ON MARKOV CHAIN POTENTIALS[1]

By John G. Kemeny and J. Laurie Snell

Dartmouth College

1. Introduction. In [3] we developed a theory of potentials for denumerable Markov Chains. The purpose of this note is to supplement these results in two ways: We will show for an important special class of Markov chains that they are normal (i.e., that the potential operators exist), and we will generalize certain results due to Spitzer [5].

While our previous paper developed a theory both for transient and for recurrent chains, our present note will deal only with the recurrent case. The key definitions, notations, and theorems for this type of chain will be summarized below. Parenthetical references to theorems will always refer to [3], Section 3.

We consider both measures (row vectors) and functions (column vectors); the former are denoted by Greek letters, the latter by ordinary lower case letters. The theory for functions is dual to that for measures. One passes from one to the other by replacing a transition matrix $\{P_{ij}\}$ by the "reverse chain" $\{\alpha_j P_{ji}/\alpha_i\}$, where $\alpha > 0$ and $\alpha P = \alpha$.

If the limit $\nu = \lim_n [\mu(I + P + \cdots + P^n)]$ exists, we say that ν is a potential, and μ is its charge. The set of states for which μ_i is non-zero is the support of the charge. If $\mathbf{1}$ is the constant function, and if $\mu\mathbf{1}$ is defined, then $\mu\mathbf{1} = 0$. Dually, one defines potential functions. If the column vector f is a charge of a potential function, and αf is finite, then $\alpha f = 0$.

Let $N_{ij}^{(n)}$ be the mean of the number of times that the process is in state j in the first n steps, starting at i. If $\lim_n [N_{jj}^{(n)} - N_{ij}^{(n)}] = C_{ij} \geqq 0$ exists for all i and j, we say that the chain is normal. Under certain assumptions, if ν exists then $\nu = -\mu C$. A sufficient condition is that μ be a weak charge, i.e., that not only μC is finite but also Cf, where $f_i = \mu_i/\alpha_i$ is the dual charge. (See Theorem 15.) For example, all charges of finite support are weak. The dual operator $G_{ij} = \lim_n [N_{ii}^{(n)} \cdot \alpha_j/\alpha_i - N_{ij}^{(n)}]$ serves a similar role for functions. All ergodic (positive recurrent) chains are normal, and the finiteness of μC suffices to assure the existence of the potential.

Many of our considerations will be relative to a given set of states E. Then B_{ij}^E is the probability of entering E at j, starting at i. ${}^E N_{ij}$ is the mean number of times in j, starting at i, before hitting E—this is taken to be 0 if i or j is in E, and we write ${}^k N_{ij}$ if $E = \{k\}$. By P_{ij}^E we mean the probability that starting from i in E we reenter E at j; P^E is itself a recurrent transition matrix, for the states in E. The submatrix of C consisting of rows and columns in E is denoted by C_E.

Received November 4, 1960; revised November 22, 1960.

[1] This research was supported by the National Science Foundation through a grant given to the Dartmouth Mathematics Projects.

Of special interest are the limits $\lim_n P^n B^E$, giving the entrance probabilities "in the long run". If these limits exist, and they always do for a normal chain, then the limits are independent of the starting states, hence the limiting matrix has identical rows λ^E. (See Theorems 16, 20.) The existence of these limits for two-point sets is equivalent to normalcy. (See Theorem 14.) Similarly, we define ${}^E\nu$ to be the common row of the limiting matrix $P^n({}^EN)$.

A chain is ergodic if the mean first passage times M_{ij} are finite, and strong ergodic if the passage times "in equilibrium," αM, are also finite.

Spitzer considered recurrent Markov chains obtained from sums of independent random variables with a common distribution. He assumed that this distribution was a two dimensional symmetric distribution. He showed that these chains are normal. In the first part of this paper we establish this result for the not necessarily symmetric one dimensional case under the assumption of a finite variance.

For transient chains the basic potential operator is $N = I + P + P^2 + \cdots$. In this case for any set E, N_E^{-1} exists and

$$I - P^E = N_E^{-1}.$$

Spitzer showed that for the recurrent chains that he considered, and any finite set E, C_E^{-1} exists. He established an elegant formula for $I - P^E$, using this inverse. In the second part of this paper we give necessary and sufficient conditions that C_E^{-1} exist for finite sets, for any normal recurrent chain. In particular this inverse exists for *all* finite sets for certain symmetric chains and for ergodic chains. We obtain a generalization of the Spitzer formula (corollary to Theorem 2) for all such chains. We use these results to shed new light on certain previous results of ours.

2. A class of normal chains. Let $\{p_j\}$ be a probability distribution on the positive and negative integers. We consider the Markov chain having transition probabilities $P_{ij} = P_{j-i}$. Assume that this chain is started in state 0. Then the resulting random variables $S_0, S_1, \cdots$ represent sums of independent random variables with a common distribution. We assume that S_1 has finite variance σ^2. The mean must be 0 for the chain to be recurrent. We assume that this is the case. We are interested first in studying B_{ij}^E for a finite set E. For convenience we assume that the smallest element of E is 0.

LEMMA 1. *$\lim_{r\to+\infty} B_{ri}^E = B_i^+$ and $\lim_{r\to-\infty} B_{ri}^E = B_i^-$ exist.*

PROOF. We shall prove that $\lim_{r\to-\infty} B_{ri}^E$ exists. We refer to the process $\{S_j\}$ determined by P as the *basic process*, and define an auxiliary process called the *ladder process* as follows: Let $\hat{S}_0 = S_0 = r$. We define $\hat{S}_{n+1}$ to be the first state $> \hat{S}_n$ reached by the basic process. We thus obtain the ladder process which represents the progress of the basic process watched only when it makes progress to the right. This ladder process is again a Markov chain with transition probabilities given by $\hat{P}_{ij} = \hat{p}_{j-i}$, equal to the probability that the basic process started in i reaches a state $> i$ for the first time at state j. Let $\mu = M_0[\hat{S}_1]$, the mean of $\hat{S}_1$ if the process starts at 0. This mean is finite by a theorem of Spitzer [6]. He proved

$$\mu = (\sigma/\sqrt{2})\cdot c$$

where

$$0 < c = \exp\left\{\sum_{1}^{\infty} (1/k)[\tfrac{1}{2} - \Pr[S_k > 0]]\right\} < \infty.$$

We denote by $\hat{B}^F_{rs}$ the probablity that the ladder process started at r reaches the set F of all non-negative integers for the first time at s. Then

$$\begin{aligned} \hat{B}^F_{rs} &= \Pr_r[\hat{S}_n = s \text{ for some } n \text{ and } \hat{S}_m < 0 \text{ for } m < n] \\ &= \Pr_r[\hat{S}_n = s \text{ for some } n] \\ &- \sum_{k=1}^{s} \Pr_r[\hat{S}_{n-1} = s - k \text{ and } \hat{S}_n = s \text{ for some } n]. \end{aligned} \tag{1}$$

By the renewal theorem,

$$\lim_{r\to-\infty} \Pr_r[\hat{S}_n = s \text{ for some } n] = 1/\mu. \tag{2}$$

Let $\hat{B}^F_s = \lim_{r\to-\infty} \hat{B}^F_{rs}$. By (1) and (2) this limit exists and

$$\hat{B}^F_s = 1/\mu\cdot\left(1 - \sum_{j=1}^{s} \hat{p}_j\right) = 1/\mu\cdot\left(\sum_{j=s+1}^{\infty} \hat{p}_j\right).$$

Note that $\sum_s \hat{B}^F_s = 1$. Now

$$B^E_{ri} = \sum_s \hat{B}^F_{rs} B^E_{si}. \tag{3}$$

Since $B^E_{si} \leqq 1$, and since $\sum_s \hat{B}^F_{rs} = \sum_s \hat{B}^F_s = 1$, we have

$$B^-_i = \lim_{r\to-\infty} B^E_{ri} = \sum_s \hat{B}^F_s B^E_{si}.$$

The proof for B^+_i is similar.

THEOREM 1. *Assume that* $\{p_j\}$ *has mean* 0 *and finite variances* σ^2. *Then*

$$\lim_n P^n B^E = \mathbf{1}\cdot\lambda^E$$

where $\lambda^E_i = \tfrac{1}{2}B^+_i + \tfrac{1}{2}B^-_i$.

PROOF. Let $\mathbf{1}^+$ be a column vector with 1 for the non-negative states and 0 otherwise. By the Central Limit Theorem

$$\lim_{n\to\infty} (P^n\mathbf{1}^+)_0 = \tfrac{1}{2}.$$

For a null chain, the probability of being in any finite set at time n tends to $\mathbf{0}$. Hence we have

$$\lim_n P^n\mathbf{1}^+ = \tfrac{1}{2}\cdot\mathbf{1}.$$

For fixed i, let f be a column with value B^+_i on the non-negative states and B^-_i on the negative states. Then

$$\lim_{n\to\infty} P^n f = (\tfrac{1}{2}B^+_i + \tfrac{1}{2}B^-_i)\cdot\mathbf{1}.$$

Let g be the ith column of B^E. Then $f - g$ has limit 0 at $+\infty$ and $-\infty$. Since $P^n \to \mathbf{0}$, it is an easy consequence that

$$\lim_{n\to\infty} P^n(f - g) = \mathbf{0}.$$

Thus

$$\lim P^n g = (\tfrac{1}{2}B_i^+ + \tfrac{1}{2}B_i^-)\cdot \mathbf{1}$$

as was to be proved.

This theorem shows that sums of independent random variables with mean 0 and finite variance always constitute a normal null chain.

3. Restricted potential operators. From now on we assume that we have an arbitrary normal chain. Quantities for the reverse chain will be denoted by ^. Thus $\hat{\lambda}$ in the following theorem is λ computed for the reverse chain.

We will assume from here on that E is a finite set of at least 2 states.

THEOREM 2. *$G_E(I - P^E) = -I + \mathbf{1}\lambda^E$; $(I - P^E)C_E = -I + l\alpha_E$, where $l_i = \hat{\lambda}_i^E/\alpha_i$.*

PROOF. We have shown (Theorem 20) that if we choose a column of $I - P^E$ as charge on a set E, then this is always a weak charge, and the resulting potential is the corresponding column of $B^E - \mathbf{1}\lambda^E$. Hence, $-G\begin{pmatrix} I - P^E \\ 0 \end{pmatrix} = B^E - \mathbf{1}\lambda^E$, where the right side has rows corresponding to all states, but columns corresponding only to the states in E. We obtain the first result by restricting the rows to those in E, while the second result is obtained by duality, i.e., by applying the first result to the reverse chain. (The vector l is the dual of λ, and α is the dual of the constant vector $\mathbf{1}$.)

COROLLARY. *If C_E^{-1} exists, then $(I - P^E) = -C_E^{-1} + l(\alpha_E C_E^{-1})$.*
If G_E^{-1} exists, then $(I - P^E) = -G_E^{-1} + (G_E^{-1}\mathbf{1})\lambda^E$.

THEOREM 3. *There is a measure ω such that $C_E(I - P^E) = -I + \mathbf{1}\omega$.*

PROOF. $(C_E P^E)_{ij} = \sum_{k\epsilon E} \lim_n [N_{kk}^{(n)} - N_{ik}^{(n)}]\cdot P_{kj}^E$. We may interchange the limit with the summation.

$$(C_E P^E)_{ij} = \lim_n \Big[\sum_{k\epsilon E} N_{kk}^{(n)} P_{kj}^E - (N_{ij}^{(n)} - \delta_{ij} + B_{ij}^{(E \text{ after } n \text{ steps})})\Big],$$
$$= C_{ij} + \delta_{ij} - \lambda_j^E + \lim_n \Big[\sum_{k\epsilon E} N_{kk}^{(n)} P_{kj}^E - N_{jj}^{(n)}\Big].$$

We let ω_j be λ_j^E minus the quantity in brackets, and the theorem follows.

COROLLARY. *If $\nu = -\mu C$ is a potential, then $\nu_E(I - P^E) = \mu_E$.*

This result is immediate from the theorem and from the fact that charges have total measure 0. We obtain an obvious dual result for potential functions.

THEOREM 4. *C_E^{-1} exists if and only if $C_E l \neq 0$. If the inverse exists, then $C_E^{-1}\mathbf{1} = cl$, where $c = \alpha_E C_E^{-1}\mathbf{1}$ is a positive constant.*

PROOF. For any finite set, $\alpha_E l = \sum_{i\epsilon E} \hat{\lambda}_i^E = 1$ (by a result in the previous paper). Hence $l \neq 0$, and thus if C_E is non-singular, $C_E l \neq 0$.

Conversely, suppose that $C_E l \neq 0$. We compute $C_E(I - P^E)C_E$ twice, once from each of the last two theorems. We obtain:

$$-C_E + (C_E l)\alpha_E = -C_E + \mathbf{1}(\omega C_E).$$

If $C_E l \neq 0$, then $\mathbf{1} = cC_E l$ and $\alpha_E = c\omega C_E$, for some constant c, which is clearly positive.

Suppose that for some measure π, $\pi C_E = 0$. Then $\pi\mathbf{1} = 0$ from the above. And multiplying the result of Theorem 3 by π we find that $0 = -\pi + (\pi\mathbf{1})\omega = -\pi$. Hence $\pi = 0$, and thus C_E is non-singular. Therefore, $C_E^{-1}\mathbf{1} = cl$, and if we multiply by α_E we obtain the value of c. This completes the proof.

The dual of this result is that G_E is non-singular if and only if $\lambda^E G_E \neq 0$, and then $\alpha_E G_E^{-1} = \hat{c}\lambda^E$, where $\hat{c} = \alpha_E G_E^{-1}\mathbf{1}$.

The Corollary to Theorem 2 provides the desired generalization of Spitzer's result, if we know that the inverses exist. His result was applicable to symmetric sums of independent random variables. We see more generally:

THEOREM 5. *If P is either ergodic, or symmetric and $P_{ii}^n = P_{jj}^n$ for all i, j, n, then C_E^{-1} and G_E^{-1} exist for all finite sets E. And for any chain, if $\lambda^E > 0$, then G_E^{-1} exists, while if $\hat{\lambda}^E > 0$, then C_E^{-1} exists.*

PROOF. We shall show the results for C_E, the others are dual. If $C_E l = 0$, then C_E must have a 0 ith column whenever $\hat{\lambda}_i^E > 0$. Hence $\hat{\lambda}^E > 0$ would require $C_E = \mathbf{0}$, which contradicts Theorem 2. For an ergodic chain $C_{ij} = M_{ij}\alpha_j$ (see Theorem 24, Corollary 1), hence all off-diagonal components of C are positive. Furthermore, we know that $C_{ij}\alpha_i/\alpha_j + C_{ji} = {}^jN_{ii} \geqq 1$ if $i \neq j$ (see Theorem 22, Corollary 1). If P is symmetric and $P_{ii}^n = P_{jj}^n$, then $\alpha_i = \alpha_j$ and $C_{ij} = C_{ji}$. Hence for $i \neq j$, $C_{ij} \geqq \frac{1}{2}$. And this completes the proof.

The significance of these results is that if C_E is non-singular, then it, together with α_E, determines P^E. This is seen from the Corollary to Theorem 2 and from Theorem 4. This is a generalization of the result in [2], that the transition matrix of a finite chain is determined by M and α.

It is easy to construct examples where $C_{ij} = 0$ if $j \geqq i$, and hence where C_E is singular for all E. The class of examples in [3] has this property in all null-recurrent cases. More generally, if in a null chain M_{ij} happens to be finite, then $C_{ij} = 0$; hence random walk in one dimension with a reflecting barrier is another example.

LEMMA 2.

$$\mathrm{Lim}_n \left[\sum_{k \varepsilon E} B_{ik}^E N_{kj}^{(n)} - N_{ij}^{(n)}\right] = d_i\alpha_j - {}^EN_{ij},$$

where $d_i = {}^E\hat{\nu}_i/\alpha_i$.

PROOF. If $i \varepsilon E$, then $B_{ik}^E = \delta_{ik}$, and both sides are 0.

For $i \not\varepsilon E$ we will show this result for the reverse chain. Using that $\hat{B}_{ik}^E = \alpha_k {}^E\bar{N}_{ki}/\alpha_i$, we find that for the reverse chain our assertion is equivalent to

$$\lim_n \left[\sum_{k \varepsilon E} N_{jk}^{(n)} {}^E\bar{N}_{ki} - N_{ji}^{(n)}\right] = {}^E\nu_i - {}^EN_{ji}.$$

In this form the result can be proven by the type of systems-theorem argument we used repeatedly in our previous paper (see the explanation preceding Lemma 6). Here ${}^E\bar{N}_{ki}$ is the mean number of entries from k into i before returning to E.

THEOREM 6. *If f is a charge with support in the finite set E, then $Cf = Gf + (\lambda^E C_E f_E)\mathbf{1}$.*

PROOF. We will show first that the relation holds on E. Then we will show that $B^E(Cf) = Cf$. The result will then follow, since $B^E(Gf) = Gf$ (see Theorem

11), and $B^E\mathbf{1} = \mathbf{1}$. If we compute $G_E(I - P^E)C_E$ in two ways from Theorem 2, we find $C_E + (G_E l)\alpha_E = G_E + \mathbf{1}\cdot(\lambda^E C_E)$. We multiply this on the right by f_E, and use the fact that $\alpha_E f_E = 0$ for a charge with support in E. We obtain the desired result inside E. Since E is finite,

$$(B^E C)_{ij} = \lim_n \sum_{k\varepsilon E} B^E_{ik}\,[N^{(n)}_{jj} - N^{(n)}_{kj}] = C_{ij} - \lim_n \Big[\sum_{k\varepsilon E} B^E_{ik} N^{(n)}_{kj} - N^{(n)}_{ij}\Big].$$

Thus, by the lemma, $B^E C = C + {}^E N - d\alpha$. But ${}^E N f = 0$, since f has its support in E, and $\alpha f = 0$ for a charge, hence $B^E C f = Cf$. Which concludes the proof.

This shows that for charges with finite support C not only serves as potential operator for measures, but it "almost" serves for functions as well.

If we are dealing with a finite Markov chain, then we may choose E to be the set of all states. Then $\lambda^E = \hat{\lambda}^E = \alpha$, hence $C\mathbf{1} = (1/c)\mathbf{1}$. This says that the row-sums of C, hence of $\{M_{ij}\alpha_j\}$ (see Theorem 24) are constant. This we proved independently in [2], and there identified the sum as the trace of $Z - A = (A - I)C$. (See p. 81 and Theorem 31. There is a discrepancy of 1 due to a difference in the definition of M.)

4. Interpretation of results. Our results are more easily interpreted for ergodic chains. Here the existence of C_E^{-1} is equivalent to the existence of M_E^{-1}. Thus we see that for ergodic chains this inverse exists for every finite set of states. This is a generalization of the existence of M^{-1} for finite chains, which we obtained in [2].

Using the fact the $l_i = \hat{\lambda}_i^E/\alpha_i = \bar{M}_{iE}$, the mean time to return to set E from i (see Theorem 27), we see that the ith row-sum of C_E^{-1} is $c\bar{M}_{iE}$. We also note that c is the sum of all the components of M_E^{-1}, and hence $\hat{\lambda}^E$ is obtained by normalizing the row-sums of M_E^{-1}.

It is also worth noting that for ergodic chains Theorem 2 is equivalent to the assertion that the mean time from i in E to reach a state j in E is the mean time to return to E plus the mean time once E is reached of hitting j.

To obtain an interpretation of one more result we will specialize to strong ergodic chains.

THEOREM 7. *For a strong ergodic chain,*

$$(\alpha M)_i + M_{ij} = (\alpha\hat{M})_j + \hat{M}_{ji}.$$

PROOF. From the relation between Z and C we obtain (see Theorem 31),

$$\hat{C}_{ij} = (\alpha C)_j + \alpha_j/\alpha_i[C_{ji} - (\alpha C)_i].$$

We then replace C_{ij} by $M_{ij}\alpha_j$, and make use of the fact that αM is the same for the reverse chain as for the original. (See Theorem 24 and Lemma 12.)

This result is interesting in itself. It says that "the time to reach j in equilibrium via i" is the same as "the time to reach i in equilibrium via j" for the reverse chain. But we can also use it to clarify a previous result, $C_E l = (1/c)\mathbf{1}$.

$$(C_E l)_i = \sum_{k\varepsilon E} C_{ik}\hat{\lambda}^E_k/\alpha_k = \sum_{k\varepsilon E} M_{ik}\hat{\lambda}^E_k = \sum_{k\varepsilon E} \hat{\lambda}^E_k[(\alpha\hat{M})_k + \hat{M}_{ki} - (\alpha\hat{M})_i].$$

The last expression is entirely in terms of the reverse chain. For this chain the last two terms yield $\sum_{k\epsilon E} \lambda_k^E M_{ki} - (\alpha M)_i$. This has a simple probabilistic interpretation. The second term is the time to reach i in equilibrium, while in the first term we start in equilibrium and count the time to reach i *after entering* E. Hence the difference is $-M_{\alpha E}$, the negative of the time to reach E in equilibrium. This explains why the sum is a constant, and we obtain that $1/\hat{c} = \sum_{k\epsilon E} \lambda_k^E (\alpha M)_k - M_{\alpha E}$. Since $(\alpha M)_k \geqq M_{\alpha E}$, we see that c is positive.

REFERENCES

[1] FELLER, W., *An Introduction to Probability Theory and its Applications*, 2nd ed. John Wiley and Sons, New York, 1959.

[2] KEMENY, JOHN G. AND SNELL, J. L., *Finite Markov Chains*, D. Van Nostrand, New Jersey, 1960.

[3] KEMENY, J. G. AND SNELL, J. L., "Potentials for denumerable Markov chains," To appear in *J. of Math. Analysis and Applications.*

[4] LAMPERTI, J., "On null-recurrent Markov chains," *Canad. J. Math.* Vol. 12 (1960), pp. 278–288.

[5] SPITZER, F., "Recurrent random walk and logarithmic potential," *Proceedings of the Fourth Berkeley Symposium on Statistics and Probability,* University of California Press, Berkeley, 1961, in press.

[6] SPITZER, F., "A Tauberian theorem and its probability interpretation," *Trans. Amer. Math. Soc.*, Vol. 94 (1960), pp. 150–169.