MATHEMATICS

ON THE LAW OF THE ITERATED LOGARITHM. I

BY

P. ERDÖS AND I. S. GÁL

(Communicated by Prof. J. F. KOKSMA at the meeting of October 30, 1954)

Introduction

The object of the present paper is to prove the following result:

THEOREM. *Let $n_1 < n_2 < \ldots < n_\nu < \ldots$ be an infinite sequence of positive numbers, satisfying the lacunarity condition $n_{\nu+1}/n_\nu \geqslant q > 1$ $(\nu = 1, 2, \ldots)$. Then*

$$\limsup_{N\to\infty} \frac{|\sum_{\nu=1}^{N} \exp 2\pi i\, n_\nu x|}{\sqrt{N \log\log N}} = 1$$

for almost all x.

This result is not unexpected in view of the law of the iterated logarithm for the sum of independent functions and the well known resemblance of $\{\exp\ 2\pi i\, n_\nu x\}$; $n_{\nu+1}/n_\nu \geqslant q > 1$ to a sequence of independent functions (see [1]). However, the proof of the above result presents considerable difficulties.

Previously R. SALEM and A. ZYGMUND [2] proved that

$$\limsup_{N\to\infty} \frac{|\sum_{\nu=1}^{N} \exp 2\pi i\, n_\nu x|}{\sqrt{N \log\log N}} \leqslant 1$$

for almost all x. In some special cases, for instance when $n_\nu = 2^\nu$ this upper estimate can be proved more easily.

Our proof is based on the asymptotic evaluation of the integral

$$I = \int_\alpha^\beta |\sum_{\nu=1}^{N} \exp 2\pi i\, n_\nu x|^{2p}\, dx$$

where $0 \leqslant \alpha < \beta \leqslant 1$, $p = O(\log\log N)$ and $N \to \infty$. This is done by finding an asymptotic formula for the number of solutions of the diophantine equation

$$x_1 + x_2 + \ldots + x_p = y_1 + y_2 + \ldots + y_p$$

and the inequality

$$s - \tfrac{1}{2} \leqslant (x_1 + \ldots + x_p) - (y_1 + \ldots + y_p) \leqslant s + \tfrac{1}{2}$$

(s = arbitrary real), where the unknowns $x_1, \ldots, x_p$ and $y_1, \ldots, y_p$ are restricted to the values $n_1, n_2, \ldots, n_N$. These investigations make up the first section of the present paper.

In the second section we use our asymptotic formula for I to obtain upper and lower estimates for the measure

$$\phi(t) = \text{meas}\, E\,\{x \mid \alpha \leqslant x \leqslant \beta ; \mid \sum_{\nu=1}^{N} \exp 2\pi i\, n_\nu x \mid \geqslant \sqrt{tN \log\log N}\}.$$

These estimates are somewhat sharper than necessary for the rest of the paper but their proof is no more difficult than that of the weaker inequalities.

The third section contains the proof of the "$\geqslant 1$ inequality". Having a lower estimate for $\phi(t)$ and noticing that the total length of those intervals of E the length of which is less than $1/n_1$ is very small, the "$\geqslant 1$ inequality" can be proved rather easily. The last section is devoted to the "$\leqslant 1$ inequality". There are no new ideas involved here, we apply the "dyadic procedure concerning higher moments" to the case of our particular sequence $\{\exp 2\pi i\, n_\nu x\}$. The literature concerning this method can be found in [3] and [4].

A number of questions can be raised in connection with our theorem and possible generalizations thereof. First, suppose $f(x)$ is a smooth function satisfying $\int_0^1 f(x)dx = 0$, $\int_0^1 f(x)^2 dx = 1$. Is it true that

$$(*) \qquad \limsup_{N\to\infty} \frac{\mid \sum_{\nu=1}^{\infty} f(n_\nu x) \mid}{\sqrt{N \log\log N}} = 1$$

for almost all x, whenever $n_{\nu+1}/n_\nu \geqslant q \geqslant 1$? It is easy to see that this equality fails even for trigonometric polynomials; in fact, the example of Erdös–Fortet (see [1]) shows that

$$\frac{\sum_{\nu=1}^{N} f(n_\nu x)}{\sqrt{N}}$$

does not necessarily have a Gaussian distribution. It seems likely that (*) holds with some correcting factor c, but c will depend in general on both $f(x)$ and the sequence $\{n_\nu\}$.

Let $\{f_\nu(x)\}$ be a sequence of independent functions satisfying $\int_0^1 f_\nu(x)dx = 0$, $\int_0^1 f_\nu(x)^2 dx = 1$. The function $\phi(n)$ is said by P. Levy to belong to the upper class if for infinitely many N's

$$(**) \qquad \mid \sum_{\nu=1}^{N} f_\nu(x) \mid > \phi(N)$$

and it belongs to the lower class if (**) holds only for a finite number of N's. In the same way functions of the upper and lower class can be defined for the sums $\sum_{\nu=1}^{N} \exp 2\pi i\, n_\nu x$ $(n_{\nu-1}/n_\nu \geqslant q > 1)$. The question can be asked whether or not these two classes of functions coincide. Our methods developed in this paper are not sufficiently strong to decide this question, though we could sharpen our theorem considerably.

Let $\sum_{\nu=1}^{\infty} a_\nu^2 = \infty$. We can prove by the methods of this paper that

$$\limsup_{N\to\infty} \frac{\left|\sum_{\nu=1}^{b(N)} \exp 2\pi i n_\nu x\right|}{\sqrt{b(N)\log\log b(N)}} = 1.$$

where $b(N) = \sum_{\nu=1}^{N} a_\nu^2$. For the sake of brevity we omit the proof, which would contain no new ideas.

The authors thank Prof. W. H. J. Fuchs for his help and corrections.

1. *Number theoretical investigations*

In this section let $n_1 = 1 < n_2 < n_3 < \ldots < n_N$ be a finite sequence of real numbers satisfying the lacunarity condition $n_{\nu+1}/n_\nu \geqslant q > 1$ $(1 \leqslant \nu < N)$. *We keep this sequence fixed throughout this section.* Our object is to estimate the number of solutions of the diophantine equation

$$A(x, y) = (x_1 + x_2 + \ldots + x_p) - (y_1 + y_2 + \ldots + y_p) = 0,$$

where $x_1, x_2, \ldots, x_p$ and $y_1, y_2, \ldots, y_p$ are restricted to the values $n_1 = 1$, $n_2, \ldots, n_N$. Moreover we want a sharp estimate for the number of solutions of the inequality $s - \frac{1}{2} \leqslant A(x, y) \leqslant s + \frac{1}{2}$; $A(x, y) \neq 0$ where s is an arbitrary real number. (If $|s| > \frac{1}{2}$ the condition $A(x, y) \neq 0$ is automatically satisfied.) The proof requires several steps. First we prove the following estimate for the number of n_ν's lying in a given interval:

Lemma 1. *If* $0 < \alpha < \beta$ *then*

$$(1) \qquad \sum_{\alpha \leqslant n_\nu \leqslant \beta} 1 \leqslant \frac{\log (\beta/\alpha) q}{\log q},$$

and if α *is real then*

$$(2) \qquad \sum_{\alpha \leqslant n_\nu \leqslant \alpha+1} 1 \leqslant \frac{\log 2q}{\log q}.$$

Proof of Lemma 1. In order to prove (1) let ν_0 be defined by the inequality $n_{\nu_0} < \alpha \leqslant n_{\nu_0+1}$ $(n_0 = 0)$ and $i \geqslant 0$ be defined by the inequality $n_{\nu_0+i} \leqslant \beta < n_{\nu_0+i+1}$. If $i = 0$ then (1) is true. If $i \geqslant 1$ then we have

$$\beta \geqslant n_{\nu_0+i} \geqslant q^{i-1} n_{\nu_0+1} \geqslant q^{i-1} \alpha.$$

Hence $\beta q/\alpha \geqslant q^i$ and (1) follows immediately. Now we prove (2): Since $n_1 = 1$ (2) is trivial for negative values of α. If $0 \leqslant \alpha \leqslant 1$ we have

$$\sum_{\alpha \leqslant n_\nu \leqslant \alpha+1} 1 \leqslant \sum_{1 \leqslant n_\nu \leqslant 2} 1,$$

whence we obtain (2) by using (1) with $\alpha = 1$, $\beta = 2$. If $\alpha > 1$ then (2) is an immediate consequence of (1).

Next we want to prove that the number of $n_k \neq n_l$ pairs satisfying the inequality $s - \frac{1}{2} \leqslant n_k - n_l \leqslant s + \frac{1}{2}$ is uniformly bounded for every real s. More precisely we prove the following:

Lemma 2. *Let* s *be an arbitrary real number and let* $\phi(1, N, s)$ *denote the number of those* $n_k \neq n_l$ *pairs which satisfy the inequality*

$$s - \tfrac{1}{2} \leqslant n_k - n_l \leqslant s + \tfrac{1}{2}.$$

Then there exists a positive constant $c = c(q)$ *independent of* s *and the choice of the sequence* $n_1 = 1, n_2, \ldots, n_N$ *such that*

$$0 \leqslant \phi(1, N, s) \leqslant c. \tag{3}$$

Remark. In fact we shall prove that (3) holds for any $c = c(q)$ satisfying

$$c \geqslant 2\,\frac{\log 2q}{\log q}\cdot\frac{\log 2q^2/(q-1)}{\log q}, \tag{4}$$

whence the independence is obvious.

Proof of Lemma 2. First of all $\phi(1, N, s) = \phi(1, N, -s)$, hence we may assume that $s \geqslant 0$. If $0 \leqslant s \leqslant \frac{1}{2}$ then $\frac{1}{2}\phi(1, N, s)$ is not more than the number of n_k, n_l pairs satisfying $0 < n_k - n_l \leqslant 1$. Since $k \geqslant l+1$ we get

$$1 \geqslant n_k - n_l \geqslant n_k(1 - q^{-1}),$$

and on the other hand $n_k \geqslant 1$. Therefore the possible n_k's satisfy the inequality $1 \leqslant n_k \leqslant q/(q-1)$ and their number can be estimated by (1):

$$\sum_{n_k} 1 \leqslant \frac{\log q^2/(q-1)}{\log q}.$$

For a fixed value of n_k the number of possible n_l's can be estimated by (2). Namely we have $n_k - 1 \leqslant n_l \leqslant n_k$, and so by (2)

$$\sum_{n_l} 1 \leqslant \frac{\log 2q}{\log q}.$$

Consequently we have for $0 \leqslant s \leqslant \frac{1}{2}$

$$\phi(1, N, s) \leqslant 2\,\frac{\log 2q}{\log q}\cdot\frac{\log q^2/(q-1)}{\log q}.$$

Now let $s \geqslant \frac{1}{2}$. We have $n_k > n_l$ and so

$$s + \tfrac{1}{2} \geqslant n_k - n_l \geqslant n_k\,(1 - q^{-1}).$$

On the other hand $n_k \geqslant \max\,(1, s - \frac{1}{2})$. Hence if $s \geqslant \frac{3}{2}$ we have the inequality

$$s - \tfrac{1}{2} \leqslant n_k \leqslant \frac{2(s-\frac{1}{2})}{(1-q^{-1})}$$

and if $\frac{1}{2} \leqslant s \leqslant \frac{3}{2}$ then

$$1 \leqslant n_k \leqslant \frac{2}{(1-q^{-1})}.$$

In the first case we may use (1) with $\alpha = s - \frac{1}{2}$, $\beta = 2(s - \frac{1}{2})q/(q-1)$ and in the second case with $\alpha = 1$, $\beta = 2q/(q-1)$. Consequently in either case

$$\sum_{n_k} 1 \leqslant \frac{\log 2q^2/(q-1)}{\log q}.$$

For a fixed value of n_k the number of possible n_l's can be estimated by (2): We have $(n_k - s) - \frac{1}{2} \leqslant n_l \leqslant (n_k - s) + \frac{1}{2}$, hence

$$\sum_{n_l} 1 \leqslant \frac{\log 2q}{\log q}.$$

Therefore we have for $s \geqslant \frac{1}{2}$

$$\phi(1, N, s) \leqslant \frac{\log 2q}{\log q} \cdot \frac{\log 2q^2/(q-1)}{\log q}.$$

This completes the proof of (3) and (4).

Now we consider the inequality

$$s-\tfrac{1}{2} \leqslant A(x, y) \leqslant s+\tfrac{1}{2}$$

where $A(x, y)$ denotes the linear form

$$A(x, y) = (x_1+x_2+\dots+x_p)-(y_1+y_2+\dots+y_p).$$

We want to prove that the number of distinct (x_p, y_p) pairs which occur among the solutions is at most $O(N)$ uniformly in s. (Here the restriction $A(x, y) \neq 0$ is omitted. Hence choosing $s = 0$ we obtain a similar result for the solutions of $A(x, y) = 0$.)

More precisely:

Lemma 3. *Let $p \geqslant 1$ and s arbitrary real. Let $\phi_p(s)$ denote the number of distinct $x_p = n_k$, $y_p = n_l$ pairs which occur among the solutions of*

$$s-\tfrac{1}{2} \leqslant A(x, y) \leqslant s+\tfrac{1}{2} \tag{5}$$

where the x's and y's take the values $n_1 = 1, n_2, \dots, n_N$ and are subject to the conditions $x_1 \leqslant x_2 \leqslant \dots \leqslant x_p$ and $y_1 \leqslant y_2 \leqslant \dots \leqslant y_p$. Then we have

$$\phi_p(s) \leqslant 8\,pN\,\frac{\log(1+q)}{\log q}. \tag{6}$$

Proof of Lemma 3. Since $\phi_p(s) = \phi_p(-s)$ we may assume that $s \geqslant 0$. We must distinguish between two types of solutions; 1⁰ those for which $x_p \leqslant 2s+1$ and 2⁰ those for which $x_p > 2s+1$. Let us consider solutions of the first kind; let $x_p = n_k$ be a possibility. Then using (5) we get $(s-\frac{1}{2}) \leqslant pn_k$, and so

$$\frac{s-\frac{1}{2}}{p} \leqslant n_k \leqslant 2s+1.$$

From this we conclude for $s \geqslant \frac{3}{2}$

$$\frac{s-\frac{1}{2}}{p} \leqslant n_k \leqslant 4\,(s-\tfrac{1}{2})$$

and for $0 \leqslant s \leqslant \frac{3}{2}$ we obtain $1 \leqslant n_k \leqslant 4$. Hence we may use (1) in both cases and get in either case

$$\sum_{x_p \leqslant 2s+1} 1 \leqslant \frac{\log 4pq}{\log q} < 4p\,\frac{\log(1+q)}{\log q}.$$

For a fixed value of $x_p = n_k$ there are at most N choices for y_p (namely $n_1, n_2, \dots, n_N$) and so

$$\sum_{\substack{(x_p, y_p) \\ x_p \leqslant 2s+1}} 1 \leqslant 4\,pN\,\frac{\log(1+q)}{\log q}.$$

Next we consider the solutions of the second kind. Here we first estimate the number of possible y_p's. Let $y_p = n_l$ be a possibility. Then using $x_p > 2s+1$ we obtain from (5)

$$x_p \leqslant py_p + s + \tfrac{1}{2} < py_p + \frac{x_p}{2},$$

that is to say $x_p/2p < y_p$. On the other hand we obtain again from (5)

$$y_p \leqslant px_p + \tfrac{1}{2} - s \leqslant px_p + \tfrac{1}{2} < 2px_p.$$

Hence $x_p/2p \leqslant y_p = n_l \leqslant 2px_p$. Consequently (1) can be used and it follows that if $x_p > 2s+1$ then

$$\sum_{y_p} 1 \leqslant \frac{\log 4p^2 q}{\log q} < 4p \frac{\log(1+q)}{\log q}.$$

There are at most N possibilities for $x_p > 2s+1$, hence

$$\sum_{\substack{(x_p y_p) \\ x_p > 2s+1}} 1 \leqslant 4pN \frac{\log(1+q)}{\log q}.$$

This establishes the inequality (6).

Now we are able to prove two inequalities concerning the number of solutions of $A(x, y) = 0$ and of $s - \frac{1}{2} \leqslant A(x, y) \leqslant s + \frac{1}{2}$; $A(x, y) \neq 0$. These inequalities form the basis of the whole proof of the law of the iterated ogarithm. They read as follows.

Lemma 4. *Let p; $1 \leqslant p \leqslant N$ be a positive integer and let $\phi(p, N)$ denote the number of solutions of the diophantine equation $A(x, y) = 0$. Furthermore let $\phi(p, N, s)$; s = arbitrary real number, denote the number of those solutions of the inequality $s - \frac{1}{2} \leqslant A(x, y) \leqslant s + \frac{1}{2}$ for which $A(x, y) \neq 0$. In both cases $x_1, \ldots, x_p$ and $y_1, \ldots, y_p$ can take the values $n_1 = 1, n_2, \ldots, n_N$ and they are subject to the conditions $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_p$ and $y_1 \leqslant y_2 \leqslant \ldots \leqslant y_p$.*

Then there exists a positive constant $c = c(q)$ independent of p, s and the choice of the sequence $n_1 = 1, n_2, \ldots, n_N$ such that

$$\binom{N}{p} \leqslant \phi(p, N) \leqslant \binom{N}{p} + (cp)^p N^{p-1} \tag{7}$$

and

$$0 \leqslant \phi(p, N, s) \leqslant (cp)^p N^{p-1} \tag{8}$$

for every $1 \leqslant p \leqslant N$ and real s.

Remark. The following proof shows that (7) and (8) holds for any $c > 0$ which satisfies (4) and

$$c \geqslant 32 \frac{\log(1+q)}{\log q}. \tag{9}$$

Hence $c(q)$ is clearly independent of p, s, and $\{n_\nu\}$.

We shall prove the inequalities (7) and (8) by induction on p. First let $p = 1$. In this case $\phi(1, N) = N$ and so (7) is obviously true. For $p = 1$ (8) had been established in Lemma 2, inequality (3). Hence $c(q) > 0$ must satisfy (4).

Induction step on (7). Let us assume now that (8) is true for

1, 2, …, $(p-1)$ and let us prove (7) for $p \geqslant 2$. Let ν $(2 \leqslant \nu \leqslant p)$ be fixed and let $\phi^{(\nu)}$ denote the number of those solutions of

$$x_1+x_2+\ldots+x_p=y_1+y_2+\ldots+y_p \tag{10}$$

which satisfy $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_p$ and $y_1 \leqslant y_2 \leqslant \ldots \leqslant y_p$ and also the additional condition

$$x_\nu \neq y_\nu \text{ and } x_{\nu+1}=y_{\nu+1}, x_{\nu+2}=y_{\nu+2}, \ldots, x_p=y_p. \tag{11}$$

Using Lemma 3 we can estimate the number of possible x_ν, y_ν pairs. For, (10) and (11) imply

$$-\tfrac{1}{2}<A(x,y)=(x_1+\ldots+x_\nu)-(y_1+\ldots+y_\nu)<\tfrac{1}{2},$$

thus (5) is satisfied for every solution and so by (6)

$$\sum_{(x_\nu, y_\nu)} 1 \leqslant 8\nu N \frac{\log(1+q)}{\log q}. \tag{12}$$

Now we fix one possible x_ν, y_ν pair; $x_\nu=n_k$ and $y_\nu=n_l$, say. Let $\phi^{(\nu)}(n_k, n_l)$ denote the number of those solutions of (10) and (11) for which $x_\nu=n_k$ and $y_\nu=n_l$. Obviously

$$\phi^{(\nu)}=\sum_{\substack{x_\nu=n_k\\ y_\nu=n_l}} \phi^{(\nu)}(n_k, n_l). \tag{13}$$

In order to estimate $\phi^{(\nu)}(n_k, n_l)$ we consider the equation system

1^0 $\quad x_1+x_2+\ldots+x_{\nu-1}=y_1+y_2+\ldots+y_{\nu-1}+(n_l-n_k)$

2^0 $\quad x_1 \leqslant x_2 \leqslant \ldots \leqslant x_{\nu-1};\ y_1 \leqslant y_2 \leqslant \ldots \leqslant y_{\nu-1}.$

3^0 $\quad x_{\nu+i}=y_{\nu+i}\ (i=1, 2, \ldots, p-\nu)$

It is obvious that $\phi^{(\nu)}(n_k, n_l)$ is majorized by the number of solutions of this system. However $n_l \neq n_k$, and so the number of solutions of 1^0 subject to the condition 2^0 is at most $\phi(\nu-1, N, n_l-n_k)$. The number of solutions of 3^0 is exactly $\binom{N}{p-\nu}$. Hence using our assumption (8) we get

$$\phi^{(\nu)}(n_k, n_l) \leqslant \phi(\nu-1, N, n_l-n_k)\binom{N}{p-\nu} \leqslant c^{\nu-1}\nu^{\nu-1}N^{p-2}.$$

Now we use (12) and (13) and obtain the estimate:

$$\phi^{(\nu)} \leqslant c^{\nu-1}\nu^{\nu-1}N^{p-2}\sum_{\substack{x_\nu=n_k\\ y_\nu=n_l}} 1 \leqslant 8\,\frac{\log(1+q)}{\log q}\,c^{\nu-1}\nu^\nu N^{p-1}.$$

Consequently $\phi^{(\nu)} \leqslant \frac{1}{4}(c\nu)^\nu N^{p-1}$, provided $c(q)$ satisfies (9). Finally we sum with respect to $\nu=2, 3, \ldots, p$:

$$\phi'=\sum_{\nu=2}^{p}\phi^{(\nu)} \leqslant \tfrac{1}{4}c^pN^{p-1}\sum_{\nu=2}^{p}\nu^\nu < \tfrac{1}{2}(cp)^pN^{p-1}.$$

Here ϕ' denotes those solutions of (10) for which at least one $x_\nu \neq y_\nu$. The number of remaining solutions is between $\binom{N}{p}$ and $\binom{N}{p}+N^{p-1}p^p$, hence in fact

$$\binom{N}{p} \leqslant \phi(p, N) \leqslant \binom{N}{p}+(cp)^pN^{p-1}.$$

Induction step on (8). We assume again that (8) is true for $1, 2, \ldots, (p-1)$ and we prove (8) for $p \geqslant 2$. We are interested in those solutions of

$$s-\tfrac{1}{2} \leqslant A(x, y)=(x_1+x_2+\ldots+x_p)-(y_1+y_2+\ldots+y_p) \leqslant s+\tfrac{1}{2}$$

for which $A(x, y) \neq 0$ and $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_p$ and $y_1 \leqslant y_2 \leqslant \ldots \leqslant y_p$. We must distinguish between two types of solutions: 1^0 those for which $x_1=y_1$ and 2^0 those for which $x_1 \neq y_1$.

If $x_1=y_1$ is fixed we have

$$s-\tfrac{1}{2} \leqslant (x_2+\ldots+x_p)-(y_2+\ldots+y_p) \leqslant s+\tfrac{1}{2}$$

and $x_2+\ldots+x_p \neq y_2+\ldots+y_p$. Hence using our assumption, for fixed $x_1=y_1$ the number of solutions is at most $\phi(p-1, N, s) \leqslant c^{p-1}p^pN^{p-2}$. There are N possibilities for $x_1=y_1$, hence the number of solutions of the first type is

$$\phi_1 \leqslant c^{p-1} p^p N^{p-1}.$$

Now we consider the solutions of the second type: If $x_{\nu+1}, \ldots, x_p$ and $y_{\nu+1}, \ldots, y_p$ are fixed then the number of possible x_ν, y_ν pairs can be estimated by Lemma 3, inequality (6). Hence using (9) the number of choices for $x_2, x_3, \ldots, x_p$ and $y_2, y_3, \ldots, y_p$ can be estimated by

$$\sum_{\substack{x_2,\ldots,x_p \\ y_2,\ldots,y_p}} 1 \leqslant \left(8\,\frac{\log(1+q)}{\log q}\right)^{p-1} p^{p-1} N^{p-1} \leqslant \tfrac{1}{4}\, c^{p-1} p^p N^{p-1}.$$

For fixed $x_2, \ldots, x_p$ and $y_2, \ldots, y_p$ the number of possible $x_1 \neq y_1$ pairs can be estimated by Lemma 2, inequality (3). Hence the number of solutions of the second type is at most

$$\phi_2 \leqslant \tfrac{1}{4}\, c^p p^p N^{p-1}.$$

Finally $\phi(p, N, s)=\phi_1+\phi_2 \leqslant (cp)^pN^{p-1}$. This completes the proof of Lemma 4.

In the following section we need a slightly modified form of Lemma 4. Namely we must drop the conditions $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_p$ and $y_1 \leqslant y_2 \leqslant \ldots \leqslant y_p$. Since every solution leads to at most $p!^2$ solutions when we drop the additional conditions we get immediately from (7) and (8) the following final estimates:

Lemma 5. *Let* $1 \leqslant p \leqslant N$ *and*

$$A(x, y)=(x_1+x_2+\ldots+x_p)-(y_1+y_2+\ldots+y_p),$$

where $x_1, \ldots, x_p$ *and* $y_1, \ldots, y_p$ *are restricted to the values* $n_1=1, n_2, \ldots, n_N$. *Then there exists a* $c=c(q)>0$ *independent of* p, s *and the sequence* $n_1=1, n_2, \ldots, n_N$ *such that*

$$p!^2\binom{N}{p} \leqslant \sum_{A(x,y)=0} 1 \leqslant p!^2\binom{N}{p}+(cp)^{3p}N^{p-1}, \tag{14}$$

and for every real s

$$0 \leqslant \sum_{\substack{s-\frac{1}{2} \leqslant A(x,y) \leqslant s+\frac{1}{2} \\ A(x,y) \neq 0}} \leqslant (cp)^{3p} N^{p-1}. \tag{15}$$

2. *On the measure of the set where the exponential sum lies between given limits*

Let throughout this section $0<n_1<\ldots<n_N$ be a fixed sequence of real numbers satisfying the lacunarity condition $n_{\nu+1}/n_\nu \geqslant q>1$ $(1 \leqslant \nu<N)$. (Notice that the condition $n_1=1$ has been dropped.) First we give an asymptotic expression for the value of the integral

$$I=\int_{\beta}^{\alpha} |\exp 2\pi i n_1 x+\exp 2\pi i n_2 x+\ldots+\exp 2\pi i n_N x|^{2p}\, dx \tag{16}$$

when $p=O(\log\log N)$; $\beta-\alpha \geqslant 1/n_1 \sqrt{N}$ and $N \to \infty$. Using this asymptotic expression we are able to obtain sharp upper and lower bounds for

$$\phi(t)=\operatorname{meas} E\{x|\alpha \leqslant x \leqslant \beta;\ F(N;x) \geqslant \sqrt{t\, N \log\log N}\}, \tag{17}$$

where for simplicity

$$F(N;x)=|\exp 2\pi i n_1 x+\ldots \exp 2\pi i n_N x|.$$

First we prove the following:

Lemma 6. *Let α, β be real and such that $\beta-\alpha \geqslant 1/n_1 \sqrt{N}$. Furthermore let p be a positive integer satisfying $1 \leqslant p \leqslant 3 \log\log N$. Then*

$$|I-(\beta-\alpha)p!\, N^p| \leqslant (\beta-\alpha)N^{p-\frac14} \tag{18}$$

for every $N \geqslant N_0(q)$ where $N_0(q)$ is independent of α, β, p and the sequence $n_1, n_2, \ldots, n_N$.

Proof of Lemma 6. We have from (16)

$$I=\sum_{1\leqslant k_\nu, l_\nu \leqslant N} \int_{\alpha}^{\beta} \exp 2\pi i \sum_{\nu=1}^{p} (n_{k_\nu}-n_{l_\nu})\, x\, dx.$$

Hence introducing the notations of the previous section we have

$$I=(\beta-\alpha)\sum_{A(x,y)=0} 1+\sum_{\substack{-(n_1/2)\leqslant A(x,y)\leqslant n_1/2\\ A(x,y)\neq 0}} \int_{\alpha}^{\beta} \exp 2\pi i A(x,y)\,\xi\, d\xi$$

$$+\sum_{s\neq 0}{}' \sum_{\substack{n_1(s-\frac12)\leqslant A(x,y)\leqslant n_1(s+\frac12)\\ A(x,y)\neq 0}} \int_{\alpha}^{\beta} \exp 2\pi i A(x,y)\,\xi\, d\xi.$$

The dash in $\sum'$ indicates that there are at most N^{2p} distinct values for $s=\pm 1, \pm 2, \ldots$ for which the contribution is not zero.

In the first sum we estimate the integral by

$$\left|\int_{\alpha}^{\beta} \exp 2\pi i A\xi\, d\xi\right| \leqslant \beta-\alpha$$

and in the second sum we use the inequality

$$\left|\int_{\alpha}^{\beta} \exp 2\pi i A\xi\, d\xi\right| \leqslant \frac{1}{\pi|A|}.$$

Hence applying the inequalities (14) and (15) of Lemma 5 we obtain

$$\left|I-(\beta-\alpha)\,p!^2\binom{N}{p}\right| \leqslant (\beta-\alpha)\,(cp)^{3p}\,N^{p-1}+(\beta-\alpha)\sum_{-(n_1/2)\leqslant A(x,y)\leqslant n_1/2} 1$$

$$+\sum_{s\neq 0}{}'\sum_{n_1(s-\frac{1}{2})\leqslant A(x,y)\leqslant n_1(s+\frac{1}{2})} A(x,y)^{-1}$$

$$\leqslant 2\,(\beta-\alpha)\,(cp)^{3p}\,N^{p-1}+\frac{(cp)^{3p}\,N^{p-\frac{1}{2}}}{n_1\sqrt{N}}\sum_{s\neq 0}{}'\frac{1}{|s-\frac{1}{2}|}.$$

Since there are at most N^{2p} distinct choices for $s=\pm 1, \pm 2, \ldots$ in $\sum'$ we get from above

$$\left|I-(\beta-\alpha)\,p!^2\binom{N}{p}\right| \leqslant 6\,(\beta-\alpha)\,(cp)^{3p+1}\,N^{p-\frac{1}{2}}\log N,$$

provided $\beta-\alpha \geqslant 1/n_1\sqrt{N}$. If p satisfies the inequality $1\leqslant p\leqslant 3\log\log N$ then the right hand side is less than $\frac{1}{2}(\beta-\alpha)N^{p-\frac{1}{2}}$ for $N\geqslant N_0(c)=N_0(q)$ and $p!^2\binom{N}{p}$ can be replaced by $p!\,N^p$ which introduces an error less than $\frac{1}{2}N^{p-\frac{1}{2}}$. Hence (18) follows.

Now we prove the following upper bounds for the measure $\phi(t)\,(0\leqslant t\leqslant N)$ defined in (17):

Lemma 7. *We have*

$$(19)\qquad \phi(t)\leqslant\begin{cases}(\beta-\alpha)\,\dfrac{18\log\log N}{(\log N)^t} & \text{for } 0\leqslant t\leqslant 3, \text{ and}\\[2mm] (\beta-\alpha)\,\dfrac{6\log\log N}{t^{2\log\log N}} & \text{for } 3\leqslant t\leqslant N\end{cases}$$

provided $\beta-\alpha\geqslant 1/n_1\sqrt{N}$ and $N\geqslant N_0(q)$ where $N_0(q)$ is indenpendent of α, β and the sequence $n_1, n_2, \ldots, n_N$.

Proof of Lemma 7. Obviously we have

$$\phi(t)\leqslant\int_{(F^2\geqslant t\log\log N)}\left(\frac{F(N;x)}{\sqrt{tN\log\log N}}\right)^{2p}dx\leqslant\frac{I}{(tN\log\log N)^p}$$

for any $t>0$, $p=1, 2, \ldots$. Hence using (18) and replacing $p!$ by $p(p/e)^p$ it follows that

$$\phi(t)\leqslant 2(\beta-\alpha)p(p/et\log\log N)^p,$$

provided $N\geqslant N_0(q)$, $p_0\leqslant p\leqslant 3\log\log N$ and $\beta-\alpha\geqslant 1/n_1\sqrt{N}$.

If $0<t\leqslant 3$ we choose $p=[t\log\log N]$ and obtain

$$\phi(t)\leqslant 6\,(\beta-\alpha)\,(\log\log N)\,e^{-[t\log\log N]}<\frac{18(\beta-\alpha)\log\log N}{(\log N)^t}.$$

If $t\geqslant 3$ choose $p=[e\log\log N]$ and get

$$\phi(t)<6\,(\beta-\alpha)\,(\log\log N)\,t^{-[e\log\log N]}<\frac{6(\beta-\alpha)\log\log N}{t^{2\log\log N}}.$$

This proves the statement of Lemma 7.

Using (18) and (19) we can find a lower bound for $\phi(t)$; $0<t<1$. Namely we prove the following:

Lemma 8. *Let ε, $0<\varepsilon<1$, be arbitrary. Then*

$$(20)\qquad \phi(1-\varepsilon)>\frac{\beta-\alpha}{(\log N)^{1-\varepsilon^2/4}}$$

for any α, β *satisfying* $\beta-\alpha \geqslant 1/n_1\sqrt{N}$ *and every* $N \geqslant N_0(q, \varepsilon)$. *This bound* $N_0(q, \varepsilon)$ *is independent of* α *and* β.

Proof of Lemma 8. For the sake of simplicity let

$$R(x) = F(N; x)^2/N \log\log N.$$

Let us introduce the following subsets of the interval $\alpha \leqslant x \leqslant \beta$:

$$\begin{aligned} E &= \{x|\alpha \leqslant x \leqslant \beta;\ 1-\varepsilon \leqslant R(x) \leqslant 1\} \\ E_1 &= \{x|\alpha \leqslant x \leqslant \beta;\ 0 < R(x) < 1-\varepsilon\} \\ E_2 &= \{x|\alpha \leqslant x \leqslant \beta;\ 1 < R(x) \leqslant 3\} \\ E_3 &= \{x|\alpha \leqslant x \leqslant \beta;\ 3 < R(x) \leqslant N\}. \end{aligned}$$

According to Lemma 6

$$\int_\alpha^\beta R(x)^p\,dx > (\beta-\alpha)\left(\frac{p}{e\log\log N}\right)^p$$

for $1 \leqslant p \leqslant 3\log\log N$, $\beta-\alpha \geqslant 1/n_1\sqrt{N}$ and $N \geqslant N_0(q)$. Hence

$$\begin{aligned} \phi(1-\varepsilon) &\geqslant \operatorname{meas} E \geqslant \int_E R(x)^p\,dx \geqslant \\ &\geqslant (\beta-\alpha)\left(\frac{p}{e\log\log N}\right)^p - \Big(\int_{E_1} + \int_{E_2} + \int_{E_3}\Big) R(x)^p\,dx, \end{aligned}$$

that is to say

$$\phi(1-\varepsilon)/(\beta-\alpha) \geqslant \left(\frac{p}{e\log\log N}\right)^p - (I_1+I_2+I_3), \tag{21}$$

where

$$I_i = \frac{1}{\beta-\alpha}\int_{E_i} R(x)^p\,dx \qquad (i = 1, 2, 3).$$

We choose $p = [(1-(\varepsilon/2))\log\log N]$ and estimate I_1, I_2 and I_3 from above. First of all using Lemma 7 we obtain

$$\begin{aligned} I_1 = -\int_0^{1-\varepsilon} t^p\,d\varphi(t) &\leqslant \frac{2p}{\beta-\alpha}\int_0^{1-\varepsilon} t^{p-1}\phi(t)\,dt \leqslant 2p\int_0^{1-\varepsilon} t^{p-1}\frac{18\log\log N}{(\log N)^t}\,dt = \\ &= 36\,p(\log\log N)^{1-p}\int_0^{(1-\varepsilon)\log\log N} u^{p-1}e^{-u}\,du. \end{aligned}$$

Since $u^{p-1}e^{-u}$ has its maximum at $u = p-1$ and $(1-\varepsilon)\log\log N \leqslant p-1$ we get for $N \geqslant N_0(q, \varepsilon)$

$$\begin{aligned} I_1 &< 36\,(\log\log N)^2\,(1-\varepsilon)^p\,e^{-(1-\varepsilon)\log\log N} \\ &< 72\,(\log\log N)^2\,(1-\varepsilon)^{(1-(\varepsilon/2))\log\log N}\,(\log N)^{-(1-\varepsilon)}. \end{aligned}$$

Finally for $N \geqslant N_0(q, \varepsilon)$

$$I_1 \leqslant \frac{72(\log\log N)^2}{(\log N)^\theta} \tag{22}$$

where

$$\theta = 1-\varepsilon-\left(1-\frac{\varepsilon}{2}\right)\log(1-\varepsilon).$$

Next we estimate I_2 by using the same procedure:

$$I_2 < 36(\log\log N)^{1-p}\,p\int_{\log\log N}^{3\log\log N} u^{p-1}e^{-u}\,du.$$

Since $p-1<u=\log\log N$ we obtain

$$I_2 < \frac{72\,(\log\log N)^2}{\log N} \tag{23}$$

for all $N \geqslant N_0(q,\varepsilon)$.

In order to estimate I_3 we proceed in a similar way, but we must apply the second, weaker estimate of Lemma 7:

$$I_3 \leqslant 2p\int_3^N t^{p-1}\,\frac{6\log\log N}{t^{2\log\log N}}\,dt \leqslant 12\,(\log\log N)\,t^{p-2\log\log N}\Big|_N^3$$
$$< 12\,(\log\log N)\,e^{-\log\log N},$$

so that

$$I_3 < \frac{12\log\log N}{\log N}. \tag{24}$$

Now we combine the inequalities (21), (22), (23) and (24). Since $p=[(1-(\varepsilon/2))\log\log N]$ we have for $N \geqslant N_0(q,\varepsilon)$

$$\left(\frac{p}{e\log\log N}\right)^p \geqslant \left(1-\frac{2}{\log\log N}\right)^p\left(\frac{1-(\varepsilon/2)}{e}\right)^p$$
$$\geqslant \left(1-\frac{2}{\log\log N}\right)^{\log\log N}\left(\frac{1-(\varepsilon/2)}{e}\right)^{(1-(\varepsilon/2))\log\log N}.$$

Hence

$$\left(\frac{p}{e\log\log N}\right)^p > \frac{1}{9\,(\log N)^{\vartheta}},$$

where

$$\vartheta = \left(1-\frac{\varepsilon}{2}\right) - \left(1-\frac{\varepsilon}{2}\right)\log\left(1-\frac{\varepsilon}{2}\right).$$

An easy computation shows that $\vartheta<1-(\varepsilon^2/4)$ for $0<\varepsilon<1$, hence we have from (23) and (24)

$$\frac{1}{2}\left(\frac{p}{e\log\log N}\right)^p > I_2 + I_3$$

for $N \geqslant N_0(q,\varepsilon)$. Consequently we obtain from (21) and (22)

$$\frac{\phi(1-\varepsilon)}{\beta-\alpha} \geqslant \frac{1}{18\,(\log N)^{\vartheta}} - \frac{72\,(\log\log N)^2}{(\log N)^{\theta}}.$$

In order to establish the statement of the lemma it is sufficient to show that $\vartheta<\theta$, i.e.

$$\frac{\varepsilon}{2} < \left(1-\frac{\varepsilon}{2}\right)\log\frac{1-(\varepsilon/2)}{1-\varepsilon},$$

where $0<\varepsilon<1$. This last inequality clearly holds, as can be seen by expanding each side of the inequality in a power series.

This establishes Lemma 8.

(To be continued)