ACTA ARITHMETICA
XLVI (1986)

A Markov process underlying the generalized Syracuse algorithm

by

G. M. Leigh (Brisbane, Australia)

1. Introduction. Let $d \geqslant 2$ and $m_0, \ldots, m_{d-1}$ be non-zero integers. Let $r_0, \ldots, r_{d-1}$ be integers with $r_i \equiv im_i \pmod d$. T: $Z \to Z$ is defined by

$$(1) \qquad T(x) = \frac{m_i x - r_i}{d} \quad \text{if} \quad x \equiv i \pmod d \quad (0 \leqslant i < d).$$

Certain cases of this mapping have been studied by Matthews and Watts in [2] and [3]. Only the case $(m_i, d) = 1$ for all i was considered in [2]. In [3] this was extended to $(m_i, d^2) = (m_i, d)$. Also in [3], they used results on Markov matrices, without observing that there was an actual Markov chain involved.

In this paper no restriction will be placed on the m_i, except that they be non-zero. It will be shown in Section 4 how the results of Matthews and Watts are derived from the general theory.

Some examples are presented in Section 6. Example 3 illustrates the heuristic technique of Section 5 for reducing infinite Markov chains to finite ones. This serves to show that the infinite case is not just a curiosity, and useful results can be obtained.

Acknowledgement. I am grateful to Dr. K. R. Matthews for introducing me to the problem, and for suggesting Example 3, which, through attempts to explain it, led to the method outlined in Section 5.

The Aim. Given a modulus $m \geqslant 2$, the aim is to predict the limiting frequencies

$$\lim_{N \to \infty} f_{Nj},$$

where

$$(2) \qquad f_{Nj} = \frac{1}{N} \operatorname{card}\{n;\ 0 \leqslant n < N,\ T^n(x) \equiv j \pmod m\},$$

for trajectories $\{T^n(x)\}_{n \geqslant 0}$ which are not eventually periodic; these are called *divergent* trajectories.

Under certain fairly general conditions, values will be found for the related limits (see Theorem 4)

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} p_{nij}$$

with

$$p_{nij} = \frac{1}{d^n}\operatorname{card}_{md^n}\{x;\ x\equiv i\ (\operatorname{mod} m),\ T^n(x)\equiv j\,(\operatorname{mod} m)\} \tag{3}$$

where $\operatorname{card}_{md^n}(A)$ denotes the number of residue classes $\operatorname{mod} md^n$ contained in A. (Compare equation (1.4) in [3].)

To permit the application of the theory of probability, we put $l = [m, d]$ and extend T to a mapping from the set of l-adic integers, Z_l, into itself. The extension is easy, being defined expressly by (1).

Markov chains $\{X_n\}$ and $\{Y_n\}$ are defined over the probability space (Z_l, P) where P is the Haar measure on Z_l. $\{X_n\}$ and $\{Y_n\}$ are related by Theorem 3 which implies that it does not matter which of them is used. $\{X_n\}$ is to be preferred in practice so as to reduce the number of states to be considered, while $\{Y_n\}$ is more useful for most theoretical purposes.

Results on l-adic integers are transferred to rational integers by means of

CONJECTURE 1. *Any condition on the distribution of iterates which has probability zero, does not occur in divergent trajectories.*

Corollaries of this conjecture include the following:

(i) If a divergent trajectory enters a positive class then

$$\lim_{N\to\infty} f_{Nj} = p_j$$

where p_j are given by (29) (see Theorem 5).

(ii) If the Markov chains are finite, every divergent trajectory enters a positive class. (See Theorem 6.)

(iii) Following on from (i), take $m = d$. Suppose a trajectory enters a positive class and

$$\prod_{j=0}^{d-1}\left|\frac{m_j}{d}\right|^{p_j} < 1. \tag{4}$$

Then the trajectory is eventually periodic. This follows from (61) of Lemma 3.

2. Notation. (a, b) and $[a, b]$ denote respectively the greatest common divisor and the least common multiple of the integers a, b.

$B(j, M)$ is the residue class $\{x\in Z_l;\ x\equiv j\,(\operatorname{mod} M)\}$ $(M\in N)$; $\mathscr{B}$ is the set of all such residue classes.

For non-zero integers x, the functions $C(x)$ and $D(x)$ are defined by:

$$C(x)\in Z,\quad D(x)\in N,\quad C(x)D(x) = x,$$

with $(C(x), d) = 1$ and $D(x)$ a divisor of some power of d. For example, if $d = 10$ then $C(-360) = -9$ and $D(-360) = 40$.

Let $c = C(m)$ and $d_i = D(m_i)$ $(0\leqslant i < d)$.

The term *trajectory* will be used interchangeably to mean a sequence $\{T^n(x)\}_{n\geqslant 0}$ for an *integer* starting point x, or the corresponding sequence of sets $\{X_n\}_{n\geqslant 0}$ or $\{Y_n\}_{n\geqslant 0}$.

Notation for Markov chains. The usual notation for conditional probabilities will be used, *viz.*

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(A\cap B)}{\mathrm{P}(A)}.$$

If a Markov chain $\{X_n\}$ has stationary transition probabilities

$$p_{AB} = \mathrm{P}(X_{n+1} = B|X_n = A),$$

the corresponding symbol $p_{AB}^{(n)}$ will represent the n-step transition probability

$$p_{AB}^{(n)} = \mathrm{P}(X_n = B|X_0 = A) = \sum_{C_1,\ldots,C_{n-1}} p_{AC_1}p_{C_1C_2}\cdots p_{C_{n-1}B}.$$

A state will be called *positive* if it is recurrent (persistent) and has a finite mean recurrence time. A *class* is positive if every state in it is positive. A positive class will commonly be denoted by $\mathscr{C}$. The *limiting probabilities* corresponding to a positive class $\mathscr{C}$ are the limits

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} p_{BB}^{(n)}\quad (B\in\mathscr{C})$$

(see Chung [1], p. 32).

A Markov chain is *ergodic* if its entire state space consists of one positive class.

A state that is recurrent with an infinite mean recurrence time will be called *null*; a state that is not recurrent will be called *transient*.

A Markov chain is *finite* if it can take on only a finite number of states, i.e. if its state space is finite.

3. Construction and application of the Markov chains. If $B = B(x, M)\in\mathscr{B}$ let

$$G(B, y) = B(y, [M, l]), \tag{5}$$

and if $d|M$,

$$H(B) = B\left(T(x),\ M\frac{d_i}{d}\right) \tag{6}$$

where $x\equiv i\,(\operatorname{mod} d)$ $(0\leqslant i < d)$.

Define the sequences of random sets $\{X_n\}_{n\geq 0}$ and $\{Y_n\}_{n\geq 0}$ as follows:

(7) $$X_0 = Z_l,$$

(8) $$Y_n = G(X_n, T^n(x)),$$

and

(9) $$X_{n+1} = H(Y_n).$$

(7) is an optional condition; it serves merely to define the sequences uniquely, and will not always be assumed.

For $A, B \in \mathscr{B}$ let

(10) $$z_{AB} = \begin{cases} \mathrm{P}(B|A) & \text{if } B = G(A, x) \text{ for some } x, \\ 0 & \text{otherwise.} \end{cases}$$

As an example take

(11) $$T(x) = \begin{cases} x/2 & \text{if } x \text{ is even,} \\ 5x+1 & \text{if } x \text{ is odd,} \end{cases}$$

$m = 3$, and $x = 7$. Then $l = 6$ and the trajectory is $\{7, 36, 18, 9, \ldots\}$. $X_0 = Z_l$, $Y_0 = B(1, 6)$, $X_1 = B(0, 6) = Y_1$, $X_2 = B(0, 3)$, $Y_2 = B(0, 6)$, $X_3 = B(0, 3)$, $Y_3 = B(3, 6)$, etc. $z_{X_0 Y_0} = 1/6$, $z_{X_1 Y_1} = 1$, $z_{X_2 Y_2} = 1/2 = z_{X_3 Y_3}$.

We now list the main results of this section, before setting down their proofs together with some lemmas.

THEOREM 1. *$\{X_n\}$ is an infinite Markov chain with stationary transition probabilities*

(12) $$p_{AB} = \sum_{H(C)=B} z_{AC}.$$

The p_{nij} of (3) *are given by*

(13) $$p_{nij} = \sum_{B\in\mathscr{B}} p^{(n)}_{B(i,m)B} \mathrm{P}(B(j, m)|B).$$

THEOREM 2. *$\{Y_n\}$ is a Markov chain with transition probabilities*

(14) $$q_{AB} = z_{H(A)B},$$

and

(15) $$p_{nij} = \frac{m}{l} \sum_{k=0}^{l/m-1} \sum_{B\subset B(j,m)} q^{(n)}_{B(i+km,l)B}.$$

THEOREM 3. *For $n \geq 1$,*

(16) $$q^{(n)}_{AB} = \sum_{C} p^{(n-1)}_{H(A)C} z_{CB}$$

and

(17) $$p^{(n)}_{AB} = \sum_{C} \sum_{H(D)=B} z_{AC} q^{(n-1)}_{CD}.$$

If $\mathscr{C}$ is a positive class of $\{X_n\}$ with limiting probabilities π_B then

(18) $$\mathscr{C}' = \{B\in\mathscr{B};\ z_{AB}\neq 0 \text{ for some } A\in\mathscr{C}\}$$

is a positive class of $\{Y_n\}$ with limiting probabilities

(19) $$\varrho_B = \sum_{C\in\mathscr{C}} \pi_C z_{CB},$$

and $H^{-1}(\mathscr{C}) - \mathscr{C}'$ contains only transient states of $\{Y_n\}$, all of which lead into $\mathscr{C}'$. If $\mathscr{C}'$ is a positive class of $\{Y_n\}$ with limiting probabilities ϱ_B then

(20) $$\mathscr{C} = H(\mathscr{C}')$$

is a positive class of $\{X_n\}$ with limiting probabilities

(21) $$\pi_B = \sum_{\substack{C\in\mathscr{C}' \\ H(C)=B}} \varrho_C.$$

THEOREM 4. *Suppose $\{X_n\}$ and $\{Y_n\}$ enter positive classes with probability one. Then*

(22) $$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{nij} = \sum_{\mathscr{C}} f_{B(i,m)\mathscr{C}} \sum_{B\in\mathscr{C}} \pi_B \mathrm{P}(B(j, m)|B)$$

(23) $$= \frac{m}{l} \sum_{k=0}^{l/m-1} \sum_{\mathscr{C}'} f'_{B(i+km,l)\mathscr{C}'} \sum_{\substack{B\in\mathscr{C}' \\ B\subset B(j,m)}} \varrho_B,$$

where p_{nij} are as in (3), *the sums are over all positive classes $\mathscr{C}$ and $\mathscr{C}'$ respectively, and*

(24) $$f_{B\mathscr{C}} = \mathrm{P}(\{X_n\} \text{ enters } \mathscr{C}|X_0 = B),$$

(25) $$f'_{B\mathscr{C}'} = \mathrm{P}(\{Y_n\} \text{ enters } \mathscr{C}'|Y_0 = B).$$

If $\{X_n\}$ and $\{Y_n\}$ are ergodic this reduces to

(26) $$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{nij} = \sum_{B} \pi_B \mathrm{P}(B(j, m)|B)$$

(27) $$= \sum_{B\subset B(j,m)} \varrho_B.$$

This result corresponds to equation (1.4) of [3].

Remarks. If $\{X_n\}$ and $\{Y_n\}$ are finite Markov chains, the condition for Theorem 4 is automatically satisfied (see Theorem 6).

If all the positive classes are aperiodic, the Cesàro limits in (22) and (26) may be replaced by ordinary limits. This was observed by Matthews and Watts but they were unable to prove it.

THEOREM 5. *If $\mathscr{C}$ and $\mathscr{C}'$ are corresponding positive classes, as in Theorem 3, then*

(28) $$P(\lim_{N\to\infty} f_{Nj} = p_j \mid \{X_n\} \text{ enters } \mathscr{C} \text{ or } \{Y_n\} \text{ enters } \mathscr{C}') = 1$$

where f_{Nj} are given by (2), *and*

(29) $$p_j = \sum_{B\in\mathscr{C}} \pi_B P(B(j, m)|B) = \sum_{\substack{B\in\mathscr{C}'\\ B\subset B(j,m)}} \varrho_B.$$

This proves Corollary (i) of Conjecture 1.

THEOREM 6. *If $\{X_n\}$ and $\{Y_n\}$ are finite Markov chains, each enters a positive class with probability one.*

This proves Corollary (ii) to Conjecture 1.

Remark. The case where $\{X_n\}$ and $\{Y_n\}$ are finite is very important. Indeed, it is unknown if there is any non-finite example in which $\{X_n\}$ and $\{Y_n\}$ enter positive classes with probability 1.

All cases considered by Matthews and Watts correspond to finite Markov chains (see Section 4).

The proofs.

LEMMA 1. *For $n \geqslant 0$,*

(30) $$P(X_0 = A_0, Y_0 = B_0, \ldots, X_n = A_n, Y_n = B_n) = z_{A_nB_n} P(X_0 = A_0, Y_0 = B_0, \ldots, X_n = A_n).$$

Proof. (30) certainly holds if the event

(31) $$X_0 = A_0, \quad Y_0 = B_0, \quad \ldots, \quad X_n = A_n$$

has probability zero. Suppose then that (31) has a positive probability. Let $x = x_0$ satisfy (31). Let

(32) $$B_i = B(T^i(x_0), M_i),$$

(33) $$K_i = [M_i, l]$$

and

(34) $$T^i(x_0) \equiv j_i (\mathrm{mod}\, d) \quad (0 \leqslant j_i < d),$$

for $0 \leqslant i \leqslant n$.

Choose $u \in N$ such that $C(u) = c$ and $x = x_0 + u$ satisfies

(35) $$T^i(x) \equiv T^i(x_0) (\mathrm{mod}\, K_i) \quad (0 \leqslant i \leqslant n-1);$$

for example $u = d^n[K_0, \ldots, K_n]$. Let

(36) $$T^i(x_0+u) - T^i(x_0) = v_i K_i \quad (0 \leqslant i \leqslant n-1)$$

for integers $v_0, \ldots, v_{n-1}$.

When $i = 0$ (36) is

$$u = v_0 K_0,$$

so

(37) $$v_0 | u.$$

If $1 \leqslant i \leqslant n-1$, (36) implies

$$\begin{aligned} v_i K_i &= T^i(x_0+u) - T^i(x_0) \\ &= \frac{m_{j_{i-1}}}{d}[T^{i-1}(x_0+u) - T^{i-1}(x_0)] \quad \text{by (1), (34) and (35)} \\ &= \frac{m_{j_{i-1}}}{d} v_{i-1} K_{i-1}. \end{aligned}$$

From (6) and (33), $K_i = \left[K_{i-1}\frac{d_{j_{i-1}}}{d}, l\right]$. Therefore

$$v_i = \frac{K_{i-1}\frac{m_{j_{i-1}}}{d}}{\left[K_{i-1}\frac{d_{j_{i-1}}}{d}, l\right]} v_{i-1},$$

and

$$D(v_i) = \frac{D(K_{i-1})\frac{d_{j_{i-1}}}{d}}{\left[D(K_{i-1})\frac{d_{j_{i-1}}}{d}, D(l)\right]} D(v_{i-1}).$$

Hence

(38) $$D(v_i) | D(v_{i-1}).$$

Combining (37) and (38), we have

(39) $$D(v_{n-1}) | D(v_{n-2}) | \ldots | D(v_0) | u.$$

Put

(40) $$U = \frac{u}{D(v_{n-1})}.$$

Since by (33) $d|K_i$, (36) and (39) imply that $x = x_0+U$ satisfies (35). By (36) and (40),

$$T^{n-1}(x_0+U)-T^{n-1}(x_0) = C(v_{n-1})K_{n-1},$$

so

$$T^n(x_0+U)-T^n(x_0) = C(v_{n-1})\frac{m_{j_{n-1}}}{d}K_{n-1} = wM_n \quad \text{by (6)}$$

where $w = C(v_{n-1})C(m_{j_{n-1}})$ satisfies

$$(w,\, d) = 1. \tag{41}$$

Furthermore,

$$T^n(x_0+aU)-T^n(x_0) = awM_n \tag{42}$$

for all $a \in Z_l$.

For Y_n equal to B_n, B_n must be a residue class mod K_n which is contained in A_n. Then

$$\begin{aligned} \mathrm{P}\{Y_n = B_n | x \equiv x_0 (\mathrm{mod}\, U)\} &= \mathrm{P}\{T^n(x)\in B_n | x \equiv x_0 (\mathrm{mod}\, U)\} \\ &= M_n/K_n = z_{A_nB_n} \quad \text{by (42) and (41)}. \end{aligned}$$

That is,

$$\mathrm{P}\{Y_n = B_n,\, x \equiv x_0 (\mathrm{mod}\, U)\} = z_{A_nB_n}\mathrm{P}\{x \equiv x_0 (\mathrm{mod}\, U)\}. \tag{43}$$

Now write $U = U(x_0)$ and define the relation $\sim$ on the x which satisfy (31) by

$$x_1 \sim x_2 \Leftrightarrow x_1 \equiv x_2 \big(\mathrm{mod}\, U(x_1)\big). \tag{44}$$

$\sim$ is easily shown to be an equivalence relation, since $U(x_1) = U(x_2)$ if $x_1 \sim x_2$. (30) then results by summing both sides of (43) over x_0, taking one value from each equivalence class.

Proof of Theorems 1 and 2. For $n \geqslant 0$,

$$\begin{aligned} \mathrm{P}(X_0 = A_0, \ldots, X_{n+1} = A_{n+1}) &= \sum_{H(B)=A_{n+1}} \mathrm{P}(X_0 = A_0, \ldots, X_n = A_n,\, Y_n = B) \\ &\qquad \text{using (9)} \\ &= \sum_{H(B)=A_{n+1}} z_{A_nB}\mathrm{P}(X_0 = A_0, \ldots, X_n = A_n) \\ &\qquad \text{by Lemma 1} \\ &= p_{A_nA_{n+1}}\mathrm{P}(X_0 = A_0, \ldots, X_n = A_n). \end{aligned}$$

Thus $\{X_n\}$ is a Markov chain as stated.

If $n \geqslant 1$,

$$\begin{aligned} \mathrm{P}(Y_0 = B_0, \ldots, Y_n = B_n) &= \mathrm{P}\{Y_0 = B_0, \ldots, Y_{n-1} = B_{n-1},\, X_n = H(B_{n-1}),\, Y_n = B_n\} \\ &= z_{H(B_{n-1})B_n}\mathrm{P}\{Y_0 = B_0, \ldots, Y_{n-1} = B_{n-1},\, X_n = H(B_{n-1})\} \\ &\qquad \text{by Lemma 1} \\ &= q_{B_{n-1}B_n}\mathrm{P}(Y_0 = B_0, \ldots, Y_{n-1} = B_{n-1}). \end{aligned}$$

Therefore $\{Y_n\}$ is a Markov chain with transition probabilities (14).

It follows from (3) that

$$\begin{aligned} p_{nij} &= \mathrm{P}\{T^n(x) \equiv j (\mathrm{mod}\, m) |\, x \equiv i (\mathrm{mod}\, m)\} \\ &= \mathrm{P}\{Y_n \subset B(j,\, m) |\, X_0 = B(i,\, m)\} \\ &= \sum_{B\subset B(j,m)} \mathrm{P}\{Y_n = B | X_0 = B(i,\, m)\} \\ &= \sum_{B\subset B(j,m)} \sum_{A\in\mathscr{B}} \mathrm{P}\{X_n = A,\, Y_n = B | X_0 = B(i,\, m)\} \\ &= \sum_{B\subset B(j,m)} \sum_{A} z_{AB}\, \mathrm{P}\{X_n = A | X_0 = B(i,\, m)\} \quad \text{by (30)} \\ &= \sum_{A} \mathrm{P}\big(B(j,\, m) |\, A\big) p^{(n)}_{B(i,m)A} \quad \text{by (10)}. \end{aligned} \tag{45}$$

This is (13).

From (45),

$$\begin{aligned} p_{nij} &= \sum_{B\subset B(j,m)} \sum_{k=0}^{l/m-1} \mathrm{P}\{Y_0 = B(i+km,\, l),\; Y_n = B | X_0 = B(i,\, m)\} \\ &= \sum_{B\subset B(j,m)} \sum_{k=0}^{l/m-1} \frac{m}{l}\mathrm{P}\{Y_n = B |\, Y_0 = B(i+km,\, l)\} \\ &= \frac{m}{l}\sum_{k=0}^{l/m-1} \sum_{B\subset B(j,m)} q^{(n)}_{B(i+km,l)B}. \end{aligned}$$

This proves (15).

Lemma 2. *If* $n \geqslant 1$,

$$\mathrm{P}(X_n = B |\, Y_0 = A) = p^{(n-1)}_{H(A)B}. \tag{46}$$

Proof. (46) holds for $n = 1$ ($p^{(0)}_{AB} = \delta_{AB}$). If it holds for some n then

$$\begin{aligned} \mathrm{P}(Y_0 = A,\, X_{n+1} = B) &= \sum_{C}\sum_{H(D)=B} \mathrm{P}(Y_0 = A,\, X_n = C,\, Y_n = D) \\ &= \sum_{C}\sum_{H(D)=B} z_{CD}\mathrm{P}(Y_0 = A,\, X_n = C) \quad \text{by (30)}. \end{aligned}$$

Hence

$$P(X_{n+1} = B|Y_0 = A) = \sum_C \sum_{H(D)=B} z_{CD} P(X_n = C|Y_0 = A)$$
$$= \sum_C p_{CB}\, p_{H(A)C}^{(n-1)} \quad \text{by (12) and (46)}$$
$$= p_{H(A)B}^{(n)}.$$

Thus (46) is true by induction.

Proof of Theorem 3.

$$q_{AB}^{(n)} = P(Y_n = B|Y_0 = A)$$
$$= \sum_C P(X_n = C,\ Y_n = B|Y_0 = A)$$
$$= \sum_C z_{CB} P(X_n = C|Y_0 = A) \quad \text{by (30)}$$
$$= \sum_C z_{CB}\, p_{H(A)C}^{(n-1)} \quad \text{by (46).}$$
$$p_{AB}^{(n)} = P(X_n = B|X_0 = A)$$
$$= \sum_C \sum_{H(D)=B} P(Y_0 = C,\ Y_{n-1} = D|X_0 = A)$$
$$= \sum_C \sum_{H(D)=B} z_{AC} P(Y_{n-1} = D|Y_0 = C),$$

which gives (17).

Now suppose $\mathscr{C}$ is a positive class of $\{X_n\}$ with limiting probabilities π_B, so that

(47) $$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{AB}^{(n)} = \pi_B \quad \text{if } A, B \in \mathscr{C}.$$

Let $\mathscr{C}'$ be given by (18). $\mathscr{C}' \subset H^{-1}(\mathscr{C})$ since $\mathscr{C}$ is closed. It can be seen from (16) that for $A \in H^{-1}(\mathscr{C})$, $q_{AB}^{(n)}$ can be non-zero only if $B \in \mathscr{C}'$.

From (16),

(48) $$\frac{1}{N} \sum_{n=1}^{N} q_{AB}^{(n)} = \sum_{C\in\mathscr{C}} \frac{1}{N} \sum_{n=1}^{N} p_{H(A)C}^{(n-1)} z_{CB} \quad (A, B \in \mathscr{C}').$$

Noting that the sum over C is finite, we can let $N \to \infty$ to get

(49) $$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} q_{AB}^{(n)} = \sum_{C\in\mathscr{C}} \pi_C z_{CB} = \varrho_B,$$

and so ϱ_B are the limiting probabilities of $\mathscr{C}'$; (49) also confirms that $\mathscr{C}'$ is a positive class, since $\varrho_B > 0$.

Next suppose $\mathscr{C}'$ is a positive class of $\{Y_n\}$ with limiting probabilities ϱ_B. Let $\mathscr{C}$ be given by (20). For $A \in \mathscr{C}$, z_{AC} can be non-zero only if $C \in \mathscr{C}'$, since $\mathscr{C}'$ is closed. Then similarly $q_{CD}^{(n-1)}$ can be non-zero only if $D \in \mathscr{C}'$. Hence by (17), $p_{AB}^{(n)}$ can be non-zero only if $B \in \mathscr{C}$; that is, $\mathscr{C}$ is closed.

Taking Cesàro limits of (17) for $A, B \in \mathscr{C}$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{AB}^{(n)} = \sum_{C\in\mathscr{C}'} \sum_{\substack{D\in\mathscr{C}' \\ H(D)=B}} z_{AC}\, \varrho_D = \sum_{H(D)=B} \varrho_D = \pi_B;$$

again, this verifies that $\mathscr{C}$ is a positive class.

Proof of Theorem 4. It follows from Theorems 3 and 4, p. 31 of [1] that

(50) $$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{AB}^{(n)} = f_{A\mathscr{C}}\, \pi_B \quad \text{if} \quad B \in \mathscr{C}.$$

Also, by Theorem 1, p. 33 of [1],

(51) $$\sum_{B\in\mathscr{C}} \pi_B = 1,$$

so that

(52) $$\sum_{\mathscr{C}} f_{A\mathscr{C}} \sum_{B\in\mathscr{C}} \pi_B = 1.$$

(22) results by taking Cesàro limits of both sides of (13), applying (50) and (52).

The proof of (23) is similar, taking limits of (15).

Proof of Theorem 5. By Corollary 1, [1], p. 87,

(53) $$P(\lim_{N\to\infty} f_{NB} = \varrho_B | \{Y_n\} \text{ enters } \mathscr{C}') = 1$$

for all $B \in \mathscr{C}'$, where

(54) $$f_{NB} = \frac{1}{N} \operatorname{card}\{n;\ 0 \leqslant n < N,\ Y_n = B\},$$

so that

(55) $$f_{Nj} = \sum_{\substack{B\in\mathscr{C}' \\ B\subset B(j,m)}} f_{NB}.$$

By using (53) and (55), with the analogue of (51) for $\{Y_n\}$

(56) $$\sum_{B\in\mathscr{C}'} \varrho_B = 1,$$

it is proved that

$$\mathrm{P}\Big(\lim_{N\to\infty} f_{Nj} = \sum_{\substack{B\in\mathscr{C}' \\ B\subset B(j,m)}} \varrho_B \,|\, \{Y_n\} \text{ enters } \mathscr{C}'\Big) = 1.$$

By (19) we also have

$$\sum_{\substack{B\in\mathscr{C}' \\ B\subset B(j,m)}} \varrho_B = \sum_{C\in\mathscr{C}} \pi_C \sum_{\substack{B\in\mathscr{C}' \\ B\subset B(j,m)}} z_{CB} = \sum_{C\in\mathscr{C}} \pi_C \,\mathrm{P}\big(B(j,m)|\,C\big),$$

using (18).

This completes the proof of Theorem 5.

Proof of Theorem 6. Let $\mathscr{P}$ be the set of positive states of $\{X_n\}$. For any initial state A we have ([1], p. 31)

$$\lim_{n\to\infty} p_{AB}^{(n)} = 0 \tag{57}$$

if $B\notin\mathscr{P}$. Hence

$$\lim_{n\to\infty} \mathrm{P}(X_n\notin\mathscr{P}|X_0 = A) = \sum_{B\notin\mathscr{P}} \lim_{n\to\infty} \mathrm{P}(X_n = B|X_0 = A) = 0,$$

that is

$$\lim_{n\to\infty} \mathrm{P}(X_n\in\mathscr{P}|X_0 = A) = 1, \tag{58}$$

so

$$\mathrm{P}(X_n\in\mathscr{P} \text{ for some } n|X_0 = A) = 1, \tag{59}$$

since $\mathscr{P}$ is finite.

Lemma 3. *If $\{T^n(x)\}$ is a divergent trajectory with $T^n(x)\equiv i_n \pmod d$ $(0\leqslant i_n < d)$ then*

$$|T^n(x)|^{1/n} \sim \prod_{k=0}^{n-1} \left|\frac{m_{i_k}}{d}\right|^{1/n} \quad \textit{as} \quad n\to\infty \tag{60}$$

and

$$\liminf_{n\to\infty} \prod_{k=0}^{n-1} \left|\frac{m_{i_k}}{d}\right|^{1/n} \geqslant 1. \tag{61}$$

This proves Corollary (iii) to Conjecture 1.

Proof. By considering only the tail of the trajectory, we may assume that $T^n(x)$ is not zero for any n. Then it is easily proved by induction that

$$T^n(x) = \frac{m_{i_0}\cdots m_{i_{n-1}}}{d^n}\, x \prod_{k=0}^{n-1} \left\{1 - \frac{r_{i_k}}{m_{i_k} T^k(x)}\right\}. \tag{62}$$

Because $|T^k(x)|\to\infty$ as $k\to\infty$,

$$\frac{r_{i_k}}{m_{i_k} T^k(x)} \to 0 \quad \text{as} \quad k\to\infty,$$

so

$$\lim_{n\to\infty} \prod_{k=0}^{n-1} \left|1 - \frac{r_{i_k}}{m_{i_k} T^k(x)}\right|^{1/n} = 1.$$

Therefore (62) implies (60).

(61) follows immediately from (60), since $|T^n(x)|\geqslant 1$.

4. The results of Matthews and Watts.

Theorem 7. *If $(m_i, d^2) = (m_i, d)$ for all i (so that $d_i|d$), the Markov chains are finite and Y_n is a residue class* $\bmod\, l$ *for all n.*

Most of the results of [3] now follow.

Proof. By (7) and (8), Y_0 is a residue class modulo l. If Y_n is a residue class $\bmod\, l$, X_{n+1} is a residue class $\bmod\, l\dfrac{d_i}{d}$ for some i. And $l\dfrac{d_i}{d}\Big| l$ since $d_i|d$. Therefore Y_{n+1} is a residue class $\bmod\, l$. Thus the theorem holds by induction.

Theorem 8. *If $(m_i, d) = 1$ for all i (so $d_i = 1$),*

$$\mathrm{P}\big(\{T^n(x)\} \text{ is uniformly distributed} \bmod d^\alpha \text{ for each } \alpha\geqslant 1\big) = 1. \tag{63}$$

This was proved in [2] (Theorem 3) using ergodic theory. It is worthwhile to give a proof via probability theory.

Proof. Take $m = d^\alpha$. Y_n is a residue class $\bmod\, d^\alpha$, by Theorem 7. The result is proved by showing that

$$q_{AB}^{(\alpha)} = 1/d^\alpha \tag{64}$$

for all A and B. Then

$$q_{AB}^{(n+\alpha)} = \sum_C q_{AC}^{(n)} q_{CB}^{(\alpha)} = \frac{1}{d^\alpha},$$

so $\{Y_n\}$ is ergodic with limiting probabilities $1/d^\alpha$. Thus (63) will be true by Theorem 5.

$$q_{AB}^{(\alpha)} = \sum_{C_1,\ldots,C_{\alpha-1}} q_{AC_1} q_{C_1C_2} \cdots q_{C_{\alpha-1}B} = \sum z_{H(A)C_1} z_{H(C_1)C_2} \cdots z_{H(C_{\alpha-1})B}.$$

$H(A)$ and $H(C_i)$ are residue classes $\bmod\, d^{\alpha-1}$, by (6) since $d_i = 1$ for all i. Therefore

(65)

$$q_{AB}^{(\alpha)} = \frac{1}{d^\alpha}\,\mathrm{card}\,\{(C_1,\ldots,C_{\alpha-1});\ C_1\subset H(A),\ C_2\subset H(C_1),\ \ldots,\ B\subset H(C_{\alpha-1})\}.$$

The condition $C_1 \subset H(A)$ gives d distinct choices for C_1, all congruent $\bmod\, d^{\alpha-1}$. Similarly there are d distinct choices for C_i corresponding to each value of C_{i-1}. This gives d^i distinct choices altogether for C_i, all congruent $\bmod\, d^{\alpha-1}$ $(2 \leqslant i \leqslant \alpha-1)$.

Of the $d^{\alpha-1}$ possible values of $H(C_{\alpha-1})$, exactly one will contain B. Thus (65) becomes

$$q_{AB}^{(\alpha)} = 1/d^{\alpha},$$

as required.

5. A method for calculating the limiting frequencies. Let D be a subset of $\mathscr{B}$ which contains Z_l as an element.

If $B = B(x, M)$ and $l|M$, let $A = \tilde{H}(B)$ minimize $P(A)$ subject to the conditions $A \in \mathscr{D}$ and $H(B) \subset A$. Define the sequence $\{\tilde{X}_n\}_{n \geqslant 0}$ by

(66) $$\tilde{X}_0 = Z_l,$$

(67) $$\tilde{X}_{n+1} = \tilde{H}\big(G(\tilde{X}_n, T^n(x))\big),$$

with G as in (5).

$\{\tilde{X}_n\}$ is not in general a Markov chain; the essence of this method is that good approximations are obtained by treating it as a Markov chain. The approximations are thought to become exact in the limit $\mathscr{D} \to \mathscr{B}$, i.e. as $\mathscr{D}$ increases to include every element of $\mathscr{B}$ (see Conjecture 2 below, and Example 3).

Let

(68) $$\tilde{p}_{AB} = \sum_{\substack{C \in \mathscr{B} \\ \tilde{H}(C) = B}} z_{AC}$$

for $A, B \in \mathscr{D}$. Let $\tilde{\pi}_B$ be the limiting probabilities of $\{\tilde{X}_n\}$ using these transition probabilities, a different set of $\tilde{\pi}_B$ corresponding to each positive class $\mathscr{C}$ of $\{\tilde{X}_n\}$. Let

(69) $$\tilde{p}_j = \sum_{B \in \mathscr{C}} \tilde{\pi}_B \, \mathrm{P}(B(j, m)|B).$$

CONJECTURE 2. *In the limit $\mathscr{D} \to \mathscr{B}$, a one-to-one correspondence emerges between classes of $\{X_n\}$ and classes of $\{\tilde{X}_n\}$. If $\mathscr{C}$ is a positive class in $\{X_n\}$ and in $\{\tilde{X}_n\}$, then*

(70) $$\lim_{\mathscr{D} \to \mathscr{B}} \tilde{p}_j = p_j,$$

with p_j as in (29). *If $\mathscr{C}$ is a positive class in $\{\tilde{X}_n\}$ only, then $\lim_{\mathscr{D} \to \mathscr{B}} \tilde{p}_j$ still exists, and* (28) *still holds with p_j equal to this limit.*

6. Examples.

EXAMPLE 1.

(71) $$T(x) = \begin{cases} 3x-1, & x \equiv 0 \pmod 3, \\ (x-16)/3, & x \equiv 1, \\ (-4x-7)/3, & x \equiv 2, \end{cases}$$

$m = 2$. Use the Markov chain $\{X_n\}$ defined by (7)–(9). Possible transitions are shown in Figure 1. The Markov chain is finite, the only states required being $B(0,2)$, $B(1,2)$, $B(5,6)$, $B(8,18)$ and $B(17,18)$ ($B(0,1)$ is omitted becuase it never re-occurs.). The matrix of transition probabilities is

(72) $$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The first and last states are transient; the limiting probabilities of the states $B(1,2)$, $B(5,6)$ and $B(8,18)$ are $\frac{3}{5}$, $\frac{1}{5}$ and $\frac{1}{5}$. $B(1,2)$ and $B(5,6)$ are both contained in $B(1,2)$, while $B(8,18) \subset B(0,2)$. Therefore $p_1 = \frac{3}{5}+\frac{1}{5} = \frac{4}{5}$, and $p_0 = \frac{1}{5}$ (p_j as in (29)).

Conjecture 1 (ii) predicts that every divergent trajectory will enter one of the states $B(1,2)$, $B(5,6)$, $B(8,18)$. Conjecture 1 (i) then says that the limiting frequences of even and odd iterates will be $\frac{1}{5}$ and $\frac{4}{5}$ respectively.

(13) also gives exact values for the numbers

(73) $$p_{nij} = \frac{1}{3^n} \operatorname{card}_{2 \cdot 3^n} \{x;\ x \equiv i \pmod 2,\ T^n(x) \equiv j \pmod 2\}$$

in terms of the elements of P^n $(i, j = 0, 1)$. (26) gives

(74) $$p_{n00},\ p_{n10} \to \tfrac{1}{5},$$

(75) $$p_{n01},\ p_{n11} \to \tfrac{4}{5} \quad \text{as} \quad n \to \infty$$

(The Cesàro limits may be omitted since P is aperiodic.)

$$B(0,1) \xrightarrow[z=1/6]{G} \begin{cases} B(0,6) \xrightarrow{H} B(17,18) \xrightarrow{H} B(5,6) \\ B(1,6) \xrightarrow{H} B(1,2) \\ B(2,6) \xrightarrow{H} B(1,2) \end{cases} \xrightarrow[z=1/3]{G} \begin{cases} B(1,6) \xrightarrow{H} B(1,2) \\ B(3,6) \xrightarrow{H} B(8,18) \\ B(5,6) \xrightarrow{H} B(1,2) \end{cases}$$

$$\begin{cases} B(3,6) \xrightarrow{H} B(8,18) \xrightarrow{H} B(5,6) \\ B(4,6) \xrightarrow{H} B(0,2) \\ B(5,6) \xrightarrow{H} B(1,2) \end{cases} \xrightarrow[z=1/3]{G} \begin{cases} B(0,6) \xrightarrow{H} B(17,18) \\ B(2,6) \xrightarrow{H} B(1,2) \\ B(4,6) \xrightarrow{H} B(0,2) \end{cases}$$

Fig. 1. Finding transition probabilities for Example 1

EXAMPLE 2.

$$T(x)=\begin{cases} x/4, & x\equiv 0 \pmod 8,\\ (x+1)/2, & x\equiv 1\\ 20x-40, & x\equiv 2\\ (x-3)/8, & x\equiv 3\\ 20x+48, & x\equiv 4\\ (3x-13)/2, & x\equiv 5\\ (11x-2)/4, & x\equiv 6\\ (x+1)/8, & x\equiv 7,\end{cases} \tag{76}$$

$m=8$. From Figure 2 it can be seen that $\{X_n\}$ has only the 5 states

$$B(0,1),\quad B(1,4),\quad B(0,2),\quad B(0,8),\quad B(0,32),$$

and the transition probability matrix is

$$\begin{bmatrix} \frac14 & \frac14 & \frac14 & \frac18 & \frac18\\ . & 1 & . & . & .\\ . & . & \frac12 & \frac14 & \frac14\\ . & . & 1 & . & .\\ . & . & . & 1 & . \end{bmatrix}$$

where zeros are represented by dots. $B(0,1)$ is transient and there are two positive classes $\mathscr{C}_1=\{B(1,4)\}$ and $\mathscr{C}_2=\{B(0,2),\ B(0,8),\ B(0,32)\}$.

Conjecture 1 (ii) predicts that every divergent trajectory enters either $\mathscr{C}_1$ or $\mathscr{C}_2$. For $\mathscr{C}_1$ we have $p_1=p_5=\frac12$, with all other $p_j=0$. We then calculate

$$(\tfrac12)^{1/2}(\tfrac32)^{1/2}=\sqrt{3}/2<1,$$

so by Conjecture 1 (iii), every trajectory which enters $\mathscr{C}_1$ should be eventually periodic.

The limiting probabilities of $\mathscr{C}_2$ are $\frac47$, $\frac27$, $\frac17$, respectively.

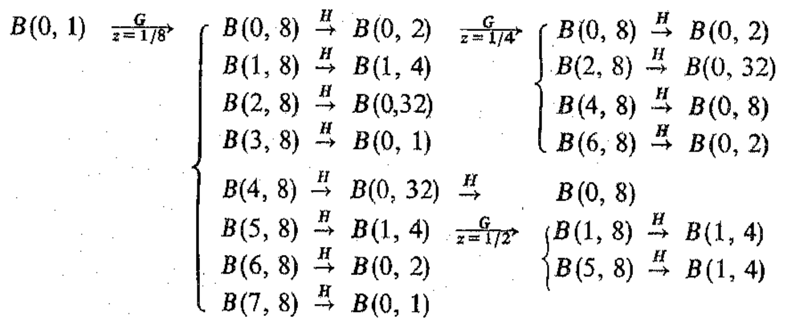

Fig. 2. Transition probabilities for Example 2

Then $p_0=\frac14\cdot\frac47+\frac27+\frac17=\frac47$ and $p_2=p_4=p_6=\frac14\cdot\frac47=\frac17$. Now apply Conjecture 1 (i). Therefore we expect all divergent trajectories to exhibit frequences $\frac47$, $\frac17$, $\frac17$, $\frac17$, of iterates congruent to 0, 2, 4 and 6 respectively modulo 8.

EXAMPLE 3.

$$T(x)=\begin{cases} 40x/3, & x\equiv 0 \pmod 6)\\ (5x-1)/2, & x\equiv 1\\ (7x-2)/6, & x\equiv 2\\ 35x/3, & x\equiv 3\\ (8x-2)/3, & x\equiv 4\\ (5x-1)/6, & x\equiv 5,\end{cases} \tag{77}$$

$m=6$. Here the Markov chain is infinite and all states appear to be transient. We approximate it using the method of Section 5. Take

$$\mathscr{D}=\{B(x,2^a3^b);\ a\leqslant 4,\ x\in Z\}. \tag{78}$$

As shown in Figure 3, this results in just 9 states

$$B(0,1),\quad B(1,2),\quad B(2,3),\quad B(0,4),\quad B(5,8),\quad B(6,8),$$
$$B(8,12),\quad B(0,16),\quad B(10,16)$$

with transition matrix

$$\begin{bmatrix} \frac13 & \frac16 & \frac16 & . & . & . & . & \frac16 & \frac16\\ \frac13 & \frac13 & \frac13 & . & . & . & . & . & .\\ 1 & . & . & . & . & . & . & . & .\\ . & \frac13 & . & . & . & . & . & \frac13 & \frac13\\ . & . & . & \frac13 & \frac13 & . & \frac13 & . & .\\ . & . & . & \frac13 & . & . & . & \frac13 & \frac13\\ . & 1 & . & . & . & . & . & . & .\\ . & . & . & . & \frac13 & . & . & \frac13 & \frac13\\ . & . & . & . & . & \frac13 & . & \frac13 & \frac13 \end{bmatrix}$$

This matrix is ergodic and its vector of limiting probabilities is

$$\frac{1}{536}[84\quad 63\quad 35\quad 30\quad 54\quad 36\quad 18\quad 108\quad 108].$$

Then by (69),

$$\tilde{p}_0=\frac{1}{536}\left(\frac16 84+\frac13 30+\frac13 36+\frac13 108+\frac13 108\right)=\frac{27}{134}=.2015,$$

$$\tilde{p}_1=\frac{53}{536}=.0989,$$

$$\tilde{p}_2 = \frac{287}{1072} = .2677,$$

$$\tilde{p}_3 = \frac{53}{536} = .0989,$$

$$\tilde{p}_4 = \frac{27}{134} = .2015,$$

$$\tilde{p}_5 = \frac{141}{1072} = .1315.$$

We expect all divergent trajectories to have limiting frequences of approximately these values. If Conjecture 2 holds, the approximations will approach the actual limiting frequencies as we take

(79) $$\mathscr{D} = \{B(x, 2^a 3^b);\ a \leqslant A\}$$

and let $A \to \infty$.

After only 1000 iterations, starting at $x = 53$, the frequencies are

.2110, .0910, .2670, .0950, .1980, .1380.

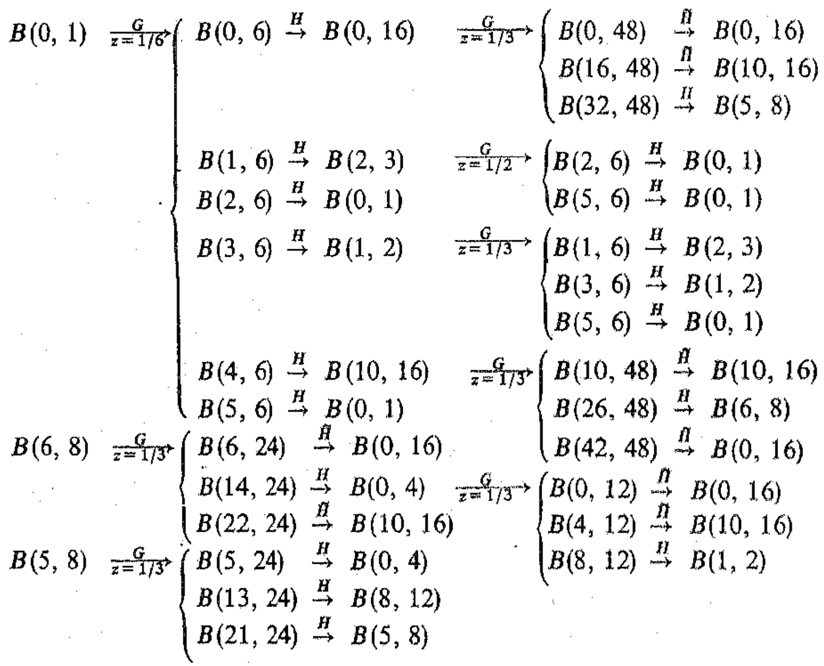

Fig. 3. Transition probabilities for Example 3

References

[1] K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, Berlin 1960.
[2] K. R. Matthews and A. M. Watts, *A generalization of Hasse's generalization of the Syracuse algorithm*, Acta Arith. 43 (1984), pp. 167–175.
[3] — — *A Markov approach to the generalized Syracuse algorithm*, ibid. 45 (1985), pp. 29–42.

Received on 7. 6. 1984 (1431)