

POTENTIALS FOR DENUMERABLE MARKOV CHAINS

BY JOHN G. KEMENY AND J. LAURIE SNELL

Communicated by J. L. Doob, September 12, 1960

Probabilistic generalizations of classical potential theory have been worked out by J. L. Doob [1] and G. A. Hunt [2]. Specialized to discrete processes, this provides a potential theory for transient Markov chains, analogous to the theory of the Newtonian potential. We provide a unified potential theory for *all* denumerable Markov chains; as applied to recurrent chains, the theory generalizes classical results on logarithmic potentials. We denote by P the transition matrix of a Markov chain with states the integers.

DEFINITION. A function of states f is called a *right potential charge* if it is integrable with respect to a given non-negative superregular measure α ($\alpha P \leq \alpha$), and if the limit $g = \lim_n (I + P + \dots + P^n)f$ exists. Then g is called a *right potential* with charge f . If μ is a finite measure and $\nu = \lim_n \mu(I + P + \dots + P^n)$ then ν is called a *left potential* with charge μ .

The mapping $f \rightarrow \mu$, with $\mu_i = f_i \alpha_i$ is an isomorphism preserving all the interesting properties. E.g., f is a right charge for P if and only if μ is a left charge for the so-called reverse chain. Since the reverse chains include all Markov chains, we automatically obtain for each theorem about functions a dual theorem about measures.

In the transient case potentials can be represented by means of a positive potential operator $G = I + P + P^2 + \dots$, as $g = Gf$ or $\nu = \mu G$.

In the recurrent case we show that $\alpha f = 0$ is a necessary condition, or dually, μ must have total measure 0.

In this case we have dual positive operators

$$G_{ij} = \lim_n [N_{ii}^{(n)} \alpha_j / \alpha_i - N_{ij}^{(n)}] \quad \text{and} \quad C_{ij} = \lim_n [N_{jj}^{(n)} - N_{ij}^{(n)}],$$

where $N_{ij}^{(n)}$ is the mean number of times that the process is in j in the first n steps, starting at i . It is shown that one exists if and only if the other does, and either is equivalent to the existence of the limiting probabilities ${}^0\lambda_j = \lim_n \sum_k P_{ik}^n {}^0h_{kj}$ for a fixed state 0. Here ${}^0h_{kj}$ is the probability, starting at k , that j is reached before 0. If any one of these conditions is fulfilled, we say that the chain is *normal*.

All ergodic (positive recurrent) chains are normal.

We know of no chain that fails to be normal, but we cannot prove that all null chains are normal. We have verified that important

special cases of null chains are. This includes random walk in one and two dimensions and sums of identically distributed independent random variables with mean 0 and finite variance.

DEFINITION. f is a weak charge if Gf and μG are both finite, where $\mu_i = f_i \alpha_i$.

THEOREM. In a normal chain, if f is a weak charge, then g exists, and $g = -Gf$ is a bounded potential.

DEFINITION. λ_i^E is the limit of the probability measure which gives the probability that after n steps the set E is entered at i .

We show that this limit always exists for a normal chain, and is independent of the starting state. However, λ^E may have total measure less than 1.

DEFINITION. If $\sum_i \lambda_i^E = 1$, then E is a small set.

THEOREM. In a normal chain a bounded function g is a potential with support in a small set E if and only if it is regular outside E and $\lambda^E g = 0$.

Analogues of such classical potential theory results as those dealing with the concepts of capacity and energy, and the principle of balayage are given. The latter takes the following form:

THEOREM. If E is a small set in a normal chain and h is a bounded function on E , then there is a unique potential with support in E that differs from h by a constant on E . The constant is $\lambda^E h$.

Energy is shown to be the average of the variance of the charge in equilibrium, with its limiting variance. (See [3, p. 86].)

Much stronger results are possible for noncyclic ergodic chains.

THEOREM. All ergodic chains are normal, and $C_{ij} = M_{ij} \alpha_j$ (M being the mean first passage time). And if $\alpha f = 0$, Gf is finite, then the potential exists and $g = -Gf$. All integrable potentials are of this form.

We also obtain a much stronger version of balayage. All sets are small in an ergodic chain, and it suffices to require that h be integrable—it need not be bounded.

For integrable functions in an ergodic chain we obtain the following simple picture. Each such function is uniquely represented as a potential plus a constant. The potentials are characterized by the fact that their integral is 0. The operator $(I - P)$ maps this space onto the space of charges, which is a proper subspace of the functions with integral 0. The inverse mapping $-G$ maps the charges onto the integrable potentials. Hence, if h is integrable, $-G[(I - P)h]$ is the

potential part of h . A matrix with the properties of $-G$ is called a *potential operator*.

It is shown that $\alpha_i M_{iE}$ (where M_{iE} is the mean time to return to the set E from its element i) is λ_i^E computed for the reverse chain. Hence Kac's theorem that $\sum_{i \in E} \alpha_i M_{iE} = 1$ is a consequence of the fact that all sets are small in an ergodic chain.

For a *strong ergodic chain* (a name we give if αM is finite), let $A = \lim_n P^n$, then we have the following even stronger result:

THEOREM. $Z = \sum_{n=0}^{\infty} (P - A)^n$ exists. $Z1 = 1$, $\alpha Z = \alpha$, and Z is a two-sided inverse of $(I - P + A)$, which may be used as the potential operator for either measures or functions. (Z is the fundamental matrix for finite ergodic chains in [3].)

Thus the analogy to the transient case is particularly good in the strong ergodic case.

A by-product of this approach is an explicit expression for first passage times. We make use of the matrix ${}^0N_{ij}$, giving the mean number of times the process is in j , starting in i , before reaching state 0. $A(k, m)$ is the number of ways k balls can be placed into m boxes, leaving none empty; and 1 is the constant function 1.

THEOREM. If f_i^k is the k th moment of the time to reach 0 from i , then

$$f^k = \sum_{m=0}^k (-1)^{k+m} A(k, m) ({}^0N)^m 1.$$

If r_n is the n th moment of the time to return to 0, then r_{n+1} and all the components of f^{n+1} exist if and only if αf^n is finite, and

$$r_{n+1} = 1/\alpha_0 \sum_{m=0}^n \binom{n+1}{m} \alpha f^m.$$

The matrix methods developed for potential theory in this paper may also be used to provide a unified approach to denumerable Markov chains.

BIBLIOGRAPHY

1. J. L. Doob, *Discrete potential and boundaries*, J. Math. Mech. vol. 8 (1959) pp. 433-458.
2. G. A. Hunt, *Markov processes and potentials*. I, II, III, Illinois J. Math. vol. 1 (1957) pp. 44-93, 316-369; vol. 2 (1958) pp. 151-213.
3. J. G. Kemeny and J. L. Snell, *Finite Markov chains*, Princeton, Van Nostrand, 1960.