{r+1}} a_u > c$, then lim sup $(1/u) \sum_{k=1}^{u} a_k = \infty$. Let now m_r be the sum of the r first l's, then clearly (3) is a consequence of (9).)

Hence it will suffice to show that for every ϵ and sufficiently large k the measure of the set in x satisfying at least one of the inequalities

$$\frac{1}{l_t^{(k)}} \left(\sum_{m \subset I_t^{(k)}} f(2^m x) \right) > c, \qquad t = 1, 2, \cdots, j_k, \tag{10}$$

is greater than $1-\epsilon$.

Put

$$f(x) = f_1(x) + f_2(x) + f_3(x)$$

where

$$f_1(x) = \sum_{s=1}^{k-1} \sum_{m=u_s+1}^{v_s} \frac{r_m(x)}{(A_s(v_s - u_s))^{1/2}}, \qquad f_2(x) = \sum_{m=u_k+1}^{v_k} \frac{r_m(x)}{(A_k(u_k - v_k))^{1/2}},$$

$$f_3(x) = \sum_{s>k} \sum_{m=u_s+1}^{v_s} \frac{r_m(x)}{(A_s(v_s - u_s))^{1/2}}.$$

A simple calculation shows that

$$\begin{aligned} \sum_{m \subset I_t^{(k)}} f_2(2^m x) &= \frac{l_t^{(k)}}{(A_k(v_k - u_k))^{1/2}} \sum r_m(x) + \sum\nolimits_1 + \sum\nolimits_2 \\ &= \sum + \sum\nolimits_1 + \sum\nolimits_2 \end{aligned} \tag{11}$$

where m runs in the interval

$$(u_k + (2t-1)v_k + l_t^{(k)}, 2tv_k)$$

and

$$\sum\nolimits_1 = \sum_{a=1}^{l_t^{(k)}} \frac{l_t^{(k)} - a}{(A_k(v_k - u_k))^{1/2}} r_{y-a}(x), \qquad y = u_k + (2t-1)v_k + l_t^{(k)},$$

$$\sum\nolimits_2 = \sum_{a=1}^{l_t^{(k)}} \frac{l_t^{(k)} - a}{(A_k(v_k - u_k))^{1/2}} r_{2tv_k+a}(x).$$

Now $\sum r_m(x)$ is the sum of

$$v_k - u_k - l_t^{(k)} > v_k/2$$

Rademacher functions (we choose $v_k > 2(u_k + l_t^{(k)})$). It is well known(4) that

(4) See, for example, P. Erdös, Ann. of Math. vol. 43 (1942) p. 420, formula (0.7). Incidentally the formula in question should read $c_1(x^2/n)e^{-2x^2/n} < \Pr\ (A_n(x)) < c_2(x^2/n)e^{-2x^2/n}$.

the measure of the set in x for which

$$\sum r_m(x) > 4c(A_k)^{1/2}(v_k)^{1/2}$$

is greater than

$$c_1 A_k e^{-32c^2 A_k} > e^{-A_k^2}$$

for sufficiently large A_k. Thus the measure of the set in x for which

$$\sum = \frac{l_t^{(k)}}{(A_k(v_k - u_k))^{1/2}} \sum r_m(x) > 4cl_t^{(k)} \tag{12}$$

is greater than $e^{-A_k^2}$. Clearly for all x

$$\left| \sum\nolimits_1 + \sum\nolimits_2 \right| < \frac{2(l_t^{(k)})^2}{(A_k(v_k - u_k))^{1/2}} < \frac{4(l_t^{(k)})^2}{(v_k)^{1/2}} < 1 \tag{13}$$

if we choose $v_k > 16(l_t^{(k)})^4$. Thus finally from (11), (12), and (13) the measure of the set in x for which

$$\sum_{m \subset I_t^{(k)}} f_2(2^m x) > 4cl_t^{(k)} - 1 > 3cl_t^{(k)} \tag{14}$$

is greater than $e^{-A_k^2}$.

If $v_k > 2l_t^{(k)}$ for all t, then the functions

$$\sum_{m \subset I_t^{(k)}} f_2(2^m x), \qquad t = 1, 2, \cdots j_k,$$

are independent (since the same $r_m(x)$ does not appear in two different sums). Thus the measure of the set in x for which one of the j_k inequalities

$$\sum_{m \subset I_t^{(k)}} f_2(2^m x) > 3cl_t^{(k)}, \qquad t = 1, 2, \cdots, j_k, \tag{15}$$

holds, is greater than

$$1 - (1 - 1/y)^z > 1 - \epsilon/2 (y = e^{A_k^2}, z = e^{A_k^3}). \tag{16}$$

Further if $l_t^{(k)} > v_{k-1}$

$$\int_0^1 \left(\sum_{m \subset I_t^{(k)}} f_1(2^m x) \right)^2 < v_{k-1}^2 (l_t^{(k)} + v_{k-1}) < 2v_{k-1}^2 l_t^{(k)}$$

since only the r_m's with $m \leqq l_t^{(k)} + v_{k-1}$ occur and the coefficients of all of them are not greater than v_{k-1}. Thus from Tchebychef's inequality we obtain that the measure of the set in x for which one of the j_k inequalities

$$\sum_{m \subset I_t^{(k)}} f_1(2^m x) > cl_t^{(k)}, \qquad t = 1, 2, \cdots, j_k, \tag{17}$$

holds is less than

$$\sum_{t=1}^{j_k} \frac{2v_{k-1}^2}{c^2 l_t^{(k)}} < \frac{4v_{k-1}^2}{c l_1^{(k)}} < \frac{\epsilon}{4}, \quad l_1^{(k)} > 16 v_{k-1}^2/c\epsilon. \tag{18}$$

Finally we have by a simple computation

$$\int_0^1 \left(\sum_{m \subset I_t^{(k)}} f_3(2^m x) \right)^2 < 4(l_t^{(k)})^2 \sum_{r>k} \frac{1}{A_r} < 1$$

if $A_{k+1}, \cdots$ are sufficiently large. Thus the measure of the set in x for which one of the inequalities

$$\sum_{m \subset I_t^{(k)}} f_3(2^m x) > c l_t^{(k)}, \qquad t = 1, 2, \cdots, j_k, \tag{19}$$

holds is less than

$$\sum_{t=1}^{j_k} \frac{1}{(c l_t^{(k)})^2} < \frac{\epsilon}{4}. \tag{20}$$

Thus finally from (15), (16), (17), (18), (19), and (20) we obtain (10) and this completes the proof of Theorem 1.

Sketch of the Proof of Theorem 2. Put $j-i=r$, then $n_j/n_i > c^r$. Denote by $a_1, b_1, a_2, b_2, \cdots$ the Fourier coefficients of $f(x)$. By (4) we evidently have

$$\int_0^1 f(n_i x) f(n_j x) = \sum_{n_i u = n_j v} (a_u a_v + b_u b_v) \leqq \left(\sum_{k=1}^{\infty} a_k^2 \sum_{k > c^r} a_k^2 \right)^{1/2} + \left(\sum_{k=1}^{\infty} b_k^2 \sum_{k > c^r} b_k^2 \right)^{1/2} < \frac{c_1}{(\log r)^{1+\epsilon/2}}.$$

Hence

$$\int_0^1 \left(\sum_{z}^{z+N} f(n_k x) \right)^2 = O\left(\frac{N^2}{(\log N)^{1+\epsilon/2}} \right),$$

or the measure of the set $M(z, N, A)$ in x for which

$$\left| \sum_{z}^{z+N} f(n_k x) \right| > A \cdot N$$

is less than

$$c/A^2 (\log N)^{1+\epsilon/2}. \tag{21}$$

Consider the sets

$$\begin{aligned} &M(1, 2^n, \delta);\ M(2^n, 2^{n-1}, 2\delta/2^2); \\ &M(2^n, 2^{n-2}, 4\delta/3^2),\ M(2^n + 2^{n-1}, 2^{n-2}, 4\delta/3^2); \cdots . \end{aligned} \tag{22}$$

There are 2^{k-1} sets of order k, that is, sets of the form

$$M(2^n + 2u2^{n-k},\ 2^{n-k},\ \delta 2^k/(k+1)^2),\ 0 \leqq u < 2^{k-1}. \tag{23}$$

From (21) it follows that the measure of any set of order k does not exceed

$$c(k+1)^4/\delta^2 2^{2k}(n-k)^{1+\epsilon/2}.$$

Thus the measure of all the sets in (23) is less than $c(k+1)^4/\delta^2 2^k(n-k)^{1+\epsilon/2}$, and the measure of all the sets M_n in (22) does not exceed

$$\sum_{k=0}^{n} \frac{c(k+1)^4}{\delta^2 2^k (n-k)^{1+\epsilon/2}} < \frac{c_1}{\delta^2 n^{1+\epsilon/2}}\,.$$

Thus

$$\sum_{n=1}^{\infty} M_n < \infty. \tag{24}$$

But if x does not belong to any of the sets (22) we have by a simple argument for all $2^n \leqq m < 2^{n+1}$ (every m is the sum of powers of 2)

$$\left|\sum_{k=1}^{m} f(n_k x)\right| < \delta 2^n + \frac{\delta 2^n}{2^2} + \frac{\delta 2^n}{3^2} + \cdots + \frac{\delta 2^n}{k^2} + \cdots < 2\delta 2^n \leqq 2\delta m. \tag{25}$$

(24) and (25) clearly prove theorem 2[5].

Syracuse University,
Syracuse, N. Y.

[5] The method used here is due to Hobson-Plancherel-Rademacher-Menchof. (See, for example, Rademacher, Math. Ann. vol. 87 (1922) p. 117–121.)