UNIFORM DISTRIBUTION IN THE \((3x + 1)\)-PROBLEM

YA. G. SINAI

Dedicated to S. P. Novikov on the occasion of his 65th birthday

Abstract. Structure theorem of the \((3x + 1)\)-problem claims that the images under \(T^n\) of arithmetic progressions with step \(2^k\) are arithmetic progressions with step \(3^m\). Here \(T\) is the basic underlying map and a given \(3^m\) progression can be the image of many different \(2^k\) progressions. This gives rise to a probability distribution on the space of \(3^m\) progressions. In this paper it is shown that this distribution is in a sense close to the uniform law.

2000 Math. Subj. Class. 60c05.

Key words and phrases. \((3x + 1)\)-problem, uniform distribution, characteristic function.

1. The \((3x + 1)\)-Problem and its Main Recurrent Relation

For any odd \(x > 0\), find an integer \(k = k(x) > 0\) such that \(y = \frac{3x + 1}{2^k}\) is again odd. In this way we get a map \(T, Tx = y\), which actually can be considered as acting on the set \(\Pi\) of positive integers not divisible by 2 and 3. We shall write \(T(k)\) if we need the dependence on \(k\). The famous \((3x + 1)\)-problem asks whether it is true that for each \(x \in \Pi\) one can find an \(n(x)\) such that \(T^{n(x)}x = 1\).

There exists a large literature devoted to the \((3x + 1)\)-problem. We shall mention only the widely known survey by J. Lagarias [L] and the book by G. Wirsching [W]. The present text is closely connected with my paper [S], where the following theorem was proved.

Let us write \(\Pi = 1 \cup \Pi^+ \cup \Pi^-\), \(\Pi^+ = \{6p + 1\}, \Pi^- = \{6p - 1\}, p \geq 1\). Take an \(m > 0\), \(\epsilon = \pm 1\), and integers \(k_1, k_2, \ldots, k_m, k_i \geq 1\). We want to describe the set \(\Sigma(k_1, \ldots, k_m, \epsilon)\) of all \(x \in \Pi^+\) to which one can successfully apply \(T^{(k_1)}, T^{(k_2)}, \ldots, T^{(k_m)}\). Denote \(k = k_1 + k_2 + \cdots + k_m\).

**Structure Theorem** (see [S]). There exists a \(q_m(k_1, \ldots, k_m, \epsilon) = q_m, 0 \leq q_m \leq 2^k\), such that the set in question has the form

\[
\Sigma(k_1, \ldots, k_m, \epsilon) = \{6(2^p + q_m) + \epsilon, p \geq 0\}. \tag{1}
\]
Moreover, for some \( r_m = r_m(k_1, \ldots, k_m, \epsilon) \), \( 0 \leq r_m \leq 3^m \), and \( \delta(k_1, \ldots, k_m, \epsilon) = \delta_m = \pm 1 \),

\[
T_m^{\delta_m} (k_1, \ldots, k_m, \epsilon) = \Lambda^{(r_m, \delta_m)} = \{ 6(3^m p + r_m) + \delta_m \}
\]

(2)

and \( p \) in (1) and (2) is the same.

The proof goes by induction on \( m \) (see [S]).

The number of possible \((q_m, \epsilon)\) equals \( 2^{k-1}(m-1) \), where \( (k-1) \) is the number of all solutions of the equation \( k = k_1 + \cdots + k_m, k_i \geq 1 \). The factor 2 is related to the set of possible values of \( \epsilon \). We have \( (k-1) = 2^{k-1}G(k, m) \), where \( G(k, m) \) has

Gaussian asymptotics in the domain \( k - 2m = O(\sqrt{m}) \). In the domain \( k \sim 2m \), the number of possible \((k_1, \ldots, k_m)\) grows in a weak sense as \( 2^{2m} \). Therefore the number of possible \( q_m \) also grows as \( 2^{2m} \). On the other hand, the number of possible \((r_m, \delta_m)\) equals \( 2 \cdot 3^m \) and it is natural to expect that the number of the \((q_m, \epsilon)\) corresponding to a typical \((r_m, \delta_m)\) via the Structure Theorem grows in a weak sense also as \( 2^{2m}/3^m \). The purpose of this paper is to give a precise meaning to this expectation.

We shall use the notation \( \Phi_m(q_m, \epsilon) = (r_m, \delta_m) \) if \( T_m^{\delta_m} (\sum(k_1, \ldots, k_m, \epsilon)) = \Lambda^{(r_m, \delta_m)} \).

In this way \( T_j^{\delta_m} (\sum(k_1, \ldots, k_j, \epsilon)) = \Lambda^{(r_j, \delta_j)} \), \( j = 1, 2, \ldots, m \). It is clear that \( \sum(k_1, \ldots, k_j, \epsilon) \subset \sum(k_1, \ldots, k_{j+1}, \epsilon) \subset \cdots \subset \sum(k_1, \ldots, k_m, \epsilon) \).

The values \( \delta_j \) can be found from \( k_j \) via the relation (see [S] and below)

\[
2^{k_j} \equiv \delta_j \pmod{3}.
\]

(3)

In other words, \( T_j^{\delta_m} (\sum(k_1, \ldots, k_j, \epsilon)) \subset \Pi^{-1}(\Pi + 1) \) for odd (even) \( k_j \).

The sequence \((r_j, \delta_j), j = m, m - 1, \ldots, 1\), can be considered as a trajectory of a kind of random walk starting from \((r_m, \delta_m)\). As was shown in [S], the pairs \((r_j, \delta_j)\) satisfy the sequence of recurrent relations

\[
2^{k_j} r_j - 3^j t_j + \frac{2^{k_j} \delta_j - 1 - 3 \delta_{j-1}}{6} = 3 r_{j-1},
\]

(4)

Here \( t_j \) are integers, \( 0 \leq t_j < 2^{k_j} \). The ratio \( c_j = c(k_j, \delta_j-1) = 2^{k_j} \delta_j - 1 - 3 \delta_{j-1} \) is an integer only if \( 2^{k_j} \delta_j \equiv 1 \pmod{3} \) (see (3)).

We call (4) the main recurrent relation of the \((3x + 1)\)-problem. For given \( k_j \) and \( \delta_j-1 \), the left-hand side of (4) must be divisible by 3, and this is the restriction on the possible values of \( k_j \). For \( j = m \), we have

\[
\frac{2^{k_m} r_m + c_m}{3} = \frac{t_m 3^{m-1} - r_{m-1}}{r_{m-1}}.
\]

(5)

and this shows that \( k_m \) should be such that \( 2^{k_m} r_m + c_m \) is divisible by 3. Therefore, for a given \((r_m, \delta_m)\), equation (3) determines the parity of \( k_m \), and \( k_m \) must belong to one of the six arithmetic progressions \( \Gamma_j = \{ j + 6p, p \geq 0 \}, 1 \leq j \leq 6 \), provided that \( \delta_{m-1} \) is given. The values \( j = 1, 3, 5 \) correspond to \( \delta_m = -1 \), while the others correspond to \( \delta_m = 1 \). Let us write

\[
c_j = c_j(k_j, \delta_j-1) = \frac{\delta_j - \delta_{j-1}}{2} + \delta_j g(k_j).
\]

(6)
2. Expressions for Solutions of (4) and the Formulation of the Main Result

We need the triadic decomposition of \( r_m \): \( r_m = h_m(0) + h_m(1) \cdot 3 + h_m(2) \cdot 3^2 + \cdots + h_m(m-1)3^{m-1} \), where \( h_m(j) \) take values 0, 1, 2. Also we use the notation

\[
\begin{align*}
&f_{m-s} = f_m(h_m(0), \ldots, h_m(s-1); k_m, k_{m-1}, \ldots, k_{m-s+1}, \delta_{m-s}) \\
&= \frac{2^{k_{m-s}} 2^{k_{m-1}} h_m(0) + c_m}{3} + h_m(1)2^{k_{m-1}+k_m} + c_{m-1} \\
&+ 2^{k_{m-2}+k_{m-1}+k_m} h_m(2) + c_{m-2} \\
&\quad + \cdots \\
&+ h_m(s-2)2^{k_{m-s+2}+\cdots+k_m} + c_{m-s+2} \\
&\frac{2^{k_{m-s+1}}}{3} + h_m(s-1)2^{k_{m-s+1}+\cdots+k_m} + c_{m-s+1}. \quad (7)
\end{align*}
\]

It is clear that

\[
f_{m-s} = \frac{2^{k_{m-s+1}} f_{m-s+1} + 2^{k_{m-s+1}+\cdots+k_m} h_m(s-1) + c_{m-s+1}}{3}. \quad (8)
\]

Using (7), we can write down an expression for \( r_{m-s-1} \):

\[
\begin{align*}
&2^{k_{m-s}} f_{m-s} + 2^{k_{m-s}+\cdots+k_m} h_m(s) + c_{m-s} \\
&= \frac{2^{k_{m-s}} h_m(s+1) + h_m(s+2)3 + \cdots + h_m(m-1)3^{m-s-2} + \cdots + t_m 2^{k_{m-s}+\cdots+k_m} + \cdots + t_{m-s} 3^{m-s-1}}{3} \\
&= r_{m-s-1}. \quad (9)
\end{align*}
\]

The numbers \( f_{m-s} \) depend on \( k_m, k_{m-1}, \ldots, k_{m-s+1}, h_m(0), h_m(1), \ldots, h_m(s-1) \) should be chosen in such a way that all ratios in (7) and (9) are integers. In particular, all \( f_{m-s} \) are integers. In this sense \( h_m(j) \) are functions of \( k_m, k_{m-1}, \ldots, k_{m-s+1} \). An equation involving \( \epsilon \) appears for \( s = m-1 \).

It follows from the results of [8] that there exists a natural probability distribution \( P \) on the space of all sequences \( k_1, k_2, \ldots, k_m, \epsilon \) for which all \( k_i \) are independent random variables having the geometric distribution with parameter \( \frac{1}{3} \), i.e., \( P[k_j = j] = \frac{1}{3^j}, j \geq 1, \) and \( \epsilon = \pm 1 \) with probabilities \( \frac{1}{2} \) independently of \( k_j \), \( 1 \leq j \leq m \). We can consider the probabilities \( P[h_m(0), \ldots, h_m(m-1), \delta_m] \) as the probabilities of all \( \{k_m, \ldots, k_1, \epsilon\} \) giving the pair \( \langle r_m, \delta_m \rangle \), i.e., these are the probabilities of \( \Phi_m^{-1}(r_m, \delta_m) \). Put \( \rho_m = \frac{r_m}{3^m} = h_m(m-1) + h_m(m-2) + \cdots + h_m(0) \) and denote by \( Q_m \) the induced probability distribution on the pairs \( \langle \rho_m, \delta \rangle, \delta = \pm 1 \). It is convenient to consider \( Q_m \) as concentrated on two intervals \( I_1 \) and \( I_{-1} \), depending on the value of \( \delta \). It follows from (3) that \( Q_m(I_{-1}) = P[k_m \text{ is odd}] = \frac{2}{3} \) and
\( Q_m(I_1) = P\{k_m \text{ is even}\} = \frac{1}{3} \); these can be considered as normalization conditions. The main result of this paper is the following theorem.

**Theorem 1.** The restrictions of \( Q_m \) to \( I_1 \) and \( I_{-1} \) converge weakly to the uniform distributions with the above-mentioned normalization conditions.

The proof of this theorem is given in Sections 3–5. We shall actually prove that

\[
P_m(h_m(m-1), h_m(m-2), \ldots, h_m(m-t), \delta_m) = \frac{1}{3^t} \cdot P(\delta_m),
\]

where \( P(-1) = \frac{2}{3} \) and \( P(1) = \frac{1}{3} \). For definiteness we consider the case of odd \( k_m \). The other case is considered in a similar way.

### 3. Reduction to an Estimate of a Characteristic Function

Let us write \( y_{m-s} = \frac{f_{m-s}}{2^{k_{m-1} + \cdots + k_{m-s+1}}} \). From (8),

\[
y_{m-s} = \frac{y_{m-s+1} + h_m(s-1)}{3} + z_{m-s+1},
\]

where \( z_{m-s+1} = \frac{c_{m-s+1}}{3 \cdot 2^{k_{m-1} + \cdots + k_{m-s+1}}} \). This gives

\[
y_{m-s} = \frac{h_m(s-1)}{3} + \frac{h_m(s-2)}{3^2} + \cdots + \frac{h_m(0)}{3^s} + \frac{z_{m-s+1}}{3} + \frac{z_{m-s+2}}{3} + \cdots + \frac{z_m}{3^{s-1}}.
\]

In particular,

\[
y_0 = \frac{h_m(m-1)}{3} + \frac{h_m(m-2)}{3^2} + \cdots + \frac{h_m(0)}{3^m} + \frac{z_1}{3} + \cdots + \frac{z_m}{3^{s-1}} = \frac{H}{3^m} + \frac{N}{3 \cdot 2^{k_m + \cdots + k_1}} = \frac{H \cdot 2^{k_m + \cdots + k_1} + N}{3 \cdot 2^{k_m + \cdots + k_1}}. \tag{10}
\]

Here

\[
H = h_m(0) + 3h_m(1) + \cdots + 3^{m-1}h_m(m-1) \tag{11}
\]

\[
N = c_m 2^{k_{m-1} + \cdots + k_1} + 3c_{m-1} 2^{k_{m-2} + \cdots + k_2} + \cdots + 3^{m-1} c_1. \tag{12}
\]

Put \( k(N) = k_m + \cdots + k_1 \). It is clear that \( 0 \leq H < 3^m \) and only the pairs \((H, N)\) for which \( 2^{k_m + \cdots + k_1}H + N \) are divisible by \( 3^m \) correspond to \( f_0 \), i.e., for each \( N \), there is only one value of \( H \) for which \( 2^{k_m + \cdots + k_1}H + N = f_0 \).

Consider the ensemble \( K \) of all sequences \((k_m, \ldots, k_1, \epsilon)\) with odd \( k_m \). Any such sequence determines uniquely the sequence

\[
f_{m-1} = \frac{2^{k_m} \cdot h_m(0) + c_m}{3},
\]

\[
f_{m-2} = \frac{2^{k_m-1} \cdot 2^{k_m} h_m(0) + c_m + 2^{k_m+k_m-1} h_m(1) + c_{m-1}}{3},
\]

\[
f_{m-3} = \frac{2^{k_m-2} \cdot 2^{k_m-1} \cdot 2^{k_m} h_m(0) + c_m + 2^{k_m+k_m-2} h_m(1) + c_{m-2}}{3},
\]

\[
\vdots
\]

\[
f_{1} = \frac{2^{k_m-k_1} \cdot 2^{k_m-1} \cdots h_m(0) + c_{k_1} + 2^{k_m+k_m-1} h_m(1) + c_{m-1}}{3}.
\]

The restrictions of \( q_{m-n+1} \) to \( I_1 \) and \( I_{-1} \) converge weakly to the uniform distributions with the above-mentioned normalization conditions.

The proof of this theorem is given in Sections 3–5. We shall actually prove that

\[
P_m(h_m(m-1), h_m(m-2), \ldots, h_m(m-t), \delta_m) = \frac{1}{3^t} \cdot P(\delta_m),
\]

where \( P(-1) = \frac{2}{3} \) and \( P(1) = \frac{1}{3} \). For definiteness we consider the case of odd \( k_m \). The other case is considered in a similar way.
\[ P(K_0) = \sum_k \frac{1}{2^k} \sum_{N: k(N) = k} \sum_{\lambda = 0}^{3^m-1} \exp \left\{ 2\pi i \frac{2^k H + N}{3^m} \lambda \right\}, \]

where \( \Delta \) is the set of \( H \) with fixed values of \( h_m(m-1), \ldots, h_m(m-t) \).

Indeed, the sum over \( \lambda \) equals to zero unless \( 2^k H + N \) is divisible by \( 3^m \). In this case it equals \( 3^m \). Therefore the whole expression (13) gives the probability of pairs \( (H, N) \) for which \( H \in \Delta \) and \( k_m \) is odd, i.e., \( P(K_0) \).

Change the order of summation in (13):

\[ P(K_0) = \sum_k \frac{1}{2^k} \frac{1}{3^m} \sum_{\lambda = 0}^{3^m-1} \sum_{H \in \Delta} \exp \left\{ 2\pi i \frac{2^k H \lambda}{3^m} \right\}, \]

We can write

\[ H = 3^{m-1-t}(3^t h_m(m-1) + 3^{t-1} h_m(m-2) + \cdots + h_m(m-t)) + H_1 \]

\[ = 3^{m-1-t} H_0 + H_1, \]

where \( 0 \leq H_1 < 3^{m-t} \). The last summation in (14) is reduced to summation over \( H_1 \), and the result is

\[ \exp \left\{ 2\pi i \frac{2^k H_1}{3^t} \lambda \right\} \cdot \exp \left\{ 2\pi i \frac{2^k \lambda}{3^t} \right\} - 1 \]

Thus we have to consider

\[ P(K_0) = \sum_k \frac{1}{2^k} \frac{1}{3^m} \sum_{\lambda = 0}^{3^m-1} \exp \left\{ 2\pi i \frac{2^k H_0 \lambda}{3^t} \right\} \cdot \exp \left\{ 2\pi i \frac{2^k \lambda}{3^t} \right\} - 1 \]

\[ \times \sum_{k_1, \ldots, k_m: k(N) = k} \exp \left\{ 2\pi i \frac{N}{3^m} \right\}. \]

For \( \lambda = 0 \), the whole expression in (14) or (15) equals

\[ \sum_k \frac{3^{m-t} \frac{1}{3^m}}{\sum_{N: k(N) = k \text{ and odd}}} \frac{1}{2^k} \]

i.e., corresponds to the required uniform distribution. The theorem will be proved if we show that the contributions of all sums with \( \lambda \neq 0 \) tend to zero as \( m \to \infty \).

We can consider only \( \lambda = 3^{t_1} \lambda_1 \), where \( \lambda_1 \) is not divisible by 3 and \( t_1 < t \); otherwise the corresponding term is zero. We shall estimate the sums

\[ \varphi(\lambda) = \sum_k \frac{1}{2^k} \sum_{N: k(N) = k \text{ and odd}} \exp \left\{ 2\pi i \frac{N}{3^m} \lambda \right\}. \]
They can be viewed as the characteristic functions of the random variable $N$ subject to the condition $k_m$ is odd. Our argument will use the fact that in some sense $N$ is close to the sum of independent variables; this will give us the needed estimate for $\phi(\lambda)$. For the proof of the theorem we need an estimate of

$$Q(K_0) = \frac{1}{3m} \sum_{\lambda_1}^{3m} \exp \left\{ 2\pi i \frac{2^k H_\alpha \lambda_1}{3^{i+1}} \right\} \frac{\exp \left\{ 2\pi i \frac{2^k \lambda_1}{3^{m+1}} \right\}}{\exp \left\{ 2\pi i \frac{2^k \lambda_1}{3^{m+1}} \right\}} - 1.$$  

Note that the term with $\lambda_1 = 0$ is not included in the last sum.

Assume for simplicity that $m$ is divisible by 4, i.e., $m = 4(i + 1)$. Other cases require trivial changes. Fix the values $k_1 + k_2 + k_3 + k_4 = k^{(1)}$, $k_1 + k_2 + \cdots + k_8 = k^{(2)}$, $\ldots$, $k_1 + k_2 + \cdots + k_4 = k^{(i)}$, and $k_1 + \cdots + k_{4(i+1)} = k^{(i+1)} = k$. Fix also the parities $\delta_{m-1}, \ldots, \delta_1$ of all $k_m, k_{m-1}, \ldots, k_1$. The crucial point is that, under the conditions $k^{(a)}$, $1 \leq a \leq i + 1$, and $\{\delta_m, \ldots, \delta_1\} = \delta$, the groups of four variables $k_{4j-1}, k_{4j+2}, k_{4j+3}, k_{4j+4}$ remain mutually independent with respect to $P$. For this reason, we can write

$$\phi_j(\lambda) = \sum_{k_{4j+1}, k_{4j+2}, k_{4j+3}, k_{4j+4}} \exp \left\{ 2\pi i \frac{2^{(j)} \lambda_1}{3^{j-i}} \left( 2^{k_{4j+1}+k_{4j+2}+k_{4j+3}+k_{4j+4}} \right) \right\}$$

$$+ \sum_{k_{4j+2}, 3 \cdot 2^{k_{4j+2}+k_{4j+3}+k_{4j+4}} + k_{4j+3}} \sum_{k_{4j+4}} \phi_{4+} \left( k^{(1)}, \ldots, k^{(i+1)} \right) \prod_{j=0}^i \phi_j(\lambda).$$

Here $P\{k^{(1)}, \ldots, k^{(i+1)}\}$ are the probabilities of the conditions and $\pi$ are conditional probabilities. Using $k^{(j)}$, we can rewrite the expression for $N$ (see (12)) as follows:

$$N = \left( c_m 2^{k_m+k_{m-1}+k_{m-2}+k_{m-3}} + 3c_{m-1} 2^{k_m+k_{m-1}+k_{m-2}+k_{m-3}} \right)$$

$$+ 3^2 c_{m-2} 2^{k_{m-2}+k_{m-3}} + 3^3 c_{m-3} 2^{k_{m-3}} \right) \cdot 2^{(j)}$$

$$+ \left( c_{m-4} 2^{k_m+k_{m-2}+k_{m-5}+k_{m-6}} + 3c_{m-5} 2^{k_{m-5}+k_{m-6}+k_{m-7}} \right) \cdot 2^{k^{(j-1)}}$$

$$+ 3^2 c_{m-6} 2^{k_{m+6}+k_{m-7}} + 3^3 c_{m-7} 2^{k_{m-7}} \right) \cdot 2^{(j-1)}$$

The expressions in parentheses are independent random variables with respect to the conditional distributions. For this reason, we can write the product $\prod_{j=0}^i \phi_j(\lambda)$ in the expression for $\phi(\lambda)$. It is clear that $|\phi_j(\lambda)| \leq 1$ since it is a characteristic function of some probability distribution. Denote by $A(\delta, \{k^{(a)}\}, 1 \leq a \leq i + 1)$ the set of $j$ for which $\delta_{4j} = \delta_{4j+1} = \delta_{4j+2} = \delta_{4j+3} = 5^{(i)}$, and $k_{4j+1} + k_{4j+2} + k_{4j+3} + k_{4j+4} = k^{(i+1)} = 6$. The first requirement implies that all $k_{4j+1}, \ldots, k_{4j+4}$ should be odd. Therefore three of them should be equal to 1 and the remaining one equals to 3. We can restrict ourselves to the situations when $|A(\delta, \{k^{(a)}\}, 1 \leq a \leq i + 1))| \geq c_0 m$ for some positive constant $c_0$. If $c_0$ is chosen
small enough, then the probability of the complement is less than \( \exp\{-\gamma m\} \) for another positive constant \( \gamma > 0 \), and this complement can be neglected.

It follows easily from the definitions (see (6)) that \( g(1) = 0 \), \( c(1, -1) = 0 \), \( g(3) = 1 \), and \( c(3, -1) = -2 \). For this reason, in each sum in the parentheses, only one of the four summands is non-zero. Here are the values of the corresponding sums:

\[
\begin{align*}
a_1 &= -2 & \text{if } c_{4j+1} &= -2, \\
a_2 &= -3 \cdot 2^5 & \text{if } c_{4j+2} &= -2, \\
a_3 &= -3^2 \cdot 2^4 & \text{if } c_{4j+3} &= -2, \\
a_4 &= -3^3 \cdot 2^3 & \text{if } c_{4j+4} &= -2.
\end{align*}
\]

The greatest common divisor of \( a_1, a_2, a_3, a_4 \) is \( 2^3 \).

Assume now that we have the inequality

\[
|\varphi_j(\lambda)| \geq 1 - \frac{1}{m^{\gamma_0}},
\]

where \( \gamma_0 < 1 \) is a constant, which will be chosen later. Let us write

\[
\frac{2k_j^{(i)}\lambda_1}{3^{i-1}} = B_j \frac{2^3}{2^3} + \theta_j
\]

where \( B_j \) is an integer and \( |\theta_j| < \frac{1}{2^7} \). If \( j \in A(\delta, \{k^{(a)}, 1 \leq a \leq i+1\}) \), we can write

\[
\varphi_j(\lambda) = \exp\{2\pi i a_1 \theta_j\} \cdot \pi_1 + \exp\{2\pi i a_2 \theta_j\} \cdot \pi_2 + \exp\{2\pi i a_3 \theta_j\} \cdot \pi_3 + \exp\{2\pi i a_4 \theta_j\} \cdot \pi_4,
\]

where \( \pi_1, \pi_2, \pi_3, \pi_4 \) are conditional probabilities, which are some constants. Further,

\[
\begin{align*}
|\varphi_j(\lambda)| &= |\pi_1 + \exp\{2\pi i (a_2 - a_1) \theta_j\} \cdot \pi_2 + \exp\{2\pi i (a_3 - a_1) \theta_j\} \cdot \pi_3 \\
&\quad + \exp\{2\pi i (a_4 - a_1) \theta_j\} \cdot \pi_4| \\
&= |1 + (\exp\{2\pi i (a_2 - a_1) \theta_j\} - 1) \pi_2 + (\exp\{2\pi i (a_3 - a_1) \theta_j\} - 1) \pi_3 \\
&\quad + (\exp\{2\pi i (a_4 - a_1) \theta_j\} - 1) \pi_4|.
\end{align*}
\]

All \( \exp\{2\pi i (a_p - a_1) \theta_j\} - 1 \) have negative real parts for \( p = 2, 3, 4 \). Therefore (18) can hold only if

\[
|\exp\{2\pi i \theta_j(a_p - a_1)\} - 1| \leq \frac{\text{const}}{m^{1/2\gamma_0}}, \quad p = 2, 3, 4.
\]

Since the greatest common divisor of the \( a_p \) is \( 2^3 \) and \( |\theta_j| < \frac{1}{2^7} \), the inequality (20) can hold only if \( |\theta_j| \leq \frac{\text{const}}{m^{1/2\gamma_0}} \). The const may depend on \( a_p, 1 \leq p \leq 4 \).

4. A Preliminary Estimate of \( \varphi(\lambda) \)

Let us call a \( j \) normal if \( |\varphi_j(\lambda)| < 1 - \frac{1}{m^{\gamma_0}} \). Otherwise it is called non-normal.

For a non-normal \( j_0 \in A(\delta, \{k^{(a)}, 1 \leq a \leq i+1\}) \), write

\[
\frac{2k_{j_0}^{(i)}\lambda_1}{3^{i-1}} = 3^{j_0} \cdot B_{j_0} \frac{2^3}{2^3} + \theta_{j_0},
\]

where

\[
\begin{align*}
a_1 &= -2 & \text{if } c_{4j+1} &= -2, \\
a_2 &= -3 \cdot 2^5 & \text{if } c_{4j+2} &= -2, \\
a_3 &= -3^2 \cdot 2^4 & \text{if } c_{4j+3} &= -2, \\
a_4 &= -3^3 \cdot 2^3 & \text{if } c_{4j+4} &= -2.
\end{align*}
\]
where \( p \geq 0 \) and \( B_{j_0} \) is an integer not divisible by \( 3^4 \). As our preceding analysis shows, \( |\theta_{j_0}| \leq \frac{\text{const}}{m^{1+\epsilon}} \). For \( j \geq j_0 \),

\[
\frac{2k(j)}{3^{4j-\epsilon_1}} = 3^{p-4(j-j_0)} \cdot \frac{\theta_{j_0} \cdot 2^{k(j)-k(j_0)}}{3^4} + \theta_{j_0} \tag{22}
\]

If \( j \leq p + j_0 \), the last expression gives a representation similar to (21) in which \( j \) replaces \( j_0 \). This means that \( \theta_j = \theta_{j_0} \cdot \frac{2^{k(j)-k(j_0)}}{3^{4(j-j_0)}} \). Typically \( \frac{2^{k(j)-k(j_0)}}{3^{4(j-j_0)}} \) grows with \( j \).

A segment \([j_0, j_1]\) is called a cycle if \( j_0 \) is non-normal, \( j_0 \in A(\delta, \{k(a), 1 \leq a \leq i + 1\}) \), and either \( j_1 - j_0 = p + 1 \) and \( \theta_{j_0} \cdot \frac{2^{k(j)-k(j_0)}}{3^{4(j-j_0)}} \geq \epsilon \), where \( \epsilon \) will be chosen later, or \( j_1 - j_0 \) is non-normal, \( j_0 \in A(\delta, \{k(a), 1 \leq a \leq i + 1\}) \). In this case these terms are uniformly bounded. If \( \epsilon \) is chosen small enough then the whole exponent in (23) for \( j = j_1 - 1 \) has an absolute value between \( \epsilon_1 \) and \( \epsilon_2 \) for some constants \( \epsilon_1 \) and \( \epsilon_2 \). This gives the estimate \(|\varphi_j(\lambda)| \leq \rho\) for some \( \rho \).

Lemma 1. There exists a constant \( \rho < 1 \) such that, for any typical cycle, \( \prod_{j_0 < j \leq j_1} |\varphi_j(\lambda)| \leq \rho \).

Proof. We consider two cases.

(a) \( p > j_1 - j_0 - 1 \). We can use (22). In (16) with \( j = j_1 - 1 \), the multiplication of \( 3^{p-4(j-j_0)} \cdot \frac{2^{k(j)-k(j_0)}}{3^4} \) by each \( a_k \) gives an integer. Therefore, for \( j \in A(\delta, \{k(a), 1 \leq a \leq i + 1\}) \), this part does not make any contribution to the characteristic function. For the product of \( \theta_{j_0} \cdot \frac{2^{k(j)-k(j_0)}}{3^{4(j-j_0)}} \) and all the other terms in parentheses in (16), we use the fact that \( j_1 - 1 \in A(\delta, \{k(a), 1 \leq a \leq i + 1\}) \). In this case the terms are uniformly bounded. If \( \epsilon \) is chosen small enough then the whole exponent in (16) for \( j = j_1 - 1 \) has an absolute value between \( \epsilon_1 \) and \( \epsilon_2 \) for some constants \( \epsilon_1 \) and \( \epsilon_2 \). This gives the estimate \(|\varphi_j(\lambda)| \leq \rho\) for some \( \rho \).

(b) \( p = j_1 - j_0 - 1 \). In this case, for \( j_1 = p + j_0 \), the first term of (21) no longer has factor 3. Therefore, for \( j_1 = j_0 + p + 1 \), we have the representation

\[
\frac{2k(j)}{3^{4j-\epsilon_1}} = B_{j_0} \cdot \frac{2^{k(j)-k(j_0)}}{3^{4j-\epsilon_1}} + \theta_{j_0} \cdot \frac{2^{k(j)-k(j_0)}}{3^{4(j-j_0)}} \tag{23}
\]

Write \( B_{j_0} \cdot \frac{2^{k(j)-k(j_0)}}{3^{4j-\epsilon_1}} = D_1 + \frac{D_2}{3^q} \) for some integers \( D_1, D_2 \not= 0 \). In the parentheses in (16), some terms are not divisible by \( 3^4 \). Therefore the multiplication of these terms by \( \frac{2^{k(j)-k(j_0)}}{3^{4(j-j_0)}} \) gives a number close to \( \exp\left\{2\pi i \frac{q}{3^4}\right\} \), where \( q \not= 0 \) is an integer not divisible by \( 3^4 \). In this case \(|\varphi_j(\lambda)| \leq \rho \) also. The last argument shows that \( j_1 \) cannot be greater than \( p + 1 + j_0 \). In case (b) we did not use the fact that \( j_1 - 1 \in A(\delta, \{k(a), 1 \leq a \leq i + 1\}) \). The lemma is proved. \( \square \)

Take a cycle \([j_0, j_1]\). It is clear that \( |\theta_{j_0}| \geq \frac{1}{2}m^{1+\epsilon} \). Therefore the length of this cycle is less than \( j_2 - j_0 \), where \( j_2 \) is a minimal index greater than \( j_0 \) for which

\[
\frac{2^{k(j_2)-k(j_0)}}{3^{4(j_2-j_0)}} \geq 3^{ij_0}. \tag{24}
\]

In what follows, we use the notation \( j_2(j) = j_2 \) for the least \( j_2 > j \) for which

\[
\frac{2^{k(j_2)-k(j_0)}}{3^{4(j_2-j_0)}} \geq 3^{4j}. \tag{25}
\]
Denote by $E_c$ the set of sequences $\{k_1, k_2, \ldots, k_m\}$ such that $j_2 \leq cj$ for all $j \geq \ln^2 m$. It is easy to show that, for all $c \geq c_1$, where $c_1$ is a constant,
\[
P\{E_c\} \geq 1 - \frac{1}{m^2}.
\]

We return back to our initial situation. Let $\nu(k_1, \ldots, k_m) = \nu$ be the number of normal indices $j$ (see (18)). If $\nu \geq 2 \ln m \cdot m^{\gamma_0}$, then
\[
|\varphi(\lambda)| \leq \prod_j |\varphi_j(\lambda)| \leq \left(1 - \frac{1}{m^2}\right)^\nu \leq \frac{\text{const}}{m^2}.
\]

We shall show later that this estimate is sufficient for our purposes.

Consider the case of $\nu < 2 \ln m \cdot m^{\gamma_0}$, i.e., of a small number of normal indices. We shall show that, in a typical situation, this can happen only if there are sufficiently many typical cycles. Then we apply Lemma 1 to get the needed estimate.

Denote by $[x_s, y_s]$ the $s$-th cycle from the right:
\[
\cdots < x_{s+1} < y_{s+1} < x_s < y_s < \cdots \leq m
\]

We want to estimate the number $\nu_1$ of typical cycles $[x_s, y_s]$ such that $x_s > m^{\frac{1+\gamma}{2}}$, $\{k_1, \ldots, k_m\} \in E_{c_1}$, and $|A(\delta, \{k^{(a)}_1, 1 \leq a \leq i + 1\})| \geq \text{const}$. Since $\nu < 2 \ln m \cdot m^{\gamma_0}$, the majority of points in $A(\delta, \{k^{(a)}_1, 1 \leq a \leq i + 1\})$ consists of non-typical indices. We take the first cycle $[x_1, y_1]$. According to the definitions, $y_1 \in A(\delta, \{k^{(a)}_1, 1 \leq a \leq i + 1\})$. Then we take $y_2$ to be the first $j < x_1$ such that $j \in A(\delta, \{k^{(a)}_1, 1 \leq a \leq i + 1\})$ and $j$ is non-normal. The index $x_2$ is the final point of the cycle. Next, $y_3$ is the first $j < x_2$, $j \in A(\delta, \{k^{(a)}_1, 1 \leq a \leq i + 1\})$ which is non-normal, and so on.

Suppose that an initial index $y_s$ of a cycle and all $k_j, j > y_s$, are given. Then the conditional probability of a cycle to be typical is greater than some constant $\pi_0 > 0$. Indeed, consider the two cases (a) and (b) defined in the proof of Lemma 1. In case (a), take logarithms: $(k^{(1)}_j - k^{(2)}_j) \ln 2 - 4(j - j_2) \ln 3$ is a trajectory of a random walk, and the event $x_s - 1 \in A(\delta, \{k^{(a)}_j, 1 \leq a \leq i + 1\})$ can be expressed in terms of the crossing of a distant level. It is a well-known probabilistic statement that this probability is greater than a constant. In case (b), the event $x_s - 1 \in A(\delta, \{k^{(a)}_j, 1 \leq a \leq i + 1\})$ is expressed in terms of the behaviour of $k_j$ for appropriate values of $j$.

We can consider only sequences for which $y_{s+1} - x_s \leq 2 \ln m \cdot m^{\gamma_0}$. The probability of other sequences is exceedingly small. For $m - x_s \geq m^{\frac{1+\gamma}{2}}$ and $\{k_1, \ldots, k_m\} \in E_{c_1}$,
\[
\frac{m - x_s - (m - y_{s+1})}{m - x_s} \leq \frac{2 \ln m \cdot m^{\gamma_0}}{m^{\frac{1+\gamma}{2}}} = \frac{2 \ln m}{m^{\frac{1+\gamma}{2}}}
\]
and
\[
m - x_{s+1} \leq c_1(m - y_{s+1})
\]
\[
= c_1(m - x_s + (x_s - y_{s+1})) \leq c_1(m - x_s) \left(1 + \frac{2 \ln m}{m^{\frac{1+\gamma}{2}}}\right).
\]
We can use this inequality until \( s \) reaches the value \( s_0 \) such that \( m - x_{s_0} \) becomes for the first time greater than \( \frac{m}{2} \). From (25),

\[
s_0 \geq \frac{\ln m - \ln 2}{\ln c_1 + \ln \left( 1 + \frac{2\ln m}{2} \right)} \geq \text{const} \ln m.
\]

It follows from Lemma 1 that \( |\varphi(\lambda)| \leq \rho^{s_0} \leq \frac{1}{m^{\gamma_1}} \) for some constant \( \gamma_1 > 0 \). This is our first important estimate.

5. More Refined Estimates

Let us write down again formula (15) without the term corresponding to \( \lambda = 0 \). Our purpose is to estimate its right-hand side

\[
Q(K_0) = \sum_k \frac{1}{2^k} \cdot \frac{1}{3^m} \sum_{\lambda = 1}^{3^{m-1}} \exp \left\{ 2\pi i \frac{k H_\lambda}{3^m} \right\} G \cdot \exp \left\{ 2\pi i \frac{k^3 \lambda}{3^m} \right\} - 1 \times \sum_{k_1, \ldots, k_m : k(N) = k, k_m \text{ is odd}} \exp \left\{ 2\pi i \frac{N \lambda}{3^m} \right\}.
\]

The absolute value of the sum over all \( \{k_1, \ldots, k_m\} \) and \( k \) for a fixed \( \lambda \) is not larger than 1, because it is a characteristic function. It is easy to see that the sum over all \( \lambda \) is not larger than \( \text{const} m \). This explains why estimate (24) is sufficient. Let us write down formula (21) for an initial \( j_o \) of some cycle, \( j_o \in A(\delta, \{k^{(a)}, 1 \leq a \leq i + 1\}) \):

\[
2\frac{k^{(j_o)}}{3^{j_o - 1}} \lambda_1 = 3^{4p} \cdot \frac{B_{j_o}}{2^3} + \theta_{j_o}.
\]

As our discussion in the previous section shows, the length of the cycle is not larger than \( p+1 \). We shall show that, for a “typical” \( \lambda \), there is a lot of small typical cycles. In view of Lemma 1, this implies a good estimate for \( |\varphi(\lambda)| \). On the other hand, the sum over the “non-typical” \( \lambda \) makes a small contribution to \( Q(K_0) \), because its probability is small.

Below we write \( j \) instead of \( j_o \). Suppose that \( j \) is non-normal. It follows from (27) that

\[
2\frac{k^{(j)}}{3^{j - t_1}} \lambda_1 = 3^{4p} \cdot \frac{B_j}{2^3} + \theta_j, \quad |\theta_j| \leq \text{const} \frac{1}{m^{1/2\gamma_0}}.
\]

To stress the dependence on \( \lambda_1 \) and, we shall write \( \theta_j(\lambda_1) \). It is easy to see that \( \theta_j(\lambda_1) = \theta_j(\lambda_1 + 3^{4(j+p)-t_1}) \). Indeed,

\[
2\frac{k^{(j)}}{3^{j - t_1}} (\lambda_1 + 3^{4(j+p)-t_1}) = 3^{4p} (B_j + 2^{k^{(j)}+3}) + \theta_j(\lambda_1).
\]

The sum \( (B_j + 2^{k^{(j)}+3}) \) can be divisible by some power of \( 3^4 \). For this reason,

\[
3^{4p} (B_j + 2^{k^{(j)}+3}) = \frac{3^{4p} \cdot B_j}{2^3}, \quad p_1 \geq p.
\]
and \( B_1 \) is not divisible by \( 3^4 \). Therefore (29) is similar to (28) and \( \theta_j(\lambda_1) = \theta_j(\lambda_1 + 3^4(j+p) - t_1) \). Thus the set of \( \lambda_1 \) for which (28) holds is the union of arithmetic progressions with difference \( 3^4(j+p) - t_1 \). The number of such progressions is not larger than \( 3^4j - t_1 \). Indeed, \( \theta_j \) has difference \( \frac{1}{3^4j - t_1} \) and the number of possible \( \theta_j \) within the interval \([\frac{\text{const}}{m^{3/2} \cdot 3^4}, \frac{\text{const}}{m^{1/2} \cdot 3^4}]\) is not larger than \( 2\text{const} \cdot m^{3/2} \cdot 3^4j - t_1 \).

Denote one of these progressions by \( \Pi \):

\[
\Pi = \{\lambda_1(0) + s \cdot 3^4(j+p) - t_1, 0 \leq s < 3^4m - 4(j+p - t_1)\}.
\]

The numbers \( \frac{2k\lambda_1}{3^4m} \) (mod \( 3^4m \)), \( \lambda_1 \in \Pi \), also constitute an arithmetic progression \( \Pi_1 \) with difference \( 3^{-4m-4(j+p)} \). Let \( \Lambda_1 \) be the first element of \( \Pi_1 \), i.e.,

\[
\Pi_1 = \{\Lambda_1 + s \cdot 3^{-4m-4(j+p)}, s < 3^4m-4(j+p-t_1)\}.
\]

It is clear that

\[
\frac{1}{3^m} \sum_{\lambda \in \Pi} \exp \left\{ \frac{2\pi i}{3^m} \frac{2k\lambda}{3^4m} \right\} \leq \text{const} \cdot \frac{1}{3^m} \sum_{\Lambda_1 + s \cdot 3^{-4m-4(j+p)}} \leq \text{const} \cdot \frac{1}{3^m} \cdot \Lambda_1 + \frac{\text{const}}{3^m} \cdot \Lambda_1 + \frac{\text{const}}{3^m} \cdot m\gamma \cdot \frac{1}{3^4p}.
\]

Since the total number of possible \( \Lambda_1 \) is not larger than \( \text{const} \cdot 3^4j \cdot \frac{1}{m^{1/2} \gamma} \), we get

\[
\sum \text{const} \cdot m \cdot \frac{3^4j}{3^4(j+p) \cdot m^{1/2} \gamma} \leq \text{const} \cdot \frac{m^{1/2} \gamma}{3^4p}.
\]

Take a \( p \leq \ln^2 m \). The last sum is over all progressions except the first elements. The last inequality shows that this sum is small and makes a negligible contribution to the whole sum (26). The values \( p < \ln^2 m \) are considered below.

It remains to estimate \( \frac{1}{3^m} \sum \frac{1}{\Lambda_1} \). Some of the terms \( \frac{1}{3^m \Lambda_1} \) can be small of order \( O(1) \). We argue now in a different way. Consider only \( j \leq m^{\gamma_2} \), where \( \gamma_2 \) is a constant which will be chosen later. For each \( j \) and \( p \), we have not more than \( 3^4j \gamma_2^{-1/2} \) progressions \( \Pi \). Therefore the total number of progressions is not larger than \( \text{const} \cdot 3^4j \gamma_2^{-1/2} \). The worst situation can happen when all the \( \Lambda \) occupy the whole interval \( \frac{1}{3^m} \), \( \text{const} \cdot m^{\gamma_2^{-1/2} \cdot 3^4j} \) with step \( \frac{1}{3^m} \). It is clear that, in this case,

\[
\frac{1}{3^m} \sum_{\Lambda_1} \frac{1}{\Lambda_1} \leq \text{const} \cdot j \leq \text{const} \cdot m^{\gamma_2}. \]

Now recall the estimate from the previous section, which says that \( |\varphi(\lambda)| \leq \frac{1}{m^{\gamma_3}} \) for some constant \( \gamma_3 > 0 \) and all \( \lambda \neq 0 \). If \( \gamma_2 \) is chosen less than \( \gamma_1 \), this part of the whole sum \( \frac{1}{3^m} \sum \frac{1}{\Lambda_1} \) is less than \( \text{const} \cdot m^{1/2} \). It remains to consider the situation when all cycles on the interval \([1, m^{\gamma_2}]\) are shorter than \( \ln^2 m \). Denote by \( \nu_2 \) the number of typical cycles among them. If \( \nu_2 \leq m^{\gamma_2/2} \), then the number of \( j \in A(\delta, \{k^{(a)}, 1 \leq a \leq i+1\}) \) which do not belong to any cycle and, hence, are normal is not less than \( \text{const} \cdot m^{\gamma_2} \), and therefore

\[
\prod_j |\varphi_j(\lambda)| \leq \prod_{k^{(a)}} |\varphi_j(\lambda)| \leq \left(1 - \frac{1}{m^{\gamma_2}}\right)^{\text{const} \cdot \gamma_2}.
\]
Here $\prod_{j}^{(n)}$ is the product over the normal $j \in A(\delta, \{k^{(a)}, 1 \leq a \leq i + 1\})$. Choose $\gamma_{0}$ so that $\gamma_{0} < \text{const} \cdot \gamma_{2}$. Then the last expression is less than $\exp\{-\text{const} \times m^{\text{const} \cdot \gamma_{2} - \gamma_{0}}\}$. If $\nu_{2} > m^{\gamma_{2}/2}$, then, by Lemma 1, $\prod_{j} |\varphi_{j}(\lambda)| \leq \rho_{m^{1/2\gamma_{2}}}$.

All these estimates imply the statement of the theorem.

6. Concluding Remarks

Let us describe the whole scheme of the proof of the main theorem. In Section 4, we fix a $\lambda$ and consider the summation over all possible $\{k_{1}, k_{2}, \ldots, k_{m}\}$. We show that it is enough to consider only typical $\{k_{1}, k_{2}, \ldots, k_{m}\}$ for which the number of typical cycles is sufficiently large.

This gives the estimate $|\varphi(\lambda)| \leq \frac{1}{m^{\gamma_{1}}}$ valid for all $\lambda \neq 0$. However, the set of typical $\{k_{1}, \ldots, k_{n}\}$ depends on $\lambda$.

In the second part (Section 5), we show that there is a large set of values of $\lambda$ for which there is a lot of small typical cycles on the interval $[0, m^{\gamma_{2}}]$. For these $\lambda$, we have a very good estimate of $\varphi(\lambda)$.

There is also a small set of values of $\lambda$ for which the denominator $\exp\{2\pi i \frac{k_{j}}{m_{j}}\}$ can be small. For these $\lambda$, we use the previous universal estimate $|\varphi(\lambda)| \leq \frac{1}{m^{\gamma_{1}}}$, which gives the final result.

The main theorem remains true if we consider the conditional distributions of $k_{1}, k_{2}, \ldots, k_{m}$ provided that $k_{1} + k_{2} + \cdots + k_{m} = k$ is fixed and takes a typical value, e.g., $|k - 2m| \leq \sqrt{m \ln m}$. In this case the distribution of $k_{j}$ for $j$ far from the boundary is close to the unconditional one.

Acknowledgment. I thank L. Koralov for careful reading of the text and many useful remarks.

References


Princeton University, Department of Mathematics, Princeton, NJ, and Landau Institute of Theoretical Physics, Russia
e-mail address: sinai@math.princeton.edu