

# An Ergodic Source Which Is not a Function of Any Finite Markov Process\*

Roger D. Traub\*\*

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(June 19, 1971)

The problem of deciding when a stationary stochastic process is a function of a finite Markov process has been considered by several authors, leading to an elegant necessary and sufficient condition. Taking a different approach, this note uses elementary ring theory to prove that a certain explicitly constructed stationary ergodic process is not a function of any finite Markov process.

Key words: Information theory; Markov process; probability theory; ring theory; stochastic process.

## 1. Introduction

The problem of deciding when a stationary stochastic process is a function of a finite Markov process has been considered by several authors ([1, 2, 4–6]),<sup>1</sup> leading to an elegant necessary and sufficient condition (Heller, [6]). This condition involves both (a) the finite dimensionality of a certain module over the real numbers, and (b) the requirement that a certain cone in this module be polyhedral (an idea used earlier by Dharmadhikari [2]). The output symbols of the stochastic process act as linear operators on the above-mentioned module.

Taking a different approach, this note uses elementary ring theory to prove that a certain specific stationary ergodic process is not a function of any finite Markov process. The structure of this process—although not the idea—resembles that of an example of Fox and Rubin [4].

Section 2 presents the principal tool for the verification, namely a sufficient condition for a “source” (in the language of information theory) *not* to be a function of any finite Markov process. This condition and its verification are purely algebraic.

Section 3 gives the explicit description of the example, plus the verification of its properties. A related conjecture is posed in section 4.

The balance of the present section 1 consists of definitions and other background material provided to make the exposition more nearly self-contained.

Let  $A$  be a fixed finite nonempty set; the sources to be considered emit “symbols” from this “alphabet”  $A$  as outputs. Let  $A^\infty$  be the set of all doubly infinite sequences  $\mathbf{x} = \{x_t\}_{-\infty}^{\infty}$  of members of  $A$ ; each such sequence can be considered a possible time history of the source’s outputs, with symbol  $x_t$  emitted at time  $t$ . The *shift* operator  $T: A^\infty \rightarrow A^\infty$  is defined by

$$T\mathbf{x} = \mathbf{y} \text{ where } y_t = x_{t+1}. \quad (1.1)$$

A *cylinder set* is a subset of  $A^\infty$  of the form

$$C(J, g) = \{\mathbf{x} \in A^\infty : x_t = g(t) \text{ for all } t \in J\}$$

AMS Subject Classification: Primary 62B5.

\* Helpful discussions with A. J. Goldman, J. Levy and M. Newman are gratefully acknowledged.

\*\* National Bureau of Standards and University of Pennsylvania Medical School.

<sup>1</sup> Figures in brackets indicate the literature references at the end of this paper.

where  $J$  is some finite subset of the integers (positive, negative, or zero) and  $g: J \rightarrow A$  is some function. If  $J = \{1, 2, \dots, k\}$  for some  $k$ , then  $C(J, g)$  is called an *elementary* cylinder set. It follows from the finiteness of  $A$  that every cylinder set is a finite union of elementary cylinder sets.

Let  $\Sigma$  be the Borel field generated by all cylinder sets (or equivalently, all elementary cylinder sets). We now *define* an information source over  $A$  to be a probability measure  $\mu$  on  $(A^\infty, \Sigma)$ . Such a source will be denoted  $[A, \mu]$ . In terms of the notation in (1.1), the source is called *stationary* if  $\mu(TS) = \mu(S)$  for every measurable subset  $S$  of  $A^\infty$ , and is called *ergodic* if  $TS = S$  (where  $S$  is measurable) implies  $\mu(S) = 0$  or  $\mu(S) = 1$ .

Let  $I_n = \{1, 2, \dots, n\}$ , and suppose there is a function  $f: I_n \rightarrow A$ , and an  $n$ -state Markov chain with transition-matrix entries  $p(i, j)$  and stationary probabilities  $p(i)$ , such that, for every elementary cylinder set

$$C = C(\{1, 2, \dots, k\}, g),$$

we have

$$\mu(C) = \sum p(i_1)p(i_1, i_2) \dots p(i_{k-1}, i_k) \quad (1.2)$$

where the sum is over all sequences  $\{i_1, i_2, \dots, i_k\}$  over  $I_n$  such that  $f(i_t) = g(t)$  for  $t = 1, 2, \dots, k$ . Note that this entirely determines  $\mu$ , so that  $[A, \mu]$  is specified by the finite set of data  $\{p(i), p(i, j)\}$  together with  $f$ . In this case,  $[A, \mu]$  is called a *function of a (finite) Markov process*; it is automatically stationary, and is ergodic if the Markov chain is.

If  $I_n$ , in the last paragraph, is replaced by the set of all natural numbers, we are led instead to the concept of "function of a denumerable-state Markov process." The example to be given later falls into this category, but not into the previous one, as will be proven.

## 2. An Algebraic Condition

Let  $\Pi_N$  denote the set of all polynomials in  $N$  variables with integer coefficients. For the information source  $[A, \mu]$ , define the following subring of the ring of real numbers:

$$V(A, \mu) = \{\rho : \rho = P(\mu(C_1), \dots, \mu(C_N)) \text{ for some } N > 0, \text{ some } P \in \Pi_N,$$

and some cylinder sets  $C_1, C_2, \dots, C_N\}$ .

Suppose in particular that  $[A, \mu]$  is a function of an  $n$ -state Markov chain with parameters  $\{p(i), p(i, j)\}$ . Then it follows by (1.2) that the measure of any elementary cylinder set is the value, at  $(p(1), \dots, p(n), p(1, 1), \dots, p(n, n))$  of a polynomial in  $n + n^2$  variables with integer coefficients. It follows that the same is true of the measure of *any* cylinder set, and so is true of every member of  $V(A, \mu)$ . That is, if we set

$$W_N(u_1, \dots, u_N) = \{\rho : \rho = P(u_1, \dots, u_N) \quad \text{for some } P \in \Pi_N\},$$

then we have,

$$V(A, \mu) \subseteq W_{n+n^2}(p(1), \dots, p(n), p(1, 1), \dots, p(n, n))$$

whenever  $[A, \mu]$  is a function of the  $n$ -state Markov chain with the indicated parameters.

This observation provides the basis for the following result.

**THEOREM 1:** *Suppose there are real numbers  $\{\rho_i\}_1^m$  such that the ring generated over the integers by  $V(A, \mu) \cup \{\rho_i\}_1^m$  includes all the rationals. Then the source  $[A, \mu]$  is not a function of any finite Markov process.*

This will be shown to follow from the next theorem, whose conclusion, in order to support an inductive proof, is stronger than is needed for Theorem 1.

**THEOREM 2:** Let  $U = (u_1, \dots, u_N)$  be a fixed  $N$ -tuple of real numbers,  $k$  a fixed positive integer,  $(D_1, \dots, D_k)$  fixed polynomials in  $\Pi_N$  with each  $D_j(U) \neq 0$ . Then the set  $R_N = \{\rho : \rho = P(U) / \Pi_1^k [D_j(U)]^{a(j)} \text{ for some } P \in \Pi_N \text{ and some integral } a(j) \geq 0\}$  does not include all the rationals.

We first show that Theorem 2 implies Theorem 1. In Theorem 2 let  $N = m + n + n^2$ , let  $U = (\rho_1, \dots, \rho_m, p(1), \dots, p(n), p(1, 1), \dots, p(n, n))$ , and let all  $D_j \equiv 1$ . It follows from (2.1) that, if  $[A, \mu]$  is a function of the Markov process with the indicated parameters  $\{p(i), p(i, j)\}$ , then  $V(A, \mu) \subseteq R_N$ . Since  $\rho_i \in R_N$ , and  $R_N$  is a ring over the integers, the ring  $R^*$  generated over the integers by  $V(A, \mu) \cup \{\rho_i\}_1^m$  also lies in  $R_N$ . By Theorem 2,  $R^*$  cannot contain all rationals, and so Theorem 1 is proved. It is interesting that the proof uses only (1.2), and not the further properties

$$p(i, j) \geq 0, \quad p(i) \geq 0,$$

$$\sum_j p(i, j) = 1, \quad \sum_i p(i) p(i, j) = p(j).$$

We turn now to the proof of Theorem 2, based on a suggestion by M. Newman. The notation  $R[z_1, \dots, z_n]$  will mean the ring of polynomials in the indeterminates  $z_1, \dots, z_n$  with coefficients in a ring  $R$ . The proof is by induction on  $N$ .

**CASE 1:  $N = 1$ .** Suppose the conclusion is false, i.e., that there is  $U = u_1$  such that  $R_1$  includes all the rationals. Let  $I$  denote the ring of integers.

$U = u_1$  must be algebraic over the rationals, since for any prime  $p$ ,  $1/p \in R_1$ , which implies an equation of the form

$$\Pi_1^k [D_j(U)]^{a(j)} - pP(U) = 0$$

in which not all coefficients can vanish if  $p$  is chosen so as not to divide all coefficients of the polynomial  $\Pi_1^k D_j$ . Now let  $u_i = u_1^{(i)}$  and let  $u_1^{(i)}$  ( $2 \leq i \leq m$ ) be the  $i$ th conjugate of  $u_1$  over the rationals, so that  $\Pi_1^m (z - u_1^{(i)})$  is the minimal polynomial of  $u_1$  over the rationals. Then for a suitable integer  $\lambda$ ,

$$g(z) = \lambda \Pi_1^m (z - u_1^{(i)}) \in \Pi_1.$$

Let  $\{\sigma_i\}_1^m$  denote the elementary symmetric functions of  $m$  variables, and put

$$s_i = \sigma_i(u_1^{(1)}, \dots, u_1^{(m)}),$$

so that

$$g(z) = \lambda (z^m - s_1 z^{m-1} + \dots + (-1)^m s_m).$$

It follows that all  $\lambda s_i \in I$ .

Let  $p$  be a prime, to be determined later. As above,  $1/p \in R_1$  leads to an equation  $f(u_1) = 0$  where

$$f \equiv \Pi_1^k (D_j)^{a(j)} - pp$$

for some appropriate  $P \in \Pi_1$  and nonnegative integers  $a(j)$ . Since  $g$  is a constant multiple of the minimal polynomial of  $u_1$ ,  $f(u_1) = 0$  implies that  $f = gh$  for some polynomial  $h$  with rational coefficients. Since  $g(u_1^{(i)}) = 0$  for all  $i$ , all  $f(u_1^{(i)}) = 0$  and so

$$1/p = P(u_1^{(i)}) / \prod_1^k [D_j(u_1^{(i)})]^{a(j)},$$

implying

$$1/p^m = \prod_1^m P(u_1^{(i)}) / \prod_{i=1}^m \prod_{j=1}^k [D_j(u_1^{(i)})]^{a(j)}. \quad (2.2)$$

In the above note that

$$\prod_1^k [D_j(u_1^{(i)})]^{a(j)}$$

cannot be zero. For otherwise some  $D_j(u_1^{(i)}) = 0$ , implying  $D_j(u_1) = 0$ , contrary to hypothesis.

Now consider the symmetric polynomials

$$P^*(z_1, \dots, z_m) = \prod_1^m P(z_i),$$

$$D_j^*(z_1, \dots, z_m) = \prod_1^m D_j(z_i).$$

By [9; p. 97], there are polynomials  $\pi, \delta_j \in I[z_1, \dots, z_m]$  such that

$$P^*(z_1, \dots, z_m) = \pi[\sigma_1(z_1, \dots, z_m), \dots, \sigma_m(z_1, \dots, z_m)],$$

$$D_j^*(z_1, \dots, z_m) = \delta_j[\sigma_1(z_1, \dots, z_m), \dots, \sigma_m(z_1, \dots, z_m)].$$

It follows from (2.2) that

$$1/p^m = \pi(s_1, \dots, s_m) / \prod_1^k [\delta_j(s_1, \dots, s_m)]^{a(j)}. \quad (2.3)$$

Let  $d(j) = \deg[\delta_j]$ , and choose integer  $b > 0$  so that

$$d = \sum_1^k d(j)a(j) + b \geq \deg[\pi];$$

since all  $\lambda s_i \in I$ , this implies that  $\lambda^d \pi(s_1, \dots, s_m)$  is an integer  $w$ . Note also that  $\lambda^{d(j)} \delta_j(s_1, \dots, s_m)$  is an integer  $w_j$  independent of  $p$ .

Now, (2.3) yields

$$1/p = p^{m-1} w / \lambda^b \prod_1^k (w_j)^{a(j)}$$

Choosing  $p$  to be a prime which divides neither  $\lambda$  nor any  $w_j$ , we obtain a contradiction. (Recall that  $\lambda$  does not depend on  $p$ .)

CASE 2:  $N > 1$ . Again we assume that  $R_N$  contains all the rationals. As in the proof for Case 1,  $U = (u_1, \dots, u_N)$  cannot be algebraically independent over the rationals. Without loss of generality, we assume  $u_N$  is algebraically dependent on  $\tilde{U} = (u_1, \dots, u_{N-1})$ . Expand the notation to read

$$R_N = R_N(U; D_1, \dots, D_k).$$

We will show that there exist  $\tilde{D}_1, \dots, \tilde{D}_{k+1} \in \Pi_{N-1}$  with all  $\tilde{D}_j(\tilde{U}) \neq 0$  such that  $R_{N-1}(\tilde{U}; \tilde{D}_1, \dots, \tilde{D}_{k+1})$  contains all the rationals. This yields the desired contradiction of the induction hypothesis.

Let  $R_0$  be the ring generated over the integers by  $\tilde{U}$ , i.e.,

$$R_0 = \{\rho : \rho = P(\tilde{U}) \text{ for some } P \in \Pi_{N-1}\},$$

and let  $F_0$  be the quotient field of  $R_0$ . Then  $u_N$  is algebraic over  $F_0$ ; let  $u_N^{(1)} = u_N$  and let  $u_N^{(i)} (2 \leq i \leq m)$  be the conjugates of  $u_N$  over  $F_0$ , so that  $\Pi_1^m(z - u_N^{(i)})$  is the minimal polynomial of  $u_N$  over  $F_0$ . Then for a suitable  $\lambda \in R_0$ ,

$$g(z) = \lambda \Pi_1^m(z - u_N^{(i)}) \epsilon R_0[z].$$

Define  $\{\sigma_i\}_1^m$  as above, and  $s_i = \sigma_i(u_N^{(1)}, \dots, u_N^{(m)})$ ; then all  $\lambda s_i \in R_0$ .

As in Case 1, we consider a prime  $p$ . The requirement  $1/p \in R_N$  leads as before first to an equation  $f(U) = 0$  where

$$f \equiv \Pi_1^k(D_j)^{a(j)} - pP,$$

for some appropriate  $P \in \Pi_N$  and nonnegative integers  $a(j)$ . This can be rewritten  $f^*(u_N) = 0$  where

$$f^*(z) \equiv \Pi_1^k[D_j(\tilde{U}, z)]^{a(j)} - pP(\tilde{U}, z).$$

Arguing as above we find that all  $f^*(u_N^{(i)}) = 0$ , implying

$$1/p = P(\tilde{U}, u_N^{(i)}) / \Pi_1^k[D_j(\tilde{U}, u_N^{(i)})]^{a(j)}$$

and thus

$$1/p^m = \Pi_1^m P(\tilde{U}, u_N^{(i)}) / \Pi_{i=1}^m \Pi_{j=1}^k [D_j(\tilde{U}, u_N^{(i)})]^{a(j)}. \quad (2.4)$$

Now consider the symmetric polynomials

$$P^*(z_1, \dots, z_m) = \Pi_1^m P(\tilde{U}, z_i) \in R_0[z_1, \dots, z_m],$$

$$D_j^*(z_1, \dots, z_m) = \Pi_1^m D_j(\tilde{U}, z_i) \in R_0[z_1, \dots, z_m].$$

As above, there are polynomials  $\pi, \delta_j \in R_0[z_1, \dots, z_m]$  such that (2.4) yields

$$1/p^m = \pi(s_1, \dots, s_m) / \Pi_1^k [\delta_j(s_1, \dots, s_m)]^{a(j)}; \quad (2.5)$$

as before  $\lambda$ , the  $s_i$  and the  $\delta_j$  are independent of  $p$ . Defining  $d(j)$ ,  $b$ , and  $d$  as above, we have

$$1/p = [p^{m-1} \lambda^d \pi(s_1, \dots, s_m)] / \lambda^b \Pi_1^k (w_j)^{a(j)} \quad (2.6)$$

where  $w_j = \lambda^{d(j)} \delta_j(s_1, \dots, s_m) \in R_0$ . We may denote the numerator of (2.6), which lies in  $R_0$ , as  $\tilde{P}(\tilde{U})$ . Similarly let  $w_j = \tilde{D}_j(\tilde{U})$  for some  $\tilde{D}_j \in \Pi_{N-1}$ , let  $\lambda = \tilde{D}_{k+1}(\tilde{U})$  where  $\tilde{D}_{k+1} \in \Pi_{N-1}$ , and let  $a(k+1) = b$ . Then (2.6) yields

$$1/p = \tilde{P}(\tilde{U}) / \Pi_1^{k+1} [\tilde{D}_j(\tilde{U})]^{a(j)} \epsilon R_{N-1}(\tilde{U}; \tilde{D}_1, \dots, \tilde{D}_{k+1}).$$

Thus if  $R_N(U; D_1, \dots, D_k)$  contains all rationals, or equivalently all prime-reciprocals  $1/p$ , then so does  $R_{N-1}(\tilde{U}; \tilde{D}_1, \dots, \tilde{D}_{k+1})$ .

This completes the proof.

### 3. The Construction

Our example is constructed as a function of a *denumerable* Markov chain  $M$ , which will be described next. The state-space of  $M$  consists of a *special* state  $S_0$ , and an infinite lower-triangular array of states  $S_{ij}$ ; the  $i$ th row of the array consists of the states  $S_{i1}, S_{i2}, \dots, S_{ii}$ . The probability of transition from  $S_0$  to a row-leading state  $S_{i1}$  is  $q_i$ , where the  $q_i$ 's are positive numbers (to be speci-

fied) with  $\sum_i q_i = 1$ ; no other states are directly accessible from  $S_0$ . The other direct transitions are from  $S_{i1}$  to  $S_{i2}$ ,  $S_{i2}$  to  $S_{i3}$ , . . . , and from  $S_{ii}$  back to  $S_0$ , each with probability 1.

We now exhibit a vector  $\alpha$ , with positive components summing to 1, such that

$$\alpha P = \alpha \quad (3.1)$$

where  $P$  is the transition matrix of  $M$ . The entries of  $\alpha$  will be denoted  $p(S_0)$  and  $p(S_{ij})$ , with obvious interpretation. For some  $x > 0$ , set

$$p(S_0) = x, \quad p(S_{ij}) = q_i x. \quad (3.2)$$

The equations of (3.1) which correspond respectively to  $S_0$ , to any  $S_{i1}$ , and to any  $S_{i,j+1}$  are

$$\sum_i p(S_{ii}) \cdot 1 = p(S_0),$$

$$p(S_0) q_i = p(S_{i1}),$$

$$p(S_{ij}) \cdot 1 = p(S_{i,j+1}),$$

and all are satisfied; hence (3.1) holds. It only remains to satisfy the normalization

$$p(S_0) + \sum_{j \leq i} p(S_{ij}) = 1,$$

which identifies  $x$  as

$$x = [1 + \sum_i q_i]^{-1}. \quad (3.3)$$

The  $q_i$  will be chosen so that the above makes sense, i.e.,

$$\sum_i q_i < \infty. \quad (3.4)$$

Take the alphabet  $A = \{0, 1\}$ . If  $\Sigma$  denotes the state-space of  $M$ , define  $f: \Sigma \rightarrow A$  by

$$f(S_0) = 0, \quad f(S_{ij}) = 1.$$

$[A, \mu]$  is now defined as the function of  $M$  determined by  $f$ .

Consider the elementary cylinder set

$$C_i = C(\{1, 2, \dots, i+2\}, g_i)$$

where  $g_i(1) = g_i(i+2) = 0$  and  $g_i = 1$  otherwise. Then  $\mu(C_i)$  is the probability of the source producing a pair of 0's separated by exactly  $i$  1's. The sum (1.2) here collapses to a single term, yielding

$$\mu(C_i) = q_i x.$$

Thus if  $\{q_i\}$  can be selected to include, for each prime  $p$ , a rational number with denominator divisible by  $p$ , then the ring generated over the integers by  $V(A, \mu) \cup \{1/x\}$  will contain all  $q_i$ , hence all reciprocals of primes, hence all rationals. By Theorem 1,  $[A, \mu]$  cannot be a function of any finite Markov process.

Let  $\pi_i$  be the  $i$ th prime ( $\pi_1 = 1$ ), and set

$$q_i = 1/2^i \pi_i \quad (i \geq 2). \quad (3.5)$$

Since  $\pi_i \geq 2$  in (3.5), we have

$$\sum_{i=2}^{\infty} i q_i < \sum_{i=2}^{\infty} i/2^{i+1} < \infty,$$

and so (3.4) holds for *any* finite  $q_i$ ; since

$$\sum_{i=2}^{\infty} q_i < \sum_{i=2}^{\infty} 1/2^{i+1} = 1/4 < 1,$$

there is a unique  $q_1 > 0$  such that  $\sum_{i=1}^{\infty} q_i = 1$ . Thus the goal described in the previous paragraph has now been accomplished.

Although the main discussion is now complete, certain technical points which were passed over above must still be verified. It must be shown that  $M$  is ergodic (as a Markov chain), that the vector  $\alpha$  defined in (3.2) is its unique stationary vector, and that  $[A, \mu]$  is (as promised) ergodic.

With  $\Sigma$  (as above) the state-space of  $M$ , the *identity* map of  $\Sigma$  induces an infinite-alphabet source  $[\Sigma, \mu^*]$  which is a function of  $M$ . Let  $\Sigma^\infty$  be the set of doubly-infinite sequences  $\{\sigma_t\}_{-\infty}^{\infty}$  over  $\Sigma$ , and  $T^*$  the shift operator on  $\Sigma^\infty$ . Then *any* function  $[A, \mu]$  of  $M$ , induced by a function  $f; \Sigma \rightarrow A$ , is ergodic if  $[\Sigma, \mu^*]$  is ergodic. This follows, in terms of the function  $f^\infty: \Sigma^\infty \rightarrow A^\infty$  arising from  $f$ , from the facts

$$f^\infty T^* = T f^\infty; \quad \mu(B) = \mu^*[(f^\infty)^{-1}(B)]$$

for any measurable subset  $B$  of  $A^\infty$ .

Thus it suffices to prove that  $[\Sigma, \mu^*]$  is an ergodic source which, for any  $S \in \Sigma$  and any integer  $t$ , satisfies

$$\mu^*\{\mathbf{x} \in A^\infty: x_t = S\} = \text{probability of } S \text{ as given by } \alpha. \quad (3.6)$$

The reader is now referred to the discussion in [8; p. 52] of limit properties of Markov chains; we shall use the terminology of that discussion. The chain  $M$  is clearly *irreducible* and *nonperiodic*. By (3.4), the mean recurrence time of  $S_0$  is finite, so that  $S_0$  (and hence all states) are *persistent*. By [8; p. 56],  $M$  has a unique invariant instantaneous distribution. Since (3.1) holds, this distribution must be  $\alpha$ ; hence (3.6) holds. In addition, letting  $P$  denote the transition matrix of  $M$ , we have for any  $\sigma, \tau \in \Sigma$

$$\lim_{n \rightarrow \infty} (P^n)_{\tau\sigma} = \alpha_\sigma. \quad (3.7)$$

The fact that  $[\Sigma, \mu^*]$  is ergodic in the sense used in this paper can now be proved from (3.7) by elementary but lengthy arguments; instead we refer the reader to Doob [3; p. 460] for an elegant proof using martingales.

This completes the discussion.

#### 4. Conjecture

According to (2.1), if  $[A, \mu]$  is a function of a finite Markov chain, then  $V(A, \mu)$  is imbedded in a finitely generated ring (of real numbers) over the integers. We conjecture that this necessary condition is *not* sufficient:

CONJECTURE: *There is an information source  $[A, \mu]$  on a finite alphabet  $A$  with the following two properties:*

- (1)  $V(A, \mu)$  is a finitely generated ring over the integers;
- (2)  $[A, \mu]$  is not a function of any finite Markov process.

## 5. References

- [1] Dharmadhikari, S. W., Functions of finite Markov chains. *Ann. Math. Stat.* **34** (1963), 1022–1032.
- [2] Dharmadhikari, S. W., Sufficient conditions for a stationary process to be a function of a finite Markov chain, *Ann. Math. Stat.* **34** (1963), 1033–1041.
- [3] Doob, J. L., *Stochastic Processes* (Wiley, New York, 1953).
- [4] Fox, M., and Rubin, H., Functions of processes with Markovian states, *Ann. Math. Stat.* **39**, 938–946 (1968).
- [5] Gilbert, E., On the identifiability problem for finite Markov chains, *Ann. Math. Stat.* **30**, 688–697 (1959).
- [6] Heller, A., On stochastic processes derived from Markov chains, *Ann. Math. Stat.* **36**, 1286–1291 (1965).
- [7] Khinchin, A. I., *Mathematical Foundations of Information Theory* (Dover, New York, 1957).
- [8] Rosenblatt, M., *Random Processes* (Oxford University Press, New York, 1962).
- [9] van der Waerden, B. L., *Modern Algebra Vol. I* (Ungar, New York, 1953).

(Paper 75B3&4–353)