In this paper the problem of automaton transformation of a discrete random process into the discrete random distribution \( (1/n, 1/n, \ldots, 1/n) \) with any arbitrarily given accuracy is investigated. It is supposed that the properties of these input processes are not completely known, i.e., they are generated by an unstable random signal source. A class of "extremally" unstable random signal sources for which the generated random processes can still be transformed into the distribution \( (1/n, 1/n, \ldots, 1/n) \) by an adder modulo \( n \) is introduced. An estimate of the distance between the obtained and the discrete uniform random distributions is also found, dependent on the number of operating steps of the adder, in the case when the input process is described by a finite Markov chain.

1. INTRODUCTION

Davis (1961) initiated a series of articles in which the structural synthesis of stochastic automata was represented as a connection of a random input signal source with a finite deterministic automaton. To this series belong also the articles by Gelenbe (1970, 1971). Though in his last article Gelenbe uses two random signal sources (RSS), it does not change the essence of his method. However, the theoretical attractiveness and universality of the mentioned approach of stochastic automaton implementation depend on some ideal properties of RSS that are difficult to realize physically. Considering the relative stability (or instability) of functioning of the RSSs employed, Dvoretzky and Wolfowitz (1951) proposed a method of stabilizing the input process by an adder modulo \( n \). The stabilizing properties of adders were investigated also in the paper by Vorobyev (1954). In automata-theoretical terms this problem of transforming a random sequence
into the uniform distribution was restated by Gill (1962). Later this stabilization method was generalized by Lorenc (1976, 1978) and his disciples Lapiński and Métra (1973; Métra, 1975). They developed a technique of transformation more effective than adders or Gill automata and also applied new more precise estimation methods of the stabilization rate of the output process.

In this paper we consider the problem of stabilization of input process under very weak conditions concerning characteristics of the primary RSS and also give new estimates of stabilization rate for the adder modulo $n$.

Our statement of the problem has many common traits with the questions discussed in the paper by Dvoretzky and Wolfowitz (1951). We will use the terms already employed by Lorenc (1976, 1978).

2. Superunstable RSSs

Lorenc and Lapiński (1975) considered a set $B_{v,a,r}$ of primary sources of random signals $0, 1, \ldots, r$ and proposed a method for transforming the generated random processes into a discrete uniform random distribution $\pi_0 = (1/n, 1/n, \ldots, 1/n)$. In this section we will consider the set $B_{v,a,r}^*$ of primary RSSs which generate random processes satisfying some weaker conditions.

Let us say that the primary RSS $G$ belongs to the set $B_{v,a,r}^*$ if the output process $\{X_t\}_{t \geq 0}$ of $G$ satisfies the following conditions:

(A1) $\{X_t\}_{t \geq 0}$ is a Markov chain over the set $\{0, 1, \ldots, r\}$ ($r \geq 2$) whose order does not exceed $v$ ($v \geq 0$), and

(A2) all transition probabilities of this Markov chain are less than or equal to $1 - \alpha$ ($\alpha = $ positive real number).

Considering the source as an actual technical device, we will suppose that only the properties (A1) and (A2) of the generated process $\{X_t\}_{t \geq 0}$ are known. Those RSSs belonging to the $B_{v,a,r}^*$ will be called superunstable.

Let $\{X_t\}_{t \geq 0}$ be the output process of $G \in B_{v,a,r}^*$ and $\mathcal{A} = \{X, Y, Z, \Delta, \lambda\}$ be a finite deterministic automaton (FDA), for which $\{0, 1, \ldots, r\} \subset X$, $Y = Z = \{0, 1, \ldots, n - 1\}$ and $\forall z \in Z (\lambda(z) = z)$ ($X$, $Y$, and $Z$ are, respectively, the input and output alphabets and the set of internal states, $\Delta : X \times X \to Z$ is the transition function and $\lambda : Z \to Y$ is the output function of $\mathcal{A}$). For any $z_0 \in Z$ let us define inductively a sequence of random variables $\{Y_t\}_{t \geq 0}$ as follows:

1. $Y_0 = \Delta(z_0, X_0),$

2. if for some $t$ ($t \geq 0$) $Y_0, Y_1, \ldots, Y_t$ are defined then $Y_{t+1} = \Delta(Y_t, X_{t+1}).$
We will say that FDA $\mathcal{A}$ transforms the output process of the source $G \in B_{v, a, r}^*$ into the uniform random distribution $\pi_0 = (1/n, 1/n, ..., 1/n)$ if every sequence of random variables $\{X_t\}_{t \geq 0}$ satisfying the properties (A1) and (A2) also fulfills the condition

$$\forall z_0 \in Z \quad \forall t \geq 0$$

$$\forall y, y_0, ..., y_t \in Y \left( \lim_{N \to \infty} P\{Y_{t+N} = y | Y_t = y_t, ..., Y_0 = y_0\} = \frac{1}{n} \right). \quad (1)$$

In the above definition we demand that the equality (1) is satisfied for all possible changes of the process $\{X_t\}_{t \geq 0}$ restricted by conditions (A1) and (A2) only. In order to underline this fact we will sometimes say that FDA $\mathcal{A}$ transforms the output process of the source $G \in B_{v, a, r}^*$ into the stable uniform random distribution $\pi_0$. Sometimes we will also say simply that FDA $\mathcal{A}$ stabilizes the process generated by the source $G \in B_{v, a, r}^*$.

In this paper we confine ourselves to the case $v = 0$. Thus the behavior of any given RSS $G \in B_{v, a, r}^*$ can be described by a sequence of independent random variables $\{X_t\}_{t \geq 0}$ whose values are in $\{0, 1, ..., r\}$ and which satisfy the condition $P\{X_t = j\} \leq 1 - \alpha$, where $\alpha$ is a fixed positive real number. Thus we know about $X_t$ that the condition $P\{X_t = j\} \geq \alpha/r$ holds for at least two values of $j$. We do not exclude the case where these values of $j$ depend on time $t$. We will consider the following question: does there exist a finite deterministic Moore automaton $\mathcal{A} = \{X, Y, Z, A, \lambda\}$ which transforms a stochastic process $\{X_t\}_{t \geq 0}$ generated by the RSS $G \in B_{v, a, r}^*$ into the stable uniform random distribution $\pi_0$? Of course, we only consider such automata $\mathcal{A}$ for which $\|Z\| \geq r + 1$ and $\{0, 1, ..., r\} \subseteq X$.

The formulated problem cannot be solved by traditional methods because the approach of unifying two or more signals can lead to a situation in which the source $G$ emits a constant signal. The so called $g$-circulant transformers form an interesting class of stabilizing transformers.

Let $n$ and $g$ be natural numbers such that $(g, n) = 1$. Then FDA $\mathcal{A} = \{X, Y, Z, A, \lambda\}$ is called a $g$-circulant transformer if $X \subseteq Y = Z = \{0, 1, ..., n-1\}$, $|X| \geq 2$, $\forall z \in Z$ $(\lambda(z) = z)$, and $\forall x \in X, z \in Z$ $(A(z, x) \equiv z \cdot g + x \pmod{n})$. In the special case if $g = 1$ and $X = \{0, 1, ..., r\}$, the $g$-circulant transformer is called a Gill automaton.

Now we will prove

**Theorem 1.** The output process of the source $G \in B_{v, a, r}^*$ can be transformed into the uniform random distribution $\pi_0 = (1/n, 1/n, ..., 1/n)$ by a

---

1 We have here in mind only such values of $y_0, ..., y_t$ for which $P\{Y_t = y_t, ..., Y_0 = y_0\} > 0$.

2 $|Z|$ is the cardinality of the set $Z$.

3 $(g, n)$ stands for the greatest common divisor of the numbers $g$ and $n$. 
Gill automaton $\mathcal{A} = \{X, Y, Z, A, \lambda\}$ with $|Z| = n$ iff $r < p$, where $p$ is the minimal prime divisor of the number $n$.

Proof. Sufficiency. For describing the functioning of $\mathcal{A}$ we will use so called transition matrices, i.e., we will describe the functioning of $\mathcal{A}$ for the input letter $x \in X$ by the $n \times n$ matrix $M_x = (a_{ij}(x))$, where

$$a_{ij}(x) = \begin{cases} 1 & \text{if } A(i - 1, x) = j - 1 (i, j = 1, 2, \ldots, n), \\ 0 & \text{otherwise.} \end{cases}$$

If $A$ is a substitution matrix representing the cycle $(1, 2, \ldots, n - 1, 0)$ then we have $M_x = A^x$. We will show that in the case $r < p$ the automaton $\mathcal{A}$ stabilizes the output process of any source $G \in B^*_0, \ldots, r$. For this purpose it is sufficient to prove the following property: for every $t$ the matrix $M(t) = \prod_{s=1}^{n-1} (p_0(t+s)A^0 + p_1(t+s)A^1 + \cdots + p_r(t+s)A^r)$ is positive.\footnote{A matrix is positive if every entry is strictly greater than 0.}

We recall that since $G \in B^*_0, \ldots, r$ then at last two of the numbers $p_0(t+s), \ldots, p_r(t+s)$ are positive. On the other hand for all sequences of natural numbers $c_s, d_s$ satisfying the condition $0 \leq c_s < d_s < p$ there exists RSS $G \in B^*_0, \ldots, r$ such that the positive elements of the set $\{p_0(t+s), \ldots, p_r(t+s)\}$ are $p_{c_s}(t+s)$ and $p_{d_s}(t+s)$ only. Therefore matrix $M(t)$ is positive (for every RSS $G \in B^*_0, \ldots, r$) iff for all sequences of natural numbers $c_s, d_s (0 \leq c_s < d_s < p)$ the matrix $M = \sum_{s=1}^{n-1} (A^{c_s} + A^{d_s})$ is positive.

Let us denote by $f_s$ the difference $d_s - c_s$. Then for every $s$ the inequality $0 < f_s < p$ holds. As the matrices $A^{c_s} + A^{d_s}$ commute, we can write $M = A^{c_1} + c_2 + \cdots + c_{n-1} \prod_{s=1}^{n-1} (A^0 + A^{f_s})$. Therefore the matrix $M(t)$ is positive iff all matrices $Q = \prod_{t=1}^{n} (A^0 + A^{f_s})$ are positive.

We also denote by $V$ a subset of the set $U = \{1, 2, \ldots, n - 1\}$ and by $S(V)$—the value $\sum_{s \in V} f_s$. For the empty subset $\emptyset$ we define $S(\emptyset) = 0$. In this notation the matrix $Q$ can be defined in the form $Q = \sum_{V \subseteq U} A^{S(V)}$.

Now it is easy to see that $Q$ is positive iff for every value of $k$, $k = 0, 1, \ldots, n - 1$, we can find a subset $V$ such that

$$S(V) \equiv k \pmod{n}. \quad (2)$$

In order to prove that Eq. (2) is solvable for all $k$ we need the following lemma.

Lemma. If $p$ is the minimal prime divisor of the natural number $n$ and $f_s < p$ for any $s$, then the equation

$$\sum_{s=1}^{n-1} f_s y_s \equiv k \pmod{n}$$

has a Boolean solution for all $k$, $k = 0, 1, \ldots, n - 1$.\footnote{A matrix is positive if every entry is strictly greater than 0.}
Proof. We will show that for any integer \( r, 1 \leq r \leq n - 1 \), among the equations

\[
\sum_{s=1}^{r} f_s y_s \equiv k \pmod{n}, \quad k = 0, 1, \ldots, n - 1,
\]

at least \( r + 1 \) have a Boolean solution. Indeed, this is evident for \( r = 1 \), because \( y_1 = 0 \) (resp. \( y_1 = 1 \)) is a solution for \( f_1 y_1 \equiv 0 \pmod{n} \) (resp. \( f_1 y_1 \equiv f_1 \pmod{n} \)). Now we assume that this statement is true for some \( r, r < n - 1 \). We will show that it is also true for \( r + 1 \). There are only two alternatives:

1. Equation (3) is solvable at least for \( r + 2 \) different values of \( k \).
2. Equation (3) is solvable exactly for \( r + 1 \) different values of \( k \).

The first one gives at once that the equation

\[
\sum_{s=1}^{r+1} f_s y_s \equiv k \pmod{n}
\]

is solvable for at least \( r + 2 \) different values of \( k \). In the second case we find out whether the number\(^5\) \( k_{\mu} = \text{Res}(\mu f_{r+1}, n) \) belongs to those values of \( k \) for which Eq. (3) has a Boolean solution. If \( k_{\mu} \) is one of these values, then the equation

\[
\sum_{s=1}^{r+1} f_s y_s \equiv k_{\mu+1} \pmod{n}
\]

has a Boolean solution. As all the numbers \( k_0, k_1, \ldots, k_{r-1} \) are different, for some \( \mu (< r) \), Eq. (3) is solvable for \( k = k_{\mu} \) and not solvable for \( k = k_{\mu+1} \). From the solvability of Eq. (5) we conclude that Eq. (4) is solvable for \( r + 2 \) different values of \( k \). Thus we have proved the induction step for the second alternative, which completes the proof of the Lemma.

We conclude from the lemma that the matrix \( Q \) and hence every matrix \( M(t) \) is positive.

Necessity. We will show that the necessity holds in a more general case, i.e., if \( X = \{x_0, x_1, \ldots, x_r\} \subseteq \{0, 1, \ldots, n - 1\} \). Let \( r \geq p \). Then there exists a pair of indices \((i, j), 0 \leq i < j \leq r\), such that \( x_i \equiv x_j \pmod{p} \). Let us suppose that the input of a Gill automaton \( \mathcal{G} \) is fed by a source \( G \in B_{\mathcal{G}}^{*} \) which has the following property: the only output signals of \( G \) with positive probabilities are the signals \( x_i \) and \( x_j \). In that case the functioning of \( \mathcal{G} \) with the input source \( G \) is described by a Markov chain with the transition matrix

\[
P(X_t = x_i|M_{x_i}) + P(X_t = x_j|M_{x_j}) = P(X_t = x_i) A^{x_i} + P(X_t = x_j) A^{x_j}
\]

at the time

\(^5\) \text{Res}(a, b)—the remainder in dividing the number \( a \) by the number \( b \).
instant \( t \). Now we will show that no power of the matrix \( M = A^{x_i} + A^{x_j} \) is a positive matrix. From this it directly follows that condition (1) cannot be satisfied and therefore the Gill automaton \( \mathcal{O} \) does not transform the output process of the source \( G \) into the uniform random distribution \( \pi_0 \). In fact, if the matrix \( M^h \) is positive, then \( (A^0 + A^{kp})^h \) is also positive. However, by the binomial theorem

\[
(A^0 + A^{kp})^h = \sum_{s=0}^{h} C^s_h A^{ksp}.
\]  

(6)

Since \( p \) is a prime divisor of the number \( n \), then for every \( s, s \geq n/p \), we can find a value of \( r, r < n/p \), such that \( ksp \equiv krp \pmod{n} \). Then (see (6))

\[
(A^0 + A^{kp})^h = \sum_{s=0}^{n/p - 1} c_s A^{ksp},
\]

where \( c_s \) are positive integers. From this we conclude that the number of positive elements of the matrix \( (A^0 + A^{kp})^h \) does not exceed \( n^2/p \). This completes the necessity proof of Theorem 1.

**Theorem 2.** The process, generated by the source \( G \in B_0^* \), can be stabilized by means of a suitable \( g \)-circulant transformer \( \mathcal{O} = \{X, Y, Z, A, \lambda\} \) iff the least prime divisor \( p \) of the number \( n = |Z| \) is greater than \( r \).

The proof of this statement directly follows from Theorem 1 and the results of Lorenc (1978, Lemma 1, Theorem 1).

Lapiński and Lorenc (1983) showed that for every \( n > r \) there exists a finite deterministic Moore automaton \( \mathcal{O} = \{X, Y, Z, A, \lambda\} \) with \( n \) internal states which stabilizes the random process generated by the source \( G \in B_0^* \). However, it may be very difficult to find an effective description of the whole class of stabilizing transformers with given pairs of \( r \) and \( n \), \( r < n \). Also of some interest is the question about those pairs of \( r \) and \( n \), for which the stabilizing transformers have commuting transition matrices. Some optimization problems arising here are also of practical value. For example, we can try to find out which of the stabilizing transformers for \( G \in B_0^* \) has a larger stabilizing rate. As a rule, problems of this kind are characterized by a high degree of complexity even for small values of \( v \), e.g., \( v = 0 \) or \( v = 1 \). A striking illustration of this situation is given in the subsequent sections.

3. **The Stabilization Rate of Binary Markov Chains by Gill Automata**

In this section we will find the lower bound for the stabilization rate of the output sequence for a source \( G \in B_{1,a,1} \) into the distribution \( \pi_0 = (1/n, 1/n, ..., 1/n) \) by a Gill automaton \( \mathcal{O} = \{X, Y, Z, A, \lambda\} \), where \( |Z| = n \). We recall that a source belongs to the set \( B_{v,a,r} \) if
its functioning is described by a Markov chain over the set \( \{0, 1, \ldots, r\} \) (\( r \geq 1 \)) whose order does not exceed \( v \) (\( v \geq 0 \)), and

all transition probabilities of the Markov chain are greater than or equal to a positive real number \( a, a < 1/(r + 1) \).

Similarly, as in the previous section, where considering the source as an actual technical device, we will suppose that only the properties (A3) and (A4) of the generated process are known.

Let us say that the sequence of random variables \( \{Y_t\}_{t \geq 0} \) with values from the set \( B = \{\beta_1, \beta_2, \ldots, \beta_m\} \) realizes the random distribution \( z = (p_1, p_2, \ldots, p_m) \) with a dilatation \( N \) and accuracy \( \varepsilon \) (\( \varepsilon > 0 \)) if

\[
\forall t \left| P\{Y_{(t+1)N} = \beta_j | Y_{tN} = y_{tN}, \ldots, Y_0 = y_0\} - p_j \right| \leq \varepsilon, \quad j = 1, 2, \ldots, m, \text{ for all } y_0, y_1, \ldots, y_{tN} \in B \text{ such that } P\{Y_{tN} = y_{tN}, \ldots, Y_0 = y_0\} > 0.
\]

In accordance with this we will assume that an automaton \( \mathcal{A} \) with an input source \( G \in B_{v, \alpha, r} \) realizes the distribution \( \pi \) with a dilatation \( N \) and accuracy \( \varepsilon \) if the output process of automaton \( \mathcal{A} \) realizes the distribution \( \pi \) with a dilatation \( N \) and accuracy \( \varepsilon \) for every \( G \in B_{v, \alpha, r} \).

Lorenc (1976) showed that by means of a Gill automaton \( \mathcal{A} \) the output process of a source \( G \in B_{v, \alpha, 1} \) can be transformed into the distribution \( \pi_0 = (1/n, 1/n, \ldots, 1/n) \) with any arbitrarily given accuracy, if the value of the dilatation \( N \) is sufficiently large. For obtaining a lower bound for this transformation rate let us suppose that the functioning of \( G \) is described by a simple homogeneous Markov chain \( \{X_t\}_{t \geq 0} \) over the state set \( \{0, 1\} \) with a transition probability matrix \( A = (\begin{array}{cc} 1 - \alpha & \alpha \\ \alpha & 1 - \alpha \end{array}) \). We denote by \( \{Y_t\}_{t \geq 0} \) the corresponding output sequence of \( \mathcal{A} \). According to Lemma 1 of Lorenc (1976) the output process of the automaton \( \mathcal{A} \) is a Markov chain of order \( \mu + 1 \) if the input process is a Markov chain of order \( \mu \), i.e., in this case the process \( \{Y_t\}_{t \geq 0} \) is a Markov chain of order 2. Let us consider a new Markov chain, whose states are output sequences of the automaton \( \mathcal{A} \) of length 2; then the output process of the Gill automaton \( \mathcal{A} \) will be described by a simple homogeneous Markov chain \( \{Y^*_t\}_{t \geq 0} \) over a state set

\[
\{(0, 0), (0, 1), (1, 1), (1, 2), \ldots, (n - 1, n - 1), (n - 1, 0)\}
\]

with a transition probabilities matrix

\[
M = \begin{bmatrix}
\alpha & 1 - \alpha & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\
0 & 0 & 1 - \alpha & \alpha & 0 & \cdots & \cdots & \cdots & 0 & 0 \\
0 & 0 & \alpha & 1 - \alpha & 0 & \cdots & \cdots & \cdots & 0 & 0 \\
0 & 0 & 0 & 1 - \alpha & \alpha & 0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 - \alpha & \alpha \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \alpha & 1 - \alpha \\
1 - \alpha & \alpha & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0
\end{bmatrix}.
\]
The characteristic equation of the matrix \( M \) is
\[
(\lambda^2 - a\lambda)^n - (a\lambda - \alpha^2 + (1 - \alpha)^2)^n = 0.
\]

Hence the eigenvalues of \( M \) are the numbers
\[
\lambda_{2j, 2j+1} = \left( a \cos \frac{\pi j}{n} \pm \sqrt{1 - 2\alpha + \alpha^2 \cos^2(\pi j/n)} \right) e^{i\pi j/n}, \quad j = 0, 1, \ldots, n - 1.
\]

(7)

In this section we will use the following abbreviations:
\[
e^{i\pi j/n} = e_j \quad (j = \text{an integer}), \quad \sqrt{1 - 2\alpha + \alpha^2 \cos^2(\pi j/n)} = \beta_j, \quad s = \lfloor j/2 \rfloor.
\]

It is easy to check that the characteristic row vector and the characteristic column vector that correspond to the eigenvalue \( \lambda_j \) are
\[
U_j = \begin{bmatrix}
e_0 \\
\lambda_j - \alpha \\
1 - \alpha e^{-2s} \\
\vdots \\
\lambda_j - \alpha e^{-2s(n-1)} \\
1 - \alpha e^{-2s(n-1)} \\
e_0
\end{bmatrix}, \quad V_j = \begin{bmatrix}
e_0 \\
\lambda_j - \alpha \\
1 - \alpha e^{2s} \\
\vdots \\
\lambda_j - \alpha e^{2s(n-1)} \\
1 - \alpha e^{2s(n-1)} \\
e_0
\end{bmatrix}.
\]

As \( \alpha < 1/2 \) then \( \beta_j > 0 \), all eigenvalues (7) are different, and for \( j \neq r \) we have \( U_j V_r = 0 \). Besides
\[
U_j V_j = n + ne_{-2s} \left( \frac{\lambda_j - \alpha}{1 - \alpha} \right)^2 = n \left( 1 + \left( \frac{\lambda_j e_{-s} - \alpha e_{-s}}{1 - \alpha} \right)^2 \right)
\]
\[
= n \left( 1 + \left( \frac{(-1)^j \beta_s + i\alpha \sin(\pi s/n)}{1 - \alpha} \right)^2 \right)
\]
\[
= 2n\beta_s \frac{\beta_s + (-1)^j i\alpha \sin(\pi s/n)}{(1 - \alpha)^2} = \frac{2n\beta_s}{\beta_s - (-1)^j i\alpha \sin(\pi s/n)}.
\]

Therefore
\[
M = T A(1) T^{-1}, \quad (8)
\]

\(^6\) \([a]\)—the largest integer not exceeding the number \( a \).

\(^7\) \( A'\)—the transpose matrix of the matrix \( A \).
where $T = (V_0 V_1 \cdots V_{2n-1})$,

$$A(N) = \begin{bmatrix}
\lambda_0^N & 0 & 0 & \cdots & 0 \\
0 & \lambda_1^N & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & \lambda_{2n-1}^N
\end{bmatrix}, \quad T^{-1} = \begin{bmatrix}
\gamma_0 U_0 \\
\gamma_1 U_1 \\
\vdots \\
\gamma_{2n-1} U_{2n-1}
\end{bmatrix}.$$

$$\gamma_j = \frac{1}{U_j Y_j} = \frac{1}{2n} - (-1)^j \frac{ia \sin(\pi s/n)}{2n \beta_s}.$$  \hspace{1cm} (9)

It follows from (8) that

$$M^N = T A(N) T^{-1}. \hspace{1cm} (10)$$

Let us denote by $m_{jr}(N)$ the entry in the $(j+1)$th row and $(r+1)$th column of the matrix $M^N$, $j, r = 0, 1, \ldots, 2n - 1$. From (9) and (10) we obtain

$$m_{0,2r}(N) = \sum_{j=0}^{2n-1} \lambda_j^N \gamma_j (j+1) e^{-2sr}$$

$$= \frac{1}{2n} + \frac{1}{2n} (2a - 1)^N + \sum_{j=1}^{n-1} |\lambda_{2j}|^N (\gamma_{2j} e^{j(N-2r)})$$

$$+ \gamma_{2(n-j)+1} e^{-j(N-2r)}$$

$$= \frac{1}{2n} + \frac{1}{2n} (2a - 1)^N + \frac{1}{n} \sum_{j=1}^{n-1} |\lambda_{2j}|^N \left( \cos \frac{\pi j(N-2r)}{n} \right)$$

$$+ a \left( \sin \frac{\pi j}{n} \sin \frac{\pi j(N-2r)}{n} \right) \left( \beta_j \right)$$

and

$$m_{0,2r-1}(N) = \sum_{j=0}^{2n-1} \lambda_j^N \frac{\lambda_j - \alpha}{1 - \alpha} e^{-2sr} = \frac{1}{2n} - \frac{1}{2n} (2a - 1)^N$$

$$+ \sum_{j=1}^{n-1} |\lambda_{2j}|^N \left( \gamma_{2j} \frac{\lambda_{2j} - \alpha}{1 - \alpha} e^{j(N-2r)} \right)$$

$$+ \gamma_{2(n-j)+1} \frac{\lambda_{2(n-j)+1} - \alpha}{1 - \alpha} e^{-j(N-2r)}$$

$$= \frac{1}{2n} - \frac{1}{2n} (2a - 1)^N + \frac{1}{2n(1 - \alpha)} \sum_{j=1}^{n-1} \frac{1}{\beta_j} |\lambda_{2j}|^N$$

$$\times \left( \left( \beta_j - ia \sin \frac{\pi j}{n} \right) \left( \alpha \cos \frac{\pi j}{n} + \beta_j - \alpha e^{-j} \right) e^{j(N-2r+1/2)} \right)$$
\[ \left( \beta_j + i \alpha \sin \frac{\pi j}{n} \right) \left( \alpha \cos \frac{\pi j}{n} + \beta_j - \alpha \epsilon_j \right) e^{-j(N-2r+1)} \]

\[ = \frac{1}{2n} - \frac{1}{2n} (2\alpha - 1)^N \]

\[ + \frac{1}{n} \sum_{j=1}^{n-1} |\lambda_{2j}|^N \left( (1 - \alpha) \cos \frac{\pi j(N - 2r + 1)}{n} \right) \beta_j. \]

Let us assume that \( m_{0,-1}(N) = m_{0,2n-1}(N) \). Since \( \forall t > 0 \) \( P\{Y_{t+N} = r \mid Y_t = 0, Y_{t-1} = 0\} = m_{0,2r}(N) + m_{0,2r-1}(N), r = 0, 1, \ldots, n - 1 \), it follows that

\[ P\{Y_{t+N} = r \mid Y_t = 0, Y_{t-1} = 0\} = \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} |\lambda_{2j}|^N \left( \cos \frac{\pi j(N-2r)}{n} \right) \beta_j. \]

A simple analysis shows that

\[ 1 = \lambda_0 > |\lambda_2| = |\lambda_{2n-1}| > |\lambda_4| = |\lambda_{2n-3}| > \cdots > |\lambda_{2n-2}| = |\lambda_3| > |\lambda_1|. \]

Let \( N = 2nL + \eta \), where \( L \) is an integer and \( 0 \leq \eta < 2n \). For even \( N \) we obtain from (11)

\[ P \left\{ Y_{t+N} = \frac{\eta}{2} \mid Y_t = 0, Y_{t-1} = 0 \right\} = \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} |\lambda_{2j}|^N \left( 1 + \frac{(1 - \alpha) \cos(\pi j/n)}{\beta_j} \right); \]

for odd \( N \) there is

\[ P \left\{ Y_{t+N} = \frac{\eta + 1}{2} \mid Y_t = 0, Y_{t-1} = 0 \right\} = \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} |\lambda_{2j}|^N \left( \cos \frac{\pi j}{n} + \frac{1 - \alpha - \alpha \sin^2(\pi j/n)}{\beta_j} \right). \]

Since \( 1 > 1 - 2\alpha \), \( (1 - \alpha)/\beta_1 \geq 1 \), and \( (1 - \alpha - \alpha \sin^2(\pi/n))/\beta_1 \geq \beta_1 > \sqrt{1 - 2\alpha} \), we have

\[ \min \left( 1 + \frac{(1 - \alpha) \cos(\pi/n)}{\beta_1}, \cos \frac{\pi}{n} + \frac{1 - \alpha - \alpha \sin^2(\pi/n)}{\beta_1} \right) \geq \cos \frac{\pi}{n} + \sqrt{1 - 2\alpha}. \]
Therefore for all sufficiently large values of $N$ there exists (see (12) + (14)) $y_N \in \{0, 1, \ldots, n-1\}$ such that

$$
\forall t \left| P\{Y_{t+N} = y_N \mid Y_t = 0, Y_{t-1} = 0\} - \frac{1}{n} \right| \geq \frac{1}{n} \left( \cos \frac{\pi}{n} + \sqrt{1 - 2\alpha} \right) \left| \lambda_2 \right|^N. \quad (15)
$$

Let us write $f(N) \geq g(N)$ if $\exists N_0 \forall N > N_0 \quad f(N) \geq g(N)$. From (15) immediately follows the lower bound for the stabilization rate of the output process of the source $G \in B_{1,\alpha,1}$ by the Gill automaton $\mathcal{G}$.

**Theorem 3.** The accuracy $\varepsilon$ of the realization of the probability distribution $\pi_0 = (1/n, 1/n, \ldots, 1/n)$ by the Gill automaton $\mathcal{G}$ with an input source $G \in B_{1,\alpha,1}$ and a dilatation $N$ satisfies the condition

$$
\varepsilon \geq \frac{1}{n} \left( \cos \frac{\pi}{n} + \sqrt{1 - 2\alpha} \right) \left( \alpha \cos \frac{\pi}{n} + \sqrt{1 - 2\alpha + \alpha^2 \cos^2(\pi/n)} \right)^N. \quad (16)
$$

The particular Markov chain $\{X_t\}_{t \geq 0}$ with a bistochastic transition probability matrix $A$ is chosen for deriving the lower bound (16) in view of the fact that in this case the transformation process converges slowly. There is still no answer to the question: is it possible to choose among the Markov chains which describe the functioning of sources from the set $B_{1,\alpha,1}$, another chain with the transformation rate of the Gill automaton $\mathcal{G}$ into the distribution $\pi_0$ still lower than for the considered Markov chain $\{X_t\}_{t \geq 0}$? For the present, the best known upper bound of the accuracy of realization of the probability distribution $\pi_0$ by the Gill automaton $\mathcal{G}$ with an input source $G \in B_{\nu,\alpha,r}$ and a dilatation $N$ is given by the formula

$$
\varepsilon \leq \frac{n - 1}{n} - k(1 - (r + 1)^{\nu} n\alpha^{(r+\nu)}^{N/(\nu(r+\nu))}, \quad (17)
$$

where $\xi = -(n - 1)/r$, $(r + 1)^{\nu} < k < (r + 1)^{\nu} n$ (see Proposition 1 of Lorenc, 1976). For $\nu = r = 1$ this bound differs greatly from that of (16), though it can most likely be brought nearer to it.

A natural generalization of the problem discussed in this section is the problem of evaluating the transforming speed of Gill automata in the case $G \in B_{1,\alpha,r}, \ r > 1$. In our opinion, of considerable interest is the study of the stabilization rate of the output process for the sources $B_{\nu,\alpha,r}$ (in particular, $B_{1,\alpha,1}$) by $g$-circulant transformers.
4. The Stabilization Rate of Complex Markov Chains by Gill Automata

Let the input of a Gill automaton \( \mathcal{A} \) be fed by a source \( G \in B_{v,\alpha,r} \). As we have mentioned above, the output process of the automaton \( \mathcal{A} \) in this case is described by a Markov chain of an order not exceeding \( v + 1 \), which realizes the probability distribution \( \pi_0 = (1/n, 1/n, ..., 1/n) \) with a dilatation \( N \) and accuracy \( \varepsilon \), satisfying the inequality (17).

We will find a lower bound of the rate of this transformation for a very general case.

**Theorem 4.** Let \( (r + 1, n) = m > 1, \ r < n \). Then the accuracy \( \varepsilon \) of realization of the probability distribution \( \pi_0 = (1/n, 1/n, ..., 1/n) \) by the Gill automaton \( \mathcal{A} \) with an input source \( G \in B_{v,\alpha,r} \) and dilatation \( N \) satisfies the inequality

\[
\varepsilon \geq \frac{m-1}{n} \left(1-\frac{(r+1)}{n}\right)^{\left[(N+v)/(1+v)\right]}.
\]

**Proof.** Let us suppose that the functioning of the source \( G \in B_{v,\alpha,r} \) is described by a Markov chain \( \{X_t\}_{t \geq 0} \) of \( v \)th order \( (v \geq 0) \) over the state set \( H_n = \{0, 1, ..., n-1\} \) with the following conditional transition probabilities:

\[
P\{X_0 = 0\} = 1 - \alpha, \quad P\{X_0 = j\} = \alpha, \ j = 1, 2, ..., r, \quad \forall t < v, \forall x, x_0, ..., x_t \in H_{r+1}
\]

\[
P\{X_{t+1} = x \mid X_t = x, ..., X_0 = x_0\} =
\begin{cases} 
1 - \alpha & \text{if } x + x_t + \cdots + x_0 \equiv 0 \pmod{r+1}, \\
\alpha & \text{otherwise},
\end{cases}
\]

\[
P\{X_{t+1} = x \mid X_t = x, ..., X_0 = x_0\} =
\begin{cases} 
1 - \alpha & \text{if } x + x_t + \cdots + x_{t-v} \equiv 0 \pmod{r+1}, \\
\alpha & \text{otherwise},
\end{cases}
\]

Let \( Y_t = X_0 \oplus X_1 \oplus \cdots \oplus X_t, \ t = 0, 1, 2, ..., \) \( Z_t \) is a random variable with values from a set \( H_{np} \) \( (p = (r+1)/m) \) defined by a condition \( Z_t = X_0 + X_1 + \cdots + X_t, \ t = 0, 1, 2, ..., \)

\( \oplus \) and \( \ominus \), respectively, are addition and subtraction modulo \( n \); \( + \) and \( - \) denote addition and subtraction modulo \( np \).
\[ D_t = \{(y_0, y_1, ..., y_t) \mid y_0, y_1, ..., y_t \in H_r, 0 \leq y_0 \leq r, \\
0 \leq y_{u+1} \ominus y_u \leq r, u = 0, 1, ..., t-1\}, \]
\[ E_t = \{(z_0, z_1, ..., z_t) \mid z_0, z_1, ..., z_t \in H_n, 0 \leq z_0 \leq r, \\
0 \leq z_{u+1} - z_u \leq r, u = 0, 1, 2, ..., t-1\}, \]
\[ D = \bigcup_{t \geq 0} D_t, \quad E = \bigcup_{t \geq 0} E_t, \]

\(\psi\) is a mapping of the set \(E\) onto the set \(D\), defined by the condition
\[ \forall t \geq 0 \forall (z_0, z_1, ..., z_t) \in E_t 
(\psi(z_0, z_1, ..., z_t) = (\text{Res}(z_0, n), ..., \text{Res}(z_t, n))). \]

Then the sequences of the random variables \(\{Y_t\}_{t \geq 0}\) and \(\{Z_t\}_{t \geq 0}\) are Markov chains of order \(v + 1\). We will show by means of mathematical induction that
\[ \forall t, N \forall (z_0, z_1, ..., z_t) \in E_t 
\exists z^* \sum_{s=0}^{N-1} P\{Z_{t+N} = z^* + j + s(r+1) \mid Z_t = z_t, ..., Z_0 = z_0\} \]
\[ = \begin{cases} 
1 - (1 - (r+1)\alpha)^{(N+v)/(1+v)} & \text{if } j = 1, 2, ..., r, \quad (18) \\
1 + r(1 - (r+1)\alpha)^{(N+v)/(1+v)} & \text{if } j = 0,
\end{cases} \]

where \(\eta = n/m\). Let \(1 \leq N \leq v + 1\). We assume that \(\tau = \max(0, t + N - v)\), \(x_t = z_t - z_{t-1}, ..., x_0 = z_0 - z_{-1}, z_{-1} = 0\). Then we have\(^9\)
\[ \sum_{s=0}^{N-1} P\{Z_{t+N} = z + s(r+1) \mid Z_t = z_t, ..., Z_0 = z_0\} \]
\[ = \sum_{s=0}^{N-1} \sum_{u=0}^{\eta-1} \sum_{x_{t+N-1}, ..., x_0 = x_0} P\{X_{t+N} = z + s(r+1) - z_t - u \mid X_{t+N-1} \]
\[ = x_{t+N-1}, ..., X_0 = x_0\} \times P\{X_{t+N-1} = x_{t+N-1}, ..., X_{t+1} = x_{t+1} \mid X_t = x_t, ..., X_0 = x_0\}. \]

Let \(z^* = z_{t-1}\). Taking into account that \(P\{X_{t+N} = x \mid X_{t+N-1} = x_{t+N-1}, ..., X_0 = x_0\} = 0\) for \(x \notin H_{r+1}\), we obtain
\(^9\) The summation is taken over all ordered sets \((x_{t+1}, ..., x_{t+N-1})\) for which \(x_j \in H_{r+1}, j = t + 1, ..., t + N - 1\), and \(x_{t+N-1} + \cdots + x_{t+1} = u\).
\[
\sum_{s=0}^{n-1} P[Z_{t+N} = z^* + s(r+1) \mid Z_t = z_t, ..., Z_0 = z_0] \\
= (1 - ra) \sum_{u=0}^{\rho n-1} \sum_{x_{t+N-1} + \cdots + x_{t+1} = u} P[X_{t+N-1} = x_{t+N-1}, ..., X_{t+1} = x_{t+1}] \\
= x_{t+1} \mid X_t = x_t, ..., X_0 = x_0 \\
= 1 - ra = \frac{1 + r(1 - (r + 1)a)}{r + 1}.
\]

Similarly, for \( j \in \{1, 2, ..., r\} \) we have

\[
\sum_{s=0}^{n-1} P[Z_{t+N} = z^* + j + s(r+1) \mid Z_t = z_t, ..., Z_0 = z_0] \\
= a \sum_{u=0}^{\rho n-1} \sum_{x_{t+N-1} + \cdots + x_{t+1} = u} P[X_{t+N-1} = x_{t+N-1}, ..., X_{t+1} = x_{t+1}] \\
= x_{t+1} \mid X_t = x_t, ..., X_0 = x_0 \\
= a = \frac{1 - (1 - (r + 1)a)}{r + 1},
\]

i.e., the equality (18) is true for \( 1 \leq N \leq v + 1 \). Let us assume that it is true for all values of \( N \) such that \( N \leq (v + 1)R \) and prove it for \( (v + 1)R < N \leq (v + 1)(R + 1) \).

\[
\sum_{s=0}^{n-1} P[Z_{t+N} = z^* + s(r+1) \mid Z_t = z_t, ..., Z_0 = z_0] \\
= \sum_{s=0}^{n-1} \sum_{v=0}^{\rho n-1} \sum_{x_{t+N-1} + \cdots + x_{t+1} = v} P[X_{t+N-1} = x_{t+N-1}, ..., X_{t+1} = x_{t+1}] \\
= x_{t+1} \mid X_t = x_t, ..., X_0 = x_0 \\
= x_{t+1} \mid X_t = x_t, ..., X_0 = x_0 \\
= \sum_{s=0}^{n-1} (1 - ra) \sum_{x_{t+N-v-1} + \cdots + x_{t+1} = z^* + s(r+1)} P[X_{t+N-v-1} = x_{t+N-v-1}, ..., X_{t+1} = x_{t+1}] \\
= x_{t+1} \mid X_t = x_t, ..., X_0 = x_0 \\
= \frac{1 + r(1 - (r + 1)a)}{r + 1}
\]
\[= x_{t+N-v-1}, \ldots, X_{t+1} = x_{t+1} \mid X_t = x_t, \ldots, X_0 = x_0\]

\[= (1 - ra) \sum_{s=0}^{\eta-1} P\{Z_{t+N-v-1} = z^* + s(r + 1) \mid Z_t = z_t, \ldots, Z_0 = z_0\} + a \sum_{j=1}^{r} \sum_{s=0}^{\eta-1} P\{Z_{t+N-v-1} = z^* + j + s(r + 1) \mid Z_t = z_t, \ldots, Z_0 = z_0\}.\]

According to the induction assumption,

\[
\exists z^* \sum_{s=0}^{\eta-1} P\{Z_{t+N-v-1} = z^* + j + s(r + 1) \mid Z_t = z_t, \ldots, Z_0 = z_0\} = \frac{1 - (1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1} \quad \text{if } j = 1, 2, \ldots, r,
\]

\[
= \frac{1 + r(1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1} \quad \text{if } j = 0.
\]

Hence

\[
\sum_{s=0}^{\eta-1} P\{Z_{t+N} = z^* + s(r + 1) \mid Z_t = z_t, \ldots, Z_0 = z_0\} = (1 - ra) \frac{1 + r(1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1} + a \frac{1 - (1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1}
\]

and for \(j \in \{1, 2, \ldots, r\}\) we have

\[
\sum_{s=0}^{\eta-1} P\{Z_{t+N} = z^* + j + s(r + 1) \mid Z_t = z_t, \ldots, Z_0 = z_0\} = (1 - ra) \frac{1 - (1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1} + a \frac{1 + r(1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1} + (r - 1)a \frac{1 - (1 - (r + 1)a)^{(N-1)/(1+v)}}{r + 1}
\]

\[= \frac{1 - (1 - (r + 1)a)^{(N+v)/(1+v)}}{r + 1}.
\]
Therefore (18) holds for every $N$. From (18) we obtain that further $\forall t, N \forall (z_0, z_1, ..., z_t) \in E_t$

$$\exists z^* \in H_n \bigg\{ \sum_{u=0}^{n-1} P\{Z_{t+N} = z^* + j + um \mid Z_t = z_t, ..., Z_0 = z_0\}$$

$$= \frac{\rho - \rho(1 - (r + 1)\alpha)^{(N+v)/(1+v)}}{r+1} \quad \text{if } j = 1, 2, ..., m - 1,$$

$$= \frac{\rho + (r + 1 - \rho)(1 - (r + 1)\alpha)^{(N+v)/(1+v)}}{r+1} \quad \text{if } j = 0. \quad (19)$$

In view of (19) and the fact that $\psi$ is a one-to-one mapping of $E_t$ onto $D_t \forall t, N \forall (y_0, ..., y_t) \in D_t$

$$\exists z^* \in H_n \bigg\{ \sum_{u=0}^{n-1} P\{Z_{t+N} = z^* + j + um \mid (Z_0, ..., Z_t) = \psi^{-1}(y_0, ..., y_t)\}$$

$$= \frac{1 - (1 - (r + 1)\alpha)^{(N+v)/(1+v)}}{m} \quad \text{if } j = 1, 2, ..., m - 1,$$

$$= \frac{1 + (m - 1)(1 - (r + 1)\alpha)^{(N+v)/(1+v)}}{m} \quad \text{if } j = 0. \quad (20)$$

Since $\forall t, N \forall y \in H_n$

$$\forall (y_0, ..., y_t) \in D_t \bigg\{ \sum_{u=0}^{n-1} P\{Z_{t+N} = y + um \mid (Z_0, ..., Z_t) = \psi^{-1}(y_0, ..., y_t)\}$$

$$= \sum_{v=0}^{n-1} P\{Y_{t+N} = y \oplus vm \mid Y_t = y_t, ..., Y_0 = y_0\},$$

we finally obtain

$$\forall t, N, \forall (y_0, ..., y_t) \in D_t$$

$$\exists y^* \in H_n \bigg\{ \sum_{v=0}^{n-1} P\{Y_{t+N} = y^* \oplus j \oplus vm \mid Y_t = y_t, ..., Y_0 = y_0\}$$

$$= \frac{1 - (1 - (r + 1)\alpha)^{(N+v)/(1+v)}}{m} \quad \text{if } j = 1, 2, ..., m - 1,$$

$$= \frac{1 + (m - 1)(1 - (r + 1)\alpha)^{(N+v)/(1+v)}}{m} \quad \text{if } j = 0. \quad (20)$$
It follows from (20) that
\[
\max_{0 \leq v < n} P\{Y_{t+N} = y^* \oplus vm \mid Y_t = y_t, \ldots, Y_0 = y_0\} \\
\geq \frac{1 + (m-1)(1 - (r+1)a)^{(N+v)/(1+v)}}{n}
\]
and hence
\[
\forall t, N \forall (y_0, y_1, \ldots, y_t) \in D_t, \max_{y \in H_n} P\{Y_{t+N} = y \mid Y_t = y_t, \ldots, Y_0 = y_0\} \geq \frac{1}{n} \\
\geq \frac{m-1}{n} (1 - (r+1)a)^{(N+v)/(1+v)}.\]

Thus Theorem 4 has been completely proved.

5. CONCLUDING REMARKS

The obtained results are of theoretical rather than practical importance; they are interesting as a theoretical device for studying the question of exactness of the structural synthesis of stochastic automata. Theorems 1 and 2 show that a random process, generated by a superunstable source, can also be used for constructing a discrete uniform random distribution by logical means. As regards the stabilization rate, it is easy to analyze it by means of eigenvalues of circulant matrices. Theorem 1 generalizes Theorem 5 of Dvoretzky and Wolfowitz (1951). In addition, the method of our paper leads to a more accurate estimation of the stabilization rate of the process, in comparison with the bounds following from Theorem 6 of Dvoretzky and Wolfowitz (1951).

Theorems 3 and 4 allow us to find the time required for obtaining the discrete uniform distribution with a given accuracy, provided that the transformer is an adder modulo \(n\). The time dilatation sufficient for that can be determined from (17). Thus we obtain the interval \([N_0, N_1]\) containing the necessary and sufficient value of the dilatation. Besides, Theorem 4 shows that for fixed \(a, \epsilon, r, n\), the necessary number \(N_0\) of stabilization steps is growing at least linearly, depending on the parameter \(v\); more precisely, we have
\[
N_0 > v \left( 1 + \frac{\ln(\epsilon n/(m-1))}{\ln(1 - (r+1)a)} \right).
\]

We would like to prove Theorem 4 for the case \((n, r+1) = m = 1\), too.
However, for the time being the authors have failed to fill this gap. They also have a certain feeling of dissatisfaction concerning unexplored questions of this kind about g-circulant transformers. As our other investigations (which are not included in this paper and which are partially given in Lorenc (1978)) have shown, many of these transformers are remarkable because of the high stabilization rates of the input process.

ACKNOWLEDGMENT

The authors are indebted to Dr. I. Strazdis for his valuable help in preparing the manuscript.

RECEIVED April 29, 1982; ACCEPTED September 9, 1984

REFERENCES


