STATISTICAL $(3x+1)$ – PROBLEM

Ya. G. Sinai*

Dedicated to the memory of J. Moser

Table of Contents

§1. Introduction

Take an odd number $x > 0$. Then $3x+1$ is even and one can find an integer $k > 0$ so that $y = \frac{3x+1}{2^k}$ is again odd. We get in this way the mapping $T, Tx = y$. It is clear that except being odd y is also not divisible by 3. By this reason the natural domain for T is the set $\sqcap$ of positive x not divisible by 2 and 3. The point $x = 1$ is the fixed point of T and it is the famous $(3x+1)$–problem which asks whether it is true that for every $x \in \sqcap$ one can find $n(x)$ such that $T^{n(x)}x = 1$. The best references concerning this problem are the expository paper by J. Lagarias [L] and the book by G. Wirsching [W], see also the annotated bibliography on $(3x+1)$–problem prepared by J. Lagarias. There one can find a lot of information about the history of the problem and its various modifications. We call the statistical $(3x+1)$–problem the basic question for x belonging to a subset of density 1. In this paper we discuss some version of the statistical $(3x+1)$–problem.

*Mathematics Department of Princeton University and Landau Institute of Theoretical Physics

The main result of this paper is the following Structure Theorem which we formulate now and prove in §2.

We have $\sqcap = 1 \cup \sqcap^{+1} \cup \sqcap^{-1}$ where $\sqcap^{+1} = \{6p+1, p>0\}$, $\sqcap^{-1} = \{6p-1, p>0\}$. Let us write if in the definition of T we divide $3x+1$ by 2^k. Fix $k_1, k_2, \dots, k_m$, $k_j > 0$ and integer, $1 \le j \le m$. We ask what is the set of $x \in \sqcap$ to which one can successively apply $T^{(k_1)}, T^{(k_2)}, \dots, T^{(k_m)}$. The Structure Theorem gives the answer.

Structure Theorem. *Let $k_1, \dots, k_m$ be given and $\epsilon_m = \pm 1$. The set of $x \in \sqcap^{\epsilon_m}$ to which one can apply $T^{(k_1)}, T^{(k_2)}, \dots, T^{(k_m)}$ is an arithmetic progression $\Sigma^{(k_1,\dots,k_m,\epsilon_m)}$ $= \{6 \cdot (2^{k_1+k_2+\dots+k_m} p + q_m) + \epsilon_m\}$ for some $q_m = q_m(k_1,\dots,k_m,\epsilon_m)$, $1 \le q_m \le 2^{k_1+\dots+k_m}$. The image $T^{(k_m)} \cdot T^{(k_{m-1})} \cdots T^{(k_1)}$ $(\Sigma^{(k_1,\dots,k_m,\epsilon_m)}) = \wedge^{(m)}_{r_m,\delta_m} = \{6(3^m p + r_m) + \delta_m\}$ for some $r_m = r_m(k_1,\dots,k_m,\epsilon_m)$, $1 \le r_m \le 3^m$, and $\delta_m = \delta_m(k_1,\dots,k_m,\epsilon_m) = \pm 1$. Even more, for each $p > 0$ $T^{(k_m)}, T^{(k_{m-1})} \cdots T^{(k_1)}$ $(6(2^{k_1+\dots+k_m} p + q_m) + \epsilon_m) = 6(3^m p + r_m) + \delta_m$ with the same p.*

The proof of this theorem goes by induction. First we check the statement for $m = 1$ and then derive it for $m+1$ assuming that it is true for m. A. Kontorovich has a shorter proof of this theorem.

This theorem plays the role of symbolic representation in dynamics.

In §3 we prove a simple statistical statement which follows directly from the Structure Theorem. Take $x_0 \in \sqcap$, $x_m = T^m x_0$, $y_m = \ell n\, x_m$ and $z_m = y_m - y_0$. Assume that $1 \le m \le M$ and

$$\omega\left(\frac{m}{M}\right) = \frac{z_m + m(2\ell n\, 2 - \ell n\, 3)}{\sqrt{M}}.$$

We show that $\omega(t), 0 \le t \le 1$, behave as Wiener trajectories. More precisely, let $M = 2^n$, $n \to \infty$, $\tau_1 = \frac{t_1}{2^n}$, $\tau_2 = \frac{t_2}{2^n}, \dots, \tau_s = \frac{t_s}{2^n}$ where $t_1, t_2, \dots, t_s$ are integers, $0 \le t_j < 2^n$, $1 \le j \le s$, and $\tau_1, \dots, \tau_s$ are fixed. Consider the following probability

$$P_n = P\{x_0 | a_1 \le \omega(\tau_1) \le b_1, \dots, a_s \le \omega(\tau_s) \le b_s\}$$

Here, $a_1, b_1, \dots, a_s, b_s$ are fixed numbers. The probability P of a set is understood as the density of this set wrt ?? provided that the density exists. The following theorem holds.

Theorem 3.1. *The probability P_n tends as $n \to \infty$ to the probability given by the Wiener measure with the zeroth drift and some diffusion constant $\sigma > 0$.*

In §3 we prove it for $s=1$. General case can be obtained in a similar way.

An analogous theorem was proven recently by K. A. Borovkov and D. Pfeifer (see [BP]).

In §4 we study some properties of $r_m(k_1,\dots,k_m,\epsilon_m)$. Technically the most important part is in §5 where we analyze the ensemble of those $(k_1,k_2,\dots,k_m,\epsilon_m)$ for which $k_1+k_2+\dots+k_m=k$ and $r_m(k_1,k_2,\dots,k_m,\epsilon_m)=r_m, \delta_m(k_1,\dots,k_m,\epsilon_m)=\delta_m$ are fixed.

During the work on $(3x+1)$–problem I had many discussions with A. Bufetov, N. Katz, A. Kontorovich, L. Koralov, J. Lagarias, and T. Suidan. It is my pleasure to thank all of them for very useful contacts. The criticism from J. Lagarias was especially important. He found a serious gap in the previous version of this paper. I am planning to discuss the related questions in a forthcoming publication. I thank also a referee for many useful remarks. The financial supports from NSF, grant DMS–0070698 and RFFI, grant N99–01–00314 are highly appreciated.

§2. The Structure Theorem

The formulation of the Structure Theorem was given in §1. Here we give its proof. First we consider the case $m=1$.

Assume that $\epsilon_m=+1$ and take $x=6p+1\in\sqcap^{+1}$. For given k_1 we should have

$$3x+1=18p+4=2^{k_1}(6t+\delta_1)$$

for some $\delta_1=\pm1$.

$a_1)$ $k_1=1$. Then

$$9p+2=6t+\delta_1.$$

This shows that p has to be odd, $p=2p_1+1$ and

$$6(3p_1+2)-1=6t+\delta_1.$$

Therefore in this case $\delta_1=-1, t=3p_1+2$, i.e. $q_1(1,+1)=1,\ r(1,+1)=2$.

$a_2)$ $k_1>1$. In this case

$$9p+2=2^{k_1-1}(6t+\delta_1).$$

This shows that p has to be even, $p = 2p_1$ and

$$6 \cdot 3p_1 = 6 \cdot 2^{k_1-1}t + \delta_1 2^{k_1-1} - 2\,. \tag{2.1}$$

The number $\delta_1 2^{k_1-1} - 2$ is even. The value δ_1 should be chosen so that $\delta_1 \cdot 2^{k_1-1} - 2 \equiv 0 \pmod 3$. If $k = 3, 5, 7, \ldots$ then δ_1 should be -1. If $k = 2, 4, 6, \ldots$ then $\delta_1 = 1$. In other words, $\delta_1 2^{k_1} \equiv 1 \pmod 3$. Since $\delta_1 = \delta_1^{-1}$ the last expression takes the form

$$\delta_1 \equiv 2^{k_1} \pmod 3\,, \tag{2.2}$$

i.e. δ_1 is uniquely determined by k_1.

Returning back to (2.1) we have

$$3p_1 - 2^{k_1-1}t = \frac{\delta_1 2^{k_1-2} - 1}{3}\,.$$

Let us write $p_1 = 2^{k_1-1}s + \bar{q}_1, t = 3s + \bar{r}_1$. For $\bar{q}_1, \bar{r}_1$ we have the equation

$$3\bar{q}_1 - 2^{k_1-1}\bar{r}_1 = \frac{\delta_1 2^{k_1-2} - 1}{3}\,. \tag{2.3}$$

If $\delta_1 = 1$ then rhs of (2.3) is non–negative and less than 2^{k_1-1}. If $k_1 = 2$ then it is zero and the solution to the equation (2.3) takes the form $\bar{q}_1 = 2, \bar{r}_1 = 3$, i.e. $q(2, +1) = 4$, $r_1(2, +1) = 3$.

If $k_1 > 2$ then rhs of (2.3) is positive. Consider the abelian group $Z_2^{k_1-1}$ of numbers mod 2^{k_1-1}. The multiplication by 3 is an automorphism of this group.

(2.3) is the equation for $\bar{q}_1$ in $Z_2^{k_1-1}$ and it has a unique solution. The value $\bar{r}_1, 1 \le \bar{r}_1 \le 3$ is also defined uniquely. Thus $q(k_1, +1) = 2\bar{q}_1$, $r(k_1, +1) = \bar{r}_1$.

If $\delta_1 = -1$ then rhs of (2.3) is negative. We rewrite (2.3) as follows:

$$3\bar{q}_1 - 2^{k_1-1}(\bar{r}_1 - 1) = 2^{k_1-1} + \frac{\delta_1 2^{k_1-2} - 1}{3}\,. \tag{2.4}$$

Now rhs is positive and we can use the same arguments as before to find $0 < \bar{q}_1 \le 2^{k_1-1}$, $0 \le \bar{r}_1 - 1 < 3$. Therefore $r(k_1, +1) = \bar{r}_1, 1 \le \bar{r}_1 \le 3$.

The case $x \in \sqcap^{-1}$ is considered in a similar way. We write $x = 6p - 1, p > 0$. For given k_1 we have the equation

$$3x + 1 = 18p - 2 = 18p' + 16 = 2^{k_1}(6t + \delta_1)$$

for $p' = p - 1 \geq 0$ and some $\delta_1 = \pm 1$.

a_1) $k_1 = 1$. Then

$$9p' + 8 = 6t + \delta_1 \,.$$

This shows that p' has to be odd, $p' = 2s + 1$ and

$$18s - 6t = -17 + \delta_1 \,.$$

Therefore $\delta_1 = -1$ and $3s + 3 = t$. From the last expression $p' = 2s + 1$, i.e. $q(2,-1) = 1$, $r(2,-1) = 3$ and $s = 0$.

a_2) $k_1 > 1$. Then

$$9p - 2 = 2^{k_1 - 1}(6t + \delta_1) \,.$$

p has to be even, $p = 2p_1, p_1 > 0$ and

$$9p_1 - 2^{k_1-2} \cdot 6t = 2^{k_1-1}\delta_1 - 8 \,.$$

The last expression shows that p_1 also has to be even, $p_1 = 2p_2, p_2 > 0$ and

$$9p_2 - 3 \cdot 2^{k_1-2}t = 2^{k_1-2}\delta_1 - 4 \,.$$

Rhs must be divisible by 3. This gives $2^{k_1-2}\delta_1 - 4 \equiv 0 \pmod 3$ or $2^{k_1} \equiv \delta_1 \pmod 3$. We get the equation

$$3p_2 - 2^{k_1-2}t = \frac{2^{k_1-2}\delta_1 - 4}{3} \tag{2.5}$$

If $k_1 = 4, \delta_1 = 1$, then *rhs* of (2.5) is zero and $p_2 = 2^{k_1-2}s, t = 3s, s > 0$. In other words, $p = 2^4 s, t = 3s, q_1(4,-1) = 0$, $r_1(4,-1) = 0$. In order to comply with the formulation of the theorem we change our choice to $q_1(4,-1) = 2^4$, $r_1(4,-1) = 3$ and $p \geq 0$.

If $k_1 = 2$ then $\delta_1 = 1$ and it is easy to check that $q(2,-1) = 3, r(2,-1) = 2$.

If $k_1 > 4, \delta_1 = 1$ we argue as before. *Rhs* of (2.5) is positive and less than 2^{k_1-2}. We put $p_2 = 2^{k_1-2}s + \bar{p}_2$, $t = 3s + \bar{t}$ and for $\bar{p}_2, \bar{t}$ we get the equation

$$3\bar{p}_2 - 2^{k_1-2}\bar{t} = \frac{2^{k_1-2}\delta_1 - 4}{3} \tag{2.6}$$

which has the unique solution $\bar{p}_2, 0 < \bar{p}_2 \leq 2^{k_1-2}$, and $\bar{t}$, $1 \leq \bar{t} \leq 3$. This gives $p = 2^{k_1 s} + 4\bar{p}_2$, $t = 3s + \bar{t}$, i.e. $q_1(k_1,-1) = 4\bar{p}_2$, $r_1(k_1,-1) = \bar{t}$.

If $\delta_1 = -1$ and rhs of (2.6) is negative we modify it as before

$$3\bar{p}_2 - 2^{k_1-2}(\bar{t}-1) = 2^{k_1-2} + \frac{2^{k_1-2}\delta_1 - 4}{3} \tag{2.6$'$}$$

Now rhs of (2.6′) is positive and we can find $1 \le \bar{p}_2 \le 2^{k_1-2}$, $0 \le \bar{t}-1 < 3$ satisfying (2.6′). Then $p = 2^{k_1}s + 4p_2$, $q(k_1,-1) = 4\bar{p}_2$ and $r_1(k_1,-1) = \bar{t}, 1 \le \bar{t} \le 3$.

The case $m > 1$ is considered by induction. Suppose that for some $m \ge 1$ the Structure Theorem is proven, i.e. $T^{(k_m)} \cdot T^{(k_{m-1})} \cdots T^{(k_1)}(6(2^{k_1+\cdots+k_m}s) + q_m(k_1,\ldots,k_m,\epsilon_m)) = 6(3^m s + r_m(k_1,\ldots,k_m,\epsilon_m)) + \delta_m(k_1,\ldots,k_m,\epsilon_m)$. Denote $x = 6(3^m s + r_m) + \delta_m$. Then $3x+1 = 2^{k_{m+1}}y$ where y is odd and

$$6(3^{m+1}s + 3r_m) + 3\delta_m + 1 = 2^{k_{m+1}}y$$

or

$$3 \cdot 3^{m+1}s + 9r_m + \frac{3\delta_m + 1}{2} = 2^{k_{m+1}-1}y\,. \tag{2.7}$$

If $k_{m+1} = 1, \delta_m = 1$ then (2.7) takes the form

$$3 \cdot 3^{m+1}s + 9r_m + 2 = y\,.$$

If r_m is even, $r_m = 2r_m^{(1)}$, then s must be odd, $s = 2s_1 + 1$,

$$3 \cdot 3^{m+1}(2s_1, +1) + 9 \cdot 2r_m^{(1)} + 2 = y\,,$$

or

$$6\left(3^{m+1}s_1 + 3r_m^{(1)} + \frac{3^{m+1}+1}{2}\right) - 1 = y\,.$$

This shows that $\delta_{m+1} = -1, r_{m+1}\ (k_1,\ldots,k_{m+1},\epsilon_{m+1}) = 3r_m^{(1)} + \frac{3^{m+1}+1}{2}$, $q_{m+1} = g_m + 2^{k_1+\cdots+k_m}$.

If r_m is odd, $r_m = 2r_m^{(1)} + 1$ then s has to be even, $s = 2s_1$ and

$$6(3^{m+1}s_1 + 3r_m^{(1)} + 2) - 1 = y\,.$$

We conclude that $\delta_{m+1} = -1$, $r_{m+1}(k_{1,\ldots,}k_{m+1},\epsilon_{m+1}) = 3r_m^{(1)} + 2, q_{m+1} = q_m$.

The case $k_{m+1} = 1, \delta_m = -1$ is considered in a similar way.

Now let $k_{m+1} > 1$. If r_m is even, $r_m = 2r_m^{(1)}$ and $\delta_m = 1$ then s has to be even, $s = 2s_1$ (see (2.7)) and from (2.7)

$$6 \cdot 3^{m+1} s_1 + 18 r_m^{(1)} + 2 = 2^{k_{m+1}-1} y\,. \tag{2.8}$$

If $y = 6t + \delta_{m+1}$ then $2^{k_{m+1}-1}\delta_{m+1} - 2$ must be divisible by 2. Therefore it has to be divisible by 6. Since it is always divisible by 2 it has to be also divisible by 3. As before, this shows that the value of k_{m+1} determines the value of δ_{m+1} for which this is true. The corresponding condition takes the form

$$2^{k_{m+1}} \equiv \delta_{m+1} \pmod 3\,. \tag{2.9}$$

From (2.8)

$$2^{k_{m+1}-1} t - 3^{m+1} s_1 = 3 r_m^{(1)} - \frac{2^{k_{m+1}-2}\delta_{m+1} - 1}{3}\,.$$

A general solution of the last equation is $t = 3^{m+1} s_2 + \bar{q}_{m+1}$, $s_1 = 2^{k_{m+1}-1} s_2 + \bar{r}_{m+1}$ or $s = 2s_1 = 2^{k_{m+1}} s_2 + 2\bar{r}_{m+1}$. This gives already one of the statements of the Structure Theorem. For $\bar{r}_{m+1}, \bar{q}_{m+1}$ we have the equation

$$2^{k_{m+1}-1} \bar{r}_{m+1} - 3^{m+1} \bar{q}_{m+1} = r_m^{(1)} + \frac{1 - \delta_{m+1} 2^{k_{m+1}-2}}{3} \tag{2.10}$$

Now we argue in the same way as in the case of $m = 1$. If rhs of (2.10) is non–negative, we can always find unique $\bar{r}_{m+1}, \bar{q}_{m+1}, 1 \leq \bar{r}_{m+1} \leq 3^{m+1}, 1 \leq \bar{q}_{m+1} \leq 2^{k_{m+1}-1}$, for which (2.10) is true.

If rhs of (2.10) is negative we modify it as follows

$$2^{k_{m+1}-1}(\bar{r}_{m+1} - 1) - 3^{m+1} \bar{q}_{m+1} = 2^{k_{m+1}-1} + r_m^{(1)} + \frac{1 - \delta_{m+1} 2^{k_{m+1}-2}}{3}\,.$$

Now the rhs is positive and we can find a solution for which $1 \leq \bar{q}_{m+1} \leq 2^{k_{m+1}-1}, 1 \leq \bar{r}_{m+1} \leq 3^{m+1}$. In all cases $q_m + 2\bar{r}_{m+1}, r_{m+1} = \bar{q}_{m+1}$.

If r_m is odd, $r_m = 2r_m^{(1)} + 1$ and $\delta_m = 1$ then

$$3 \cdot 3^{m+1} \cdot s + 18 r_m^{(1)} + 11 = 2^{k_{m+1}-1} y \tag{2.11}$$

and s has to be even, $s = 2s_1 + 1$. This yields

$$6 \cdot (3^{m+1} s_1 + 3 r_m^{(1)} + 2) + 3^{m+2} - 1 = 2^{k_{m+1}-1}(6t + \delta_{m+1})$$

and thus $3^{m+2} - 1 - 2^{k_{m+1}-1}\delta_{m+1}$ must be divisible by 6. Therefore δ_{m+1} should be such that $2^{k_{m+1}-1}\delta_{m+1} + 1$ is divisible by 3 which is equivalent to $2^{k_{m+1}} \equiv \delta_{m+1}$ (mod 3). It is clear that $3^{m+2} - 1 - 2^{k_{m+1}-1}\delta_{m+1}$ is even.

Now we write as before

$$t = 3^{m+1}s_2 + \bar{q}_{m+1}, s_1 = 2^{k_{m+1}-1}s_2 + \bar{r}_{m+1}\,,$$

and get for $\bar{q}_{m+1}, \bar{r}_{m+1}$ the equation

$$2^{k_{m+1}-1}\bar{q}_{m+1} - 3^{m+1}r_{m+1} = 3r_m^{(1)} + 2 + \frac{3^{m+2} - 1 - 1^{k_{m+1}-1}\delta_{m+1}}{6}\,.$$

This shows that $r_{m+1} = \bar{q}_{m+1}, q_{m+1} = 2\bar{r}_{m+1} + q_m$. The case $\delta_m = -1$ is considered in a similar way. The Structure Theorem is proven.

§3. A Corollary of the Structure Theorem

Take $x_0 \in \sqcap$ and put $x_m = T^m x_0$, $y_m = \ell n\, x_m, z_m = y_m - y_0, m \geq 1$.

Consider the probability

$$P_m(a,b) = P\left\{x_0 \middle| a \leq \frac{z_m + m(2\ell n\; 2 - \ell n\; 3)}{\sqrt{\sigma m}} \leq b\right\}\,.$$

Here a, b are fixed numbers, $\sigma > 0$ is a constant which will be described during the proof, the probability means the normalized wrt $\sqcap$ density, i.e. P_m is the relative (wrt $\sqcap$) density of $x_0 \in \sqcap$ satisfying the above inequalities.

Theorem 3.1.

$$\lim_{m\to\infty} P_m(a,b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-u^2/2}du\,.$$

Proof. . Consider any progression $\Sigma^{(k_1,\cdots,k_m,\epsilon_m)}$ (see the formulation of the Structure Theorem in §1). Then its probability in the sense mentioned above

$$P\{\Sigma^{(k_1,\cdots,k_m,\epsilon_m)}\} = 3\cdot\frac{1}{6\cdot 2^{k_1+\cdots+k_m}} = \frac{1}{2^{k_1+\cdots+k_m+1}}\,. \tag{3.1}$$

Actually the factor 3 is connected with the normalization density ($\sqcap$) = density ($\sqcap^{+1}$) + density ($\sqcap^{-1}$) $= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$ and additional 1 is connected with uniform distribution of $\epsilon_m = \pm 1$ independent on k_j.

Take large $x_0 \in \Sigma^{(k_1 \dots k_m, \epsilon_m)}$, $x_0 = (6(2^{k_1+\dots+k_m} p + q_m) + \epsilon_m)$. Then $x_0 = 6p \cdot 2^{k_1+\dots+k_m}(1+\circ(1))$, $x_m = 6 \cdot p \cdot 3^m(1+\circ(1))$ and $z_m = \ell n \, \frac{x_m}{x_0} = (k_1+\dots+k_m)\ell n \, 2 - m \, \ell n \, 3 + 0(1)$. Therefore $z_m + m(2\ell n \, 2 - \ell n \, 3) = (k_1+\dots+k_m - 2m) \, \ell n \, 2(1+\circ(1))$.

It follows from (3.1) that $k_1, \dots, k_m$ are independent random variables having geometrical distribution with parameter $\frac{1}{2}$ and $\frac{k_1+\dots+k_m-2m}{\sqrt{\sigma \, m}}$ for some $\sigma > 0$ has limiting Gaussian distribution. This implies the statement of the theorem. In an analogous way one can prove a limiting theorem for finite–dimensional distributions of z_m mentioned in §1.

Theorem 3.1 says that for very large x_0 typical $z_m = \ell n \frac{x_m}{x_0}$ decrease with the drift coefficient $-(2 \; \ell n \; 2 - \ell n \; 3)$. This means that x_m also decrease and this gives some reasons to expect that $(3x+1)$–problem is true. However, the main difficulty lies in the dynamics on the intermediate scales.

§4. $(\mathbf{r_m}, \delta_\mathbf{m})$ as a Random Walk

Take (r_m, δ_m) and a progression $\Sigma^{(k_1, \dots, k_m, \epsilon)}$ such that $T^m \Sigma^{(k_1, \dots, k_m, \epsilon)} = \wedge^{(r_m, \delta_m)}$ (see the Structure Theorem). In principle it can happen that for some (r_m, δ_m) there are no such $\Sigma^{(k_1, \dots, k_m, \epsilon)}$. But if they exist then $T^j \Sigma^{(k_1, \dots, k_j, \epsilon)} = \wedge^{(r_j, \delta_j)}$, $1 \le j \le m$, and the sequence (r_j, δ_j), $1 \le j \le m$ can be viewed as a trajectory of some random walk which ends at (r_m, δ_m). Different $\Sigma^{(k_1, \dots, k_m, \epsilon)}$ generate different trajectories.

We shall use the notation $\Phi_m(k_1, k_1, \dots, k_m, \epsilon) = (r_m(k_1, \dots, k_m, \epsilon), \delta_m(k_1, \dots, k_m, \epsilon))$. It is clear that $\Sigma^{(k_1, \dots, k_m, \epsilon)} \subset \Sigma^{(k_1, \dots, k_{m-1}, \epsilon)}$ and $\Phi_j(k_1, \dots, k_j, \epsilon) = (r_j, \delta_j)$ where $T^j \Sigma^{(k_1, \dots, k_j, \epsilon)} = \wedge^{(r_j, \delta_j)}$, $1 \le j \le m$. Sometimes we shall use also the equivalent writing $(r_m(k_1, \dots, k_m, \epsilon), \, \delta_m(k_1, \dots, k_m, \epsilon)) = \Phi_m(q_m(k_1, \dots, k_m, \epsilon), \epsilon)$. The value of δ_j can be found from (2.9):

$$2^{k_j} \equiv \delta_j \pmod 3 \tag{4.1}$$

which imposes some restrictions on possible values of k_j provided that δ_j are given. We shall show that there is another restriction of a similar type.

As in §2 we have the equation

$$3[6(3^{m-1}p + r_{m-1}) + \delta_{m-1}] + 1 = 2^{k_m} y \, ,$$

$p \ge 0$ and $6(3^{m-1}p + r_{m-1}) + \delta_{m-1} \in \wedge^{(r_{m-1}, \delta_{m-1})}$.

Since $y \in \sqcap$ we write $y = 6s + \delta_m$ where δ_m is found from (4.1) and

$$6[2^{k_m} s - 3^m p] + 2^{k_m}\delta_m - 3\delta_{m-1} - 1 = 18 r_{m-1}\,.$$

Define t by setting $s = 3^m t + r_m$, and then define t_m by $p = 2^{k_m} t + t_m$. Then

$$6[2^{k_m} r_m - 3^m t_m] = 18 r_{m-1} + 3\delta_{m-1} + 1 - 2^{k_m}\delta_m\,. \tag{4.2}$$

(4.2) shows that for given δ_{m-1} the value of k_m should be such that

$$3\delta_{m-1} + 1 \equiv 2^{k_m}\cdot\delta_m \qquad (\mathrm{mod}\ 6)$$

or $2^{k_m-1}\cdot\delta_m \equiv \frac{3\delta_{m-1}+1}{2}$ (mod 3). Using (4.1) we can write

$$2^{k_m} = \delta_m + 3a_m^{(1)}$$

for some odd $a_m^{(1)}$. Then

$$\frac{2^{k_m}\delta_m - 3\delta_{m-1} - 1}{6} = \frac{a_m^{(1)}\delta_m - \delta_{m-1}}{2}\,.$$

Returning back to (4.2) we get

$$2^{k_m} r_m - 3^m t_m = 3 r_{m-1} - \frac{a_m^{(1)}\delta_m - \delta_{m-1}}{2}\,. \tag{4.3}$$

This shows that for given r_m the value of k_m should be such that

$$2^{k_m} r_m + \frac{a_m^{(1)}\delta_m - \delta_{m-1}}{2} \equiv 0 \qquad (\mathrm{mod}\ 3)\,. \tag{4.4}$$

Since $a_m^{(1)}$ is odd, $a_m^{(1)} = 2a_m^{(2)} + 1$ and $a_m^{(2)} = g_m + 3a_m^{(3)}$. Remark that $a_m^{(1)}$, $a_m^{(2)}$, $a_m^{(3)}$, g_m are functions of k_m only. Actually

$$2^{k_m} = \delta_m + 3 + 6g_m + 18a_m^{(3)}\,. \tag{4.5}$$

For r_m we can write $r_m = h_m + 3r_m^{(1)}$ where h_m can take values $0, 1, 2$. The last expression can be considered as the definition of $h_m, r_m^{(1)}$ as functions of r_m. From (4.4)

$$h_m\delta_m + g_m\delta_m + \frac{\delta_m - \delta_{m-1}}{2} \equiv 0 \qquad (\mathrm{mod}\ 3) \tag{4.6'}$$

or

$$h_m + g_m + \frac{1-\delta_{m-1}\delta_m}{2} \equiv 0 \qquad (\text{mod } 3)\ . \qquad (4.6'')$$

The equations (4.6′), (4.6″) have an important interpretation. Suppose that we are given r_m, δ_m, δ_{m-1}. Then the value of δ_m determines the parity of k_m, the value of r_m gives the value of h_m and (4.6″) allows us to find the value of g_m.

Take again (4.3). It shows how to find r_{m-1} knowing r_m, k_m, δ_{m-1}. From (4.5), (4.6′), (4.6″)

$$2^{k_m} r_m^{(1)} + a_m^{(2)}\delta_m + \frac{\delta_m - \delta_{m-1}}{2} - 3^{m-1}t_m = r_{m-1}\ . \qquad (4.7)$$

Using the analogy with Markov processes we can call (4.7) the backward system of equations.

§5. The Ensemble $\Phi_{\mathbf{m}}^{-1}(\mathbf{r_m}, \delta_{\mathbf{m}})$.

As it follows from §3 it is natural to consider the probability distribution P for which $\epsilon = \pm 1$ with probabilities $\frac{1}{2}$ and $k_1, k_2, \dots, k_m, \dots$ is a sequence of independent, random variables, also independent on ϵ and having the geometric distribution with exponent $\frac{1}{2}$. All probabilities which we consider below are induced by this distribution. For example, with respect to this distribution the probability of any $q_m(k_1, \dots, k_m, \epsilon)$ equals $\frac{1}{2^{k_1+\cdots+k_m+1}}$ and the probability of a pair (r_m, δ_m) is the probability of all $(q_1, \dots, q_m, \epsilon)$ which give (r_m, δ_m) under the mapping Φ_m. The main purpose of this section is to study the probabilities of pairs (r_m, δ_m).

The pair (r_m, δ_m) can take 2.3^m values. On the other hand the number of typical $(k_1, \dots, k_m, \epsilon)$ grows (in a weak sense) as 2^{2m}. Therefore it is natural to expect that typically $\Phi_m^{-1}(r_m, \delta_m)$ contains $2^{2m}\cdot 3^{-m}$ elements.

Put $-c(k_m, \delta_m, \delta_{m-1}) = \frac{a_m^{(1)}\delta_m - \delta_{m-1}}{2}$, where (see above) $a_m^{(1)} = \frac{2^{k_m}-\delta_m}{3}$. Thus $-c(k_m, \delta_m, \delta_{m-1}) = \frac{2^{k_m}\delta_m - 1 - 3\delta_{m-1}}{5}$. In particular, $-c(1, -1, \delta_{m-1}) = -\frac{1+\delta_{m-1}}{2}$, $-c(2, 1, \delta_{m-1}) = \frac{1-\delta_{m-1}}{2}$, and so on. It is clear that $c(k_m, \delta_m, \delta_{m-1})$ can be positive or negative and for large k

$$-c(k_m, \delta_m, \delta_{m-1}) \sim \frac{2^{k_m}\delta_m}{6}\ .$$

Denote $\rho_m = \frac{r_m}{3^m}$. Then $0 \le \rho_m \le 1$ and possible values of ρ_m go with the step $\frac{1}{3^m}$. From (4.3)

$$\rho_m = \frac{t_m}{2^{k_m}} + \frac{1}{2^{k_m}}\cdot \rho_{m-1} + \frac{c(k_m, \delta_m, \delta_{m-1})}{2^{k_m}\cdot 3^m}\ . \qquad (5.1)$$

The iteration of the last equality yields

$$\rho_m = \frac{t_m}{2^{k_m}} + \frac{t_{m-1}}{2^{k_m+k_{m-1}}} + \cdots + \frac{t_1}{2^{k_m+k_{m-1}+\cdots+k_1}} + \sum_{s=1}^{m} \frac{c(k_s,\delta_s,\delta_{s-1})}{2^{k_m+\cdots+k_s}\cdot 3^s}\ . \tag{5.2}$$

It follows easily from (4.3) and from §2 that if $q_m = q_m(k_1,\dots,k_m,\epsilon) \in \Phi_m^{-1}(r_m,\delta_m)$ then

$$q_m = t_m \cdot 2^{k_{m-1}+\cdots+k_1} + t_{m-1}2^{k_{m-2}+\cdots+k_1} + \cdots + t_2\cdot 2^{k_1} + t_1\ . \tag{5.3}$$

Put $\kappa_m = q_m 2^{-(k_m+\cdots+k_1)}$. We have

$$\kappa_m = q_m 2^{-(k_m+\cdots+k_1)} = \frac{t_m}{2^{k_m}} + \frac{t_{m-1}}{2^{k_m+k_{m-1}}} + \cdots + \frac{t_1}{2^{k_m+k_{m-1}+\cdots+k_1}} \tag{5.4}$$

and from (5.2)

$$\rho_m = \kappa_m + \sum_{s=1}^{m}\frac{c(k_s,\delta_s,\delta_{s-1})}{2^{k_m+\cdots+k_s}\cdot 3^s} = \kappa_m + \frac{1}{3^m}\sum_{s=1}^{m}\frac{3^{m-s}c(k_s,\delta_s,\delta_{s-1})}{2^{k_m+\cdots+k_s}}\ . \tag{5.5}$$

Since k_s are independent random variables having geometric distribution with parameter $\frac{1}{2}$, $\delta_s = -1$ or $+1$ depending on the parity of k_s the sum $k_m+\cdots+k_s$ grows typically as $2(m-s)$. By this reason the last sum in (5.5) is converging, at least in probability, takes values $O(1)$ and has limiting distribution as $m\to\infty$. The formula (5.5) shows that for $(k_1,\dots,k_m,\epsilon)\in\Phi_m^{-1}(r_m,\delta_m)$ the difference $\rho_m-\kappa_m = O(\frac{1}{3^m})$. Write $k = k_1+\cdots+k_m$. It is a well–known combinatorial fact that the number $H_m(k)$ of solutions of the last equation with $k_i\ge 1$ equals to

$$H_m(k) = \binom{k-1}{m-1} = 2^{k-1}\cdot G_m(k-2m) \tag{5.6}$$

where $G_m(k-2m)$ have Gaussian asymptotics

$$G_m(k-2m) \sim \frac{1}{\sqrt{2\pi\sigma m}}\exp\left\{-\frac{(k-2m)^2}{2\sigma m}\right\}$$

for some constant $\sigma>0$ and not too large $|k-2m|$.

Put

$$\theta_m = \sum_{s=1}^{m}\frac{3^{m-s}c(k_s,\delta_s,\delta_{s-1})}{2^{k_m+\cdots+k_s}}$$

and

$$A_{m,i} = \left\{ ((k_1, \dots, k_m), \epsilon) \middle| \frac{i}{10} \le \theta_m < \frac{i+1}{10} \right\} .$$

Instead of 10 we could take any large enough integer. It is clear that the value of θ_m is basically determined by the last $k_m, k_{m-1}, \dots$. It follows from (5.5) that $((k_1, \dots, k_m), \epsilon) \in A_{m,i} \cap \Phi_m^{-1}((r_m, \delta_m))$ iff

$$\rho_m - \frac{(i+1)}{10 \cdot 3^m} < \kappa_m \le \rho_m - \frac{i}{10 \cdot 3^m} . \tag{5.7}$$

It is easy to show that one can find such constant $\gamma_\circ > 0$ that

$$P\{|\theta_m| > m^{\gamma_\circ}\} \le \frac{1}{m} .$$

We shall use the notation D'_m for the set of $(k_1, \dots k_m, \epsilon)$ for which $|\theta_m| \le m^{\gamma_\circ}$.

For any value of k the number of possible $(k_1, \dots, k_m, \epsilon) \in \Phi_m^{-1}(r_m, \delta_m) \cap A_{m,i}$ with the given k is at most $\frac{2^k}{10.3^m}$ because the interval (5.7) has width $1/10 \cdot 3^m$ and each κ_m is rational with denominator 2^k and all κ_m are distinct by (5.3).

Therefore the probability of this set is not greater than $\frac{1}{2 \cdot 10 \cdot 3^m} = \frac{1}{20 \cdot 3^m}$. As was mentioned above $P\{(r_m, \delta_m)\} = \sum_{\Phi_m(q_m, \epsilon) = (r_m, \delta_m)} P\{(q_m, \epsilon)\}$. Actually we can consider the partition ξ_m of the space Ω_m of pairs (κ_m, ϵ) onto pre–images $\Phi_m^{-1}((r_m, \delta_m))$. Denote by H_m the entropy of this partition, i.e. $H_m = -\sum P((r_m, \delta_m)) \ln P((r_m, \delta_m))$. Below the letter H is used for the entropy of a partition.

Theorem 5.1. $H_m \ge m \ln 3 - (2\gamma_\circ + 7) \ln m$

Proof. The proof is based upon the fact that if the entropy is small then there should be elements of partition having a big measure. This is impossible in our case. Let $B_k = \{(k_1, \dots, k_m, \epsilon) | k_1 + \dots + k_m = k\}$. It follows easily from the combinatorial formula above that we can find a constant γ_1 for which for all sufficiently large m

$$P\left\{ \bigcup_{|k-2m| \ge \gamma_1 \sqrt{m \ln m}} B_k \right\} \le \frac{1}{m^{\gamma_\circ + 2}} .$$

Introduce the partition α_m which has two elements

$$C'_m = \bigcup_{|k-2m| \le \gamma_1 \sqrt{m \ln m}} B_k , \qquad C''_m = \bigcup_{|k-2m| > \gamma_1 \sqrt{m \ln m}} B_k ,$$

and another partition β_m also onto two elements, D'_m,D''_m where D''_m is the complement of D'_m. Then

$$H_m \geq H(\xi_m|\alpha_m \vee \beta_m) = -\Big(\sum_{(r_m,\delta_m)} P((r_m,\delta_m)\big|C'_m \cap D'_m) \\ \ln P((r_m,\delta_m)|C'_m \cap D'_m) \cdot P(C'_m \cap D'_m) + \dots \Big) \tag{5.8}$$

where dots mean similar sums multiplied by small probabilities $P(C'_m \cap D''_m)$, $P(C''_m \cap D'_m)$, $P(C''_m \cap D''_m)$ respectively. All conditional entropies are less than $m \ln 3 + \ln 2$ because the partition has not more than 2.3^m elements. Therefore because of the estimates of the measures all these terms in (5.7) have absolute values less than a constant. Assume that the first sum is smaller than $m \ln 3 - (2\gamma_\circ + 6) \ln m$. By Chebyshev inequality

$$P\big\{ -\ln P(\Phi_m^{-1}(r_m,\delta_m)|C'_m \cap D'_m) \geq m \ln 3 - (2\gamma_\circ + 3) \ln m \big\} \\ \leq \frac{m \ln 3 - (2\gamma_\circ + 6) \ln m}{m \ln 3 - (2\gamma_\circ + 3) \ln m} = 1 - \frac{(2\gamma_\circ + 3) \ln m}{m \ln 3 - (2\gamma_\circ + 3) \ln m} \, .$$

Therefore

$$P\big\{ -\ln P\big(\Phi_m^{-1}(r_m,\delta_m)\big|C'_m \cap D'_m\big) < m \ln 3 - (2\gamma_\circ + 3) \ln m \big\} \\ \geq \frac{(2\gamma_\circ + 3) \ln m}{m \ln 3 - (2\gamma_\circ + 3) \ln m}$$

and by this reason the set of (r_m, δ_m) for which $-\ln P(\Phi_m^{-1}(r_m,\delta_m)|C'_m \cap D'_m) < m \ln 3 - (2\gamma_\circ + 3) \ln m$ or, equivalently, $P(\Phi_m^{-1}(r_m,\delta_m)|C'_m \cap D'_m) > \frac{m^{2\gamma_\circ + 3}}{3^m}$ is not empty. We shall show that this is impossible.

By definition

$$P\big(\Phi_m^{-1}(r_m,\delta_m)\big|C'_m \cap D'_m\big\} = \frac{P\big\{\Phi_m^{-1}(r_m,\delta_m) \cap C'_m \cap D'_m\big)}{P(C'_m \cap D'_m)} \\ = \frac{1}{P(C'_m \cap D'_m)} \sum_{\substack{|k-2m| \leq \gamma_1 \sqrt{m \ln m} \\ |i| \leq m^{\gamma_\circ}}} \\ P\big(\Phi_m^{-1}(r_m,\delta_m) \cap A_{m,i} \cap B_k\big) \, .$$

Therefore one can find $i_\circ$, $k_\circ$ such that

$$P\big(\Phi_m^{-1}(r_m,\delta_m)\cap A_{m,i_0}\cap B_{k_0}\big)\geq \frac{m^2}{3^m}$$

for all sufficiently large m. But it was already shown above that this probability cannot be greater than $\frac{1}{20.3^m}$. This contradiction proves the theorem.

Theorem 5.1 shows in what sense the distribution $\big\{P\big(\Phi_m^{-1}(r_m,\delta_m)\big)\big\}$ is close to the uniform. We believe that actually $H_m \geq m$ ln $3-$const.

References

[BP] K. A. Borovkov and D. Pfeifer, *Estimates for the Syracuse Problem via a probabilistic model*, Theory of Probability and its Applications **45, N2** (2000), 300–310.

[L] J. C. Lagarias, *The* $(3x+1)$*–problem and its generalizations*, American Mathematical Monthly **92** (1985), 3–23.

[W] G. J. Wirsching, *The Dynamical System generated by the* $(3x+1)$*–function*, Lecture Notes in Mathematics, N1681, Springer–Verlag, Berlin, 1998, 158p.