ADDITIVE FUNCTIONALS AND EXCESSIVE FUNCTIONS[1]

By R. K. Getoor

Stanford Univerversity and University of Washington

0. Introduction and summary. During the past five or six years there has been much work done on the problem of representing excessive functions as the potentials of additive functionals. This is a natural generalization to Markov processes of the classical theorem of Riesz dealing with the representation of non-negative superharmonic functions as the potentials of measures. These results (due mainly to Meyer [9], Šur [12], and Volkonski [13]) are now in a fairly definitive state, and the purpose of this paper is to give a cohesive account of them. In addition to their intrinsic importance these representation theorems have been very useful in several recent applications, for example, in the theory of time changes [2] and the theory of local times [4].

Although the general theory of Markov processes is extremely rich, it is necessary to set up a rather large amount of notation and machinery before coming to grips with the problems of interest. Consequently Section 1 contains a compact summary of the definitions and basic theorems of what is now called the theory of Hunt processes, that is, Markov processes satisfying Hypothesis (A) of Hunt's fundamental memoir [8]. The author feels that such a summary is worthwhile in itself, since it can serve as a source of definitions and notations for many current research papers. The reader familiar with this material should begin with Section 2 and refer back to Section 1 only as needed. The proofs of most of the results quoted in Section 1 can be found in [1], [6], [11], or, of course, [8].

Beginning with Section 2 the exposition becomes more leisurely. We have tried not only to state definitions and theorems, but also to give some insight into them and indicate some of the more important (in our opinion) open problems. The proofs of all results quoted in Sections 2–5 (with the exception of 2.6.1(ii)) can be found in [7], as well as in the original papers cited in connection with each theorem.

1. Preliminaries on Markov processes.

1.1. *The state space.* Let E be a locally compact separable metric space with metric, d, and let Δ be a point adjoined to E as the point at infinity if E is not compact, or as an isolated point if E is compact. We write $E_\Delta = E \cup \{\Delta\}$, and we adopt the convention that any numerical (A numerical function is an extended real valued function, that is, a function taking values in the closed interval $[-\infty, \infty]$.) function f on E is extended to E_Δ by $f(\Delta) = 0$ unless explicitly stated otherwise. Let $\mathbf{C}$ denote the space of bounded (real valued) continuous functions on E, $\mathbf{C}_0$ the continuous functions on E vanishing at infinity, and $\mathbf{C}_K$

Received 13 October 1964.

[1] This paper was supported, in part, by the National Science Foundation, NSF-GP 1737.

the continuous functions on E with compact support. If $\mathbf{L}$ is any space of numerical functions on E, then $\mathbf{L}^+$ denotes the non-negative elements in $\mathbf{L}$. If f is any numerical function on E we let $\|f\| = \sup |f(x)|$. In particular $\mathbf{C}$ and $\mathbf{C}_0$ are Banach spaces under this norm.

Let $\mathcal{B}$ and $\mathcal{B}_\Delta$ denote the Borel sets of E and E_Δ respectively. If μ is a finite (non-negative) measure on $(E, \mathcal{B})$ let $\mathcal{B}^\mu$ denote the completion of $\mathcal{B}$ with respect to μ. The σ-algebra, $\mathcal{A}$, of universally measurable subsets of E is defined by $\mathcal{A} = \cap \mathcal{B}^\mu$ where the intersection is over all finite measures μ on $(E, \mathcal{B})$. A numerical function f is universally measurable (i.e. measurable with respect to $\mathcal{A}$) if and only if for each μ there exists a Borel measurable function that agrees μ almost everywhere with f. The σ-algebra $\mathcal{A}_\Delta$ is defined similarly relative to $(E_\Delta, \mathcal{B}_\Delta)$. If $(M, \mathcal{M})$ is any measurable space $\mathbf{B}(M, \mathcal{M})$ denotes the bounded real valued $\mathcal{M}$ measurable functions on M.

1.2. *Definition of a (stationary) Markov process with state space E.* Consider the following objects:

(1) A set Ω with a distinguished point ω_Δ.

(2) A σ-algebra $\mathcal{M}$ of subsets of Ω.

(3) For each $\omega \in \Omega$ a map $t \to X_t(\omega)$ from $[0, \infty]$ to E_Δ such that $X_\infty(\omega) = \Delta$, $X_t(\omega_\Delta) = \Delta$ for all $t \geqq 0$, and if $X_t(\omega) = \Delta$ then $X_s(\omega) = \Delta$ for all $s \geqq t$.

(4) For each $t \in [0, \infty]$ a map $\omega \to \theta_t(\omega)$ from Ω into Ω such that $\theta_\infty(\omega) = \omega_\Delta$ for all ω.

(5) For each x in E_Δ a probability measure P^x on $(\Omega, \mathcal{M})$.

The collection $X = \{\Omega, \mathcal{M}, P^x, X_t, \theta_t\}$ is called a (stationary) Markov process with state space E provided the following conditions are satisfied:

(M_1) Let $\mathcal{F}^0$ be the σ-algebra of subsets of Ω generated by $\{X_t \in B\}$ where $t \in [0, \infty]$ and $B \in \mathcal{B}$ (One obtains the same σ-algebra $\mathcal{F}^0$ if B is allowed to vary over $\mathcal{B}_\Delta$.) We assume that $\mathcal{F}^0 \subset \mathcal{M}$ and $x \to P^x(\Lambda)$ is $\mathcal{B}_\Delta$ measurable for each Λ in $\mathcal{F}^0$.

(M_2) $X_{t+h}(\omega) = X_t[\theta_h(\omega)]$ for $t, h \in [0, \infty]$.

(M_3) $P^x[X_0 = x] = 1$ for each x in E_Δ.

(M_4) For each $t \in [0, \infty]$ let $\mathcal{F}_t^0$ be the σ-algebra of subsets of Ω generated by $\{X_s \in B\}$, $s \leqq t$ and $B \in \mathcal{B}$, then for each $t, s \in [0, \infty]$, $x \in E_\Delta$, and $B \in \mathcal{B}$

$$P^x[X_{t+s} \in B \mid \mathcal{F}_s^0] = P^{X(s)}[X_t \in B]$$

P^x almost surely. Note that $\mathcal{F}_\infty^0 = \mathcal{F}^0$.

The function $P_t(x, B) = P^x[X_t \in B]$ defined for $t \geqq 0$, $x \in E$, and $B \in \mathcal{B}$ is called the transition function of X. Two Markov processes with the same state space E are said to be *equivalent* provided that they have the same transition function.

It follows easily from (M_2) that the *shift operators*, θ_t, are measurable in the sense that $\theta_t^{-1}\mathcal{F}^0 \subset \mathcal{F}^0$, and more precisely $\theta_t^{-1}\mathcal{F}_s^0 \subset \mathcal{F}_{t+s}^0$. The functions $t \to X_t(\omega)$ are called the *trajectories* or the *paths* of X. Sometimes it will be convenient to write $X(t, \omega)$ in place of $X_t(\omega)$. The function $\zeta(\omega) = \inf\{t: X_t(\omega) = \Delta\}$ is called the *lifetime* of the process. If F is an $\mathcal{M}$ measurable numerical function and $\Lambda \in \mathcal{M}$, $E^x\{F; \Lambda\}$ will denote $\int_\Lambda F(\omega) P^x(d\omega)$ whenever this integral exists.

1.3. *The measures P^μ.* For each finite measure μ on $(E_\Delta, \mathcal{B}_\Delta)$ we may define a measure P^μ on $(\Omega, \mathfrak{F}^0)$ by $P^\mu(\Lambda) = \int P^x(\Lambda)\mu(dx)$. We use E^μ to denote integrals with respect to P^μ. If $\mu = \epsilon_x$ (unit mass at x), then P^μ reduces to P^x. We now define $\mathfrak{F}$ to be the intersection over all μ of the P^μ-completions of $\mathfrak{F}^0$. Each of the measures P^μ extends uniquely to $\mathfrak{F}$. If $\mathcal{G}$ is any σ-algebra contained in $\mathfrak{F}$ we define the σ-algebra $\tilde{\mathcal{G}}$ as follows: $\Lambda \varepsilon \tilde{\mathcal{G}}$ if for each μ there exists Λ_μ in $\mathcal{G}$ such that $\Lambda \setminus \Lambda_\mu$ and $\Lambda_\mu \setminus \Lambda$ are in $\mathfrak{F}$ and $P^\mu(\Lambda \setminus \Lambda_\mu) = P^\mu(\Lambda_\mu \setminus \Lambda) = 0$. We next define the *definitive* σ-algebras $\mathfrak{F}_t$ of X by $\mathfrak{F}_t = \tilde{\mathfrak{F}}_t^0$ for each $t \geqq 0$. One also defines $\mathfrak{F}_{t+} = \bigcap_{s>t} \mathfrak{F}_s$ for each $t \geqq 0$. It follows that $x \to E^x(F)$ is universally measurable whenever F is in $\mathbf{B}(\Omega, \mathfrak{F})$. Moreover $\theta_t^{-1}\mathfrak{F} \subset \mathfrak{F}$ and the Markov property (M_4) extends as follows: Let F be in $\mathbf{B}(\Omega, \mathfrak{F})$, then for each $t \geqq 0$, μ, and $\Lambda \varepsilon \mathfrak{F}_t$ one has

$$E^\mu\{F \circ \theta_t ; \Lambda\} = E^\mu\{E^{X(t)}(F); \Lambda\}.$$

1.4. *Stopping times.* A mapping $T: \Omega \to [0, \infty]$ is called a *stopping time* (relative to $\{\mathfrak{F}_{t+}\}$) provided $\{T < t\} \varepsilon \mathfrak{F}_t$ for each $t \geqq 0$. The σ-algebra $\mathfrak{F}_T$ of a stoppng time consists of all $\Lambda \varepsilon \mathfrak{F}$ such that $\Lambda \cap \{T < t\} \varepsilon \mathfrak{F}_t$ for all $t \geqq 0$. Note that $\tilde{\mathfrak{F}}_T = \mathfrak{F}_T$. This notation is not yet consistent since if $T \equiv t$ then $\mathfrak{F}_T = \mathfrak{F}_{t+}$, however we will shortly impose conditions which imply that $\mathfrak{F}_{t+} = \mathfrak{F}_t$. The shift operator θ_T associated with a stopping time is defined by $\theta_T(\omega) = \theta_{T(\omega)}(\omega)$.

1.5. *Hunt processes.* Let $X = (\Omega, \mathfrak{M}, P^x, X_t, \theta_t)$ be a Markov process with state space E. We assume:

(M_5) The path functions $t \to X_t(\omega)$ are right continuous and have left-hand limits on $[0, \infty)$ almost surely. Here and in the sequel almost surely means almost surely with respect to each P^x.

Condition (M_5) implies that $\theta_T^{-1}\mathfrak{F} \subset \mathfrak{F}$ and that $\omega \to X_{T(\omega)}(\omega)$ is $\mathfrak{F} - \mathcal{B}_\Delta$ measurable for each stopping time T. It also implies that ζ is a stopping time. A Markov process X is called a *Hunt* process if in addition to (M_1) through (M_5) it satisfies:

(M_6) *Strong Markov property.* For each $F \varepsilon \mathbf{B}(\Omega, \mathfrak{F})$ and stopping time T one has

$$E^\mu\{F \circ \theta_T ; \Lambda\} = E^\mu\{E^{X(T)}(F); \Lambda\}$$

for all $\Lambda \varepsilon \mathfrak{F}_T$ and μ.

(M_7) *Quasi-left continuity.* If $\{T_n\}$ is an increasing sequence of stopping times with limit T, then $X(T_n) \to X(T)$ almost surely on $\{T < \zeta\}$.

For a Hunt process the σ-algebras $\mathfrak{F}_t$ are right continuous, that is, $\mathfrak{F}_t = \mathfrak{F}_{t+}$. Moreover one has the following characterization of the σ-algebra $\mathfrak{F}_T$ of a stopping time [5]. Let $\mathcal{G}_T$ be the σ-algebra generated by sets of the form $\{X(t \wedge T) \varepsilon B\}$ where $t \geqq 0$ and $B \varepsilon \mathcal{B}$. Of course $t \wedge s = \min(t, s)$. Then $\mathfrak{F}_T = \tilde{\mathcal{G}}_T$.

1.6. *Hitting times.* We will assume throughout the rest of the paper that X is a Hunt process. If $A \subset E_\Delta$, then $T_A(\omega) = \inf\{t > 0: X_t(\omega) \varepsilon A\}$ is called the hitting time (or entry time) of A. It is understood that if the set in braces is empty, then $T_A(\omega) = \infty$. We next state the fundamental result that T_A is a stopping time whenever A is an analytic subset of E_Δ. If A is analytic $P^x(T_A = 0)$

is either 1 or 0. In the first case x is said to be regular for A, in the second case irregular for A. If A^r denotes the regular points of A, then A^r is $\mathcal{A}_\Delta$ measurable. Note that due to the right continuity of the paths A^r contains the interior of A and is contained in the closure of A.

A set $B \subset E_\Delta$ is said to be *nearly* Borel (analytic) relative to X provided that for each μ there exist Borel (analytic) sets B_1 and B_2 such that $B_1 \subset B \subset B_2$ and $P^\mu[X_t \varepsilon B_2 \setminus B_1$ for some $t \geqq 0] = 0$. The collection of nearly Borel sets forms a σ-algebra which we will denote by $\mathcal{B}^n$. If A is nearly analytic, then T_A is a stopping time, and one can show that A^r is nearly Borel measurable. If A is nearly analytic, then for each μ there exists an increasing sequence $\{K_n\}$ of compact subsets of A such that $T_{K_n} \downarrow T_A$ almost surely P^μ. Moreover if $\mu(A \setminus A^r) = 0$, then there exists a decreasing sequence $\{G_n\}$ of open sets containing A such that $T_{G_n} \uparrow T_A$ almost surely P^μ on $\{T_A < \infty\}$.

1.7. *Existence of Hunt processes.* A transition function $P_t(x, B)$ on E is a non-negative function defined for $t \geqq 0$, $x \varepsilon E$, and $B \varepsilon \mathcal{B}$ such that:

(i) $x \to P_t(x, B)$ is $\mathcal{B}$ measurable for each t and B.

(ii) $B \to P_t(x, B)$ is a (non-negative) measure on $(E, \mathcal{B})$ for each t and x with $P_t(x, E) \leqq 1$.

(iii) The following relation holds identically

$$P_{t+s}(x, B) = \int P_t(x, dy)P_s(y, B).$$

In most applications of the present theory the basic data from which one starts is a transition function. Thus it is of fundamental importance to give conditions under which a Hunt process corresponds to a given transition function. The following is the basic result in this direction. Let $P_t(x, B)$ be a transition function on E and let $P_tf(x) = \int P_t(x, dy)f(y)$ for $f \varepsilon \mathbf{B}(E, \mathcal{B})$. If $P_t\mathbf{C}_0 \subset \mathbf{C}_0$ for each $t \geqq 0$ and $\|P_tf - f\| \to 0$[2] as $t \to 0$ for each $f \varepsilon \mathbf{C}_0$, then there exists a Hunt process X with state space E whose transition function is the given $P_t(x, B)$. In actual fact the process constructed will satisfy a stronger condition than (M_7); namely the convergence asserted in (M_7) will take place almost surely on $\{T < \infty\}$ and not merely on $\{T < \zeta\}$. In some of our earlier work a Hunt process was assumed to satisfy this stronger condition, but (M_7) suffices for the results quoted in this paper.

1.8. *The fine topology.* If A is an arbitrary subset of E_Δ, then a point x is called irregular for A provided that there exists a Borel set $B \supset A$ such that x is irregular for B, i.e., $P^x(T_B > 0) = 1$. A set A is finely open if each x in A is irregular for $A^c = E_\Delta \setminus A$. Roughly speaking a set A is finely open if a trajectory starting from a point in A remains in A for an initial interval of time with probability one. The collection of all finely open sets is a topology for E_Δ and is called the fine topology. Note that any open set is finely open. One can show that the fine topology is completely regular and that if $A \varepsilon \mathcal{B}^n$, then $A \cup A^r$ is the fine closure of A.

[2] If $P_t\mathbf{C}_0 \subset \mathbf{C}_0$ for all $t \geqq 0$, then it can be shown that $P_tf(x) \to f(x)$ as $t \to 0$ for each x in E and f in $\mathbf{C}_0$ implies that $\|P_tf - f\| \to 0$ as $t \to 0$ for each f in $\mathbf{C}_0$.

1.9. *Transition and potential operators.* We have already introduced the notation

$$P_t f(x) = \int P_t(x, dy) f(y) = E^x f(X_t).$$

More generally for any $\lambda > 0$ we define

$$P_t^\lambda f(x) = e^{-\lambda t} P_t f(x) = E^x\{e^{-\lambda t} f(X_t)\}.$$

For any $\lambda \geqq 0$ the family $\{P_t^\lambda; t \geqq 0\}$ is a semi-group of non-negative contraction operators on $\mathbf{B}(E, \mathcal{B})$ or $\mathbf{B}(E, \mathcal{A})$ with P_0^λ being the identity operator. If T is a stopping time we define

$$P_T^\lambda f(x) = E^x\{e^{-\lambda T} f(X_T)\} = E^x\{e^{-\lambda T} f(X_T); T < \zeta\}.$$

The last equality holds since by convention $f(\Delta) = 0$. If $\lambda = 0$ we write P_T in place of P_T^0, and when $T = T_A$ we write P_A^λ for $P_{T_A}^\lambda$ or just P_A when $\lambda = 0$.

Our notation for the resolvent of the semi-group is

$$U^\lambda f(x) = \int_0^\infty P_t^\lambda f(x)\, dt = E^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt.$$

Also U^λ is called the *λ-potential operator* and as above we will write U in place of U^0 when $\lambda = 0$. If $\lambda > 0$, U^λ is a non-negative operator on $\mathbf{B}(E, \mathcal{B})$ or $\mathbf{B}(E, \mathcal{A})$ of norm not exceeding $1/\lambda$. Of course, the resolvent equation, $U^\alpha - U^\beta = (\beta - \alpha) U^\alpha U^\beta$, holds for $\alpha, \beta > 0$. If $f \geqq 0$ is $\mathcal{A}$ measurable then $U^\lambda f$ is called the λ-potential of f (the potential of f when $\lambda = 0$). The operators U^λ are given by kernels, $U^\lambda(x, dy)$, called the *potential kernels* of X, i.e. $U^\lambda f(x) = \int U^\lambda(x, dy) f(y)$.

1.10. *Excessive functions.* A non-negative $\mathcal{A}$ measurable function f is *λ-excessive* provided $P_t^\lambda f \leqq f$ for all $t \geqq 0$ and $P_t^\lambda f \to f$ as $t \to 0$ pointwise. As usual when $\lambda = 0$ we say that f is excessive. If $g \geqq 0$ is $\mathcal{A}$ measurable, then the λ-potential of g, $U^\lambda g$, is λ-excessive for any $\lambda \geqq 0$. Any statement involving λ will be understood to hold for all $\lambda \geqq 0$ unless explicitly stated otherwise. We will use $\mathbf{E}^\lambda$ to denote the λ-excessive functions ($\mathbf{E}$ when $\lambda = 0$). If $f \in \mathbf{E}^\lambda$, then so is $P_A^\lambda f$ for any nearly analytic set A. Excessive functions may be characterized in terms of the resolvent $\{U^\lambda\}$ as follows: $f \in \mathbf{E}^\lambda$ if and only if (i) $\alpha U^{\lambda+\alpha} f \leqq f$ for all $\alpha > 0$ and (ii) $\alpha U^{\lambda+\alpha} f \to f$ pointwise as $\alpha \to \infty$.

The following is the main result on excessive functions [8].

1.10.1. THEOREM. *Let f be λ-excessive, then*

(i) *f is nearly Borel measurable,*

(ii) *the function $t \to f(X_t)$ is right continuous, and it is finite for all $t \geqq s$ if $f(X_s)$ is finite, both statements holding almost surely,*

(iii) *if μ is a probability measure on $(E, \mathcal{A})$ such that $\int f\, d\mu < \infty$ then $\{e^{-\lambda t} f(X_t), \mathfrak{F}_t, P^\mu\}$ is a supermartingale and $t \to f(X_t)$ has left hand limits on $[0, \infty)$ almost surely.*

A consequence of this theorem is the fact that if f and g are in $\mathbf{E}^\lambda$ so is $\min(f, g)$. Also any f in $\mathbf{E}^\lambda$ is finely continuous. In fact for any $\lambda > 0$ the fine topology is the coarsest topology relative to which the elements of $\mathbf{E}^\lambda$ are continuous.

1.11. *Exceptional sets.* A set A is *approximately null* if it is contained in a universally measurable set B such that $U^\lambda(x, B) = U^\lambda I_B(x)$ vanishes identically in x for one, and hence, all, $\lambda \geqq 0$. Here I_B is the indicator function of B. A set A is *thin* if it is contained in a nearly analytic set B for which no point is regular, i.e., B^r is empty. A set is *semi-polar* if it can be covered by a countable union of thin sets. A set A is *polar* provided it is contained in a nearly analytic set B such that $P^x(T_B < \infty) = 0$ for all x. Thus, roughly speaking, a polar set is a set that the process never enters at positive times almost surely. The following implications hold,

$$\text{polar} \Rightarrow \text{semi-polar} \Rightarrow \text{approximately null}.$$

The converse implications are not valid in general.

Two λ-excessive functions which agree except on an approximately null set are identical. In particular the complement of an approximately null set is finely dense. If a λ-excessive function f is finite except on an approximately null set, then $\{f = \infty\}$ is actually polar. If A is nearly analytic, then $A \setminus A^r$ is semi-polar. If B is semi-polar then X_t is in B for at most countably many values of t almost surely.

2. Additive functionals and their potentials. Throughout the remainder of this paper $X = (\Omega, \mathfrak{M}, P^x, X_t, \theta_t)$ is a fixed Hunt process with state space E.

2.1. *Motivation.* Let f be a non-negative bounded Borel measurable function on E and consider

$$A_t(\omega) = A(t, \omega) = \int_0^t f[X_s(\omega)]\, ds.$$

Clearly $t \to A_t$ is continuous, non-decreasing, and $A_0 = 0$. Moreover A_t is $\mathfrak{F}_t$ measurable. Also we may calculate A_{t+s} as follows

$$\begin{aligned} A(t + s, \omega) &= \int_0^t f[X_u(\omega)]\, du + \int_t^{t+s} f[X_u(\omega)]\, du \\ &= A(t, \omega) + \int_0^s f[X_{u+t}(\omega)]\, du \\ &= A(t, \omega) + A(s, \theta_t\omega), \end{aligned}$$

since $X_{u+t} = X_u \circ \theta_t$. The potential of f is easily expressed in terms of A,

$$Uf(x) = E^x \int_0^\infty f(X_t)\, dt = E^x A(\infty),$$

where $A(\infty) = \lim_{t\uparrow\infty} A(t)$, and more generally the λ-potential of f is given by

$$\begin{aligned} U^\lambda f(x) &= E^x \int_0^\infty e^{-\lambda t} f(X_t)\, dt \\ &= E^x \int_0^\infty e^{-\lambda t}\, dA(t). \end{aligned}$$

Notice finally that $A(t)$ is constant on $[\zeta, \infty]$ since $X_t = \Delta$ on this interval and $f(\Delta) = 0$, that is, $A(\infty) = \int_0^\zeta f(X_t)\, dt$.

In the classical potential theory of three dimensional Euclidean space (i.e. three dimensional Brownian motion) the potential kernel $U(x, dy)$ has the form $U(x, dy) = u(x, y)\, dy$ where dy is three dimensional Lebesgue measure and $u(x, y)$ is the Newtonian kernel, $c|x - y|^{-1}$. Thus in addition to the potential of a function one can define the potential of a (non-negative) measure μ by $U\mu(x) = \int u(x, y)\mu(dy)$. In fact if one wishes to represent non-negative superharmonic functions as a harmonic part plus a potential (Riesz Theorem) it is necessary to use the potential of a measure, the potential of a function will not suffice. For a general Hunt process the excessive functions are analogous to the non-negative superharmonic functions for Brownian motion. Thus it is natural to try and represent excessive functions (or, at least, certain large classes of excessive functions) as "potentials."

This immediately raises the question "potentials of what?". As we have seen even in the classical case potentials of functions do not suffice, and for a general Hunt process X the potential kernel $U(x, dy)$ can not be used to define the potential of a measure as a function. We will see that a reasonable answer to this question is furnished by the "potentials" of additive functionals.

2.2. *Definitions.* A family $\{A_t ; t \geqq 0\}$ of numerical random variables on $(\Omega, \mathfrak{F})$ is called a (non-negative, right continuous) *additive functional* of X provided:

(i) The following statements hold almost surely; $t \to A_t(\omega)$ is right continuous, non-decreasing, $A_0 = 0$, and $A_t(\omega) = \lim_{s\uparrow\zeta(\omega)} A_s(\omega)$ for all $t \geqq \zeta(\omega)$.

(ii) A_t is $\mathfrak{F}_t$ measurable.

(iii) For each $t, s \geqq 0$ one has $A_{t+s} = A_t + A_s \circ \theta_t$ almost surely.

We define $A_\infty(\omega) = \lim_{t\uparrow\infty} A_t(\omega)$. This exists almost surely and $A_\infty = A_\zeta = A_{\zeta-}$ almost surely. We will sometimes write $A(t)$ for A_t and $A(t, \omega)$ for $A_t(\omega)$. The family of random variables defined in Section (2.1) is an example of a (continuous) additive functional. We emphasize that the exceptional set in (iii) depends, in general, on t and s. If it can be chosen independent of t and s the additive functional is called *perfect.*

A *continuous* additive functional (CAF) is an additive functional such that $t \to A_t$ is almost surely continuous. An additive functional is called *natural* (NAF) provided $t \to A_t$ and $t \to X_t$ have no common discontinuities almost surely. Note that any CAF is natural, while if X has continuous paths every additive functional is natural. The importance of the concept of NAF will appear shortly. Finally two additive functionals $A = \{A_t\}$ and $B = \{B_t\}$ are said to be *equivalent* provided $P^x[A_t \neq B_t] = 0$ for all x and t. In view of the right continuity of A and B this is equivalent to the statement that the functions $t \to A_t$ and $t \to B_t$ are identical almost surely.

2.3. *The strong Markov property.* The following result of Meyer [9] is of basic importance for the present theory.

2.3.1. THEOREM. *Let A be an additive functional, then if T is a stopping time and R any non-negative random variable on $(\Omega, \mathfrak{F})$ one has*

$$A[T(\omega) + R(\omega), \omega] = A[T(\omega), \omega] + A[R(\omega), \theta_T\omega]$$

almost surely.

In the future we will omit the ω's when writing such a relationship. For example the above equation would be written $A(T + R) = A(T) + A(R, \theta_T)$.

2.4. *Potentials of additive functionals.* Let $A = \{A_t\}$ be an additive functional of X. If f is a non-negative nearly Borel measurable function on E, we define the λ-potential of f with respect to A by

$$\begin{aligned} U_A{}^\lambda f(x) &= E^x \int_0^\infty e^{-\lambda t} f(X_t)\, dA(t) \\ &= E^x \int_0^\zeta e^{-\lambda t} f(X_t)\, dA(t). \end{aligned}$$

In particular if $f \equiv 1$ we write $U_A{}^\lambda$ in place of $U_A{}^\lambda 1$, and the function $U_A{}^\lambda$ is called the λ-potential of A. As usual when $\lambda = 0$ we drop it entirely from our notation. One can show that the definition given above makes sense; in fact $\int_0^\infty e^{-\lambda t} f(X_t)\, dA_t$ is $\mathfrak{F}$ measurable and so $U_A{}^\lambda f$ is $\mathfrak{A}$ measurable. If $U_A{}^\lambda$ is finite, then $f \to U_A{}^\lambda f$ is given by kernel which we denote by $U_A{}^\lambda(x, dy)$, that is

$$U_A{}^\lambda f(x) = \int U_A{}^\lambda(x, dy) f(y).$$

If $f \geqq 0$ is nearly Borel measurable then $U_A{}^\lambda f$ is λ-excessive and hence nearly Borel measurable. The operators $f \to U_A{}^\lambda f$ are not in general a resolvent. However the following equation is analogous to the resolvent equation. If $\alpha, \beta \geqq 0$ and $U_A{}^\alpha f$ and $U_A{}^\beta f$ are finite, then

$$U_A{}^\alpha f - U_A{}^\beta f = (\beta - \alpha) U^\beta U_A{}^\alpha f = (\beta - \alpha) U^\alpha U_A{}^\beta f.$$

2.5. *Uniqueness theorems.* The results of this subsection are due to Meyer [9]. They give conditions under which the potential of an additive functional determines the additive functional.

2.5.1. THEOREM. *Let A and B be additive functionals with finite λ-potentials for some fixed $\lambda \geqq 0$. If $U_A{}^\lambda f = U_B{}^\lambda f$ for all f in $\mathbf{C}_K{}^+$; then A and B are equivalent.*

2.5.2. THEOREM. *Let A and B be natural additive functionals with finite λ-potentials for some fixed $\lambda \geqq 0$. If $U_A{}^\lambda = U_B{}^\lambda$, then A and B are equivalent.*

It is easy to see that Theorem 2.5.2 is *not* valid if the word "natural" is omitted from its statement. For example, take X to be the Poisson process with parameter 1 and let $A(t) = t$ and $B(t)$ be the number of jumps of the sample path in the interval $[0, t]$. A simple computation shows that $U_A{}^\lambda = U_B{}^\lambda = \lambda^{-1}$ for all $\lambda > 0$.

2.6. *A characterization of* CAF's *and* NAF's. Let A be an additive functional with a finite λ-potential for some $\lambda \geqq 0$. If A is natural and G is an open subset of E, then $P_G{}^\lambda U_A{}^\lambda f = U_A{}^\lambda f$ whenever f is a nearly Borel measurable function vanishing outside of G. If A is continuous one may replace the open set G by an arbitrary nearly analytic set D in the above statement. These statements are almost obvious. However we also have the following converses.

2.6.1. THEOREM. *Let A be an additive functional with a finite λ-potential for some fixed $\lambda \geqq 0$.*

(i) *If for each f in $\mathbf{C}_K{}^+$ and open neighborhood G of the support of f one has $P_G{}^\lambda U_A{}^\lambda f = U_A{}^\lambda f$, then A is natural.*

(ii) *If $P_K{}^\lambda U_A{}^\lambda I_K = U_A{}^\lambda I_K$ for all compact K where I_K is the indicator function of K, then A is continuous.*

Statement (i) is due to Meyer [9]. Meyer also proved (ii) under an additional hypothesis (Hyp (L) of Section 5). However, (ii) can be established in general by an argument which is completely different from the one in [9].

It is easy to see that for any additive functional with finite λ-potential

$$U_A^\lambda I_K(x) - P_K^\lambda U_A^\lambda I_K(x) = E^x\{e^{-\lambda T_K}[A(T_K) - A(T_K-)]\}.$$

Thus Theorem 2.6.1(ii) states that if A is not continuous, then there exist an x and a compact set K such that $t \to A_t$ is discontinuous at T_K with positive P^x probability.

3. Representation theorems. In the remainder of this paper we will state results only in the case $\lambda = 0$. However, all of these results remain valid for arbitrary λ with obvious modifications.

3.1. *Motivation.* Suppose A is a CAF with a finite potential. If $\{T_n\}$ is an increasing sequence of stopping times with limit T, then using the strong Markov property we may compute

$$\begin{aligned} E^x U_A(X_{T_n}) &= E^x\{E^{X(T_n)}A(\infty)\} \\ &= E^x\{A(\infty, \theta_{T_n})\} \\ &= E^x\{A(\infty) - A(T_n)\} \\ &\to E^x\{A(\infty) - A(T)\} \\ &= E^x U_A(X_T). \end{aligned}$$

Note that if A is not assumed to be continuous but $T = \zeta$ almost surely P^x, then the same computation shows that

$$E^x U_A(X_{T_n}) \to E^x U_A(X_\zeta) = 0,$$

since $U_A(\Delta) = 0$.

3.2. *The representation theorems.* A finite excessive function f is called a *regular potential* provided $E^x f(X_{T_n}) \to E^x f(X_T)$ for all x whenever $\{T_n\}$ is an increasing sequence of stopping times and $T = \lim T_n$. We saw in Section 3.1 that the finite potential of a CAF is a regular potential. The next result is due to Šur [12] for bounded f. His proof contained a slight gap which was filled in [2]. The extension to finite f is standard. See [7]. The general result was also obtained by Meyer [9] but under an additional hypothesis (Hyp(L) of Section 5).

3.2.1. THEOREM. *Let f be a finite regular potential, then there exists a unique* (*up to equivalence*) *continuous additive functional A such that $f = U_A$.*

A finite excessive function f is called a *potential of class* (D) if for each x whenever $\{T_n\}$ is a sequence of stopping times that increases to ζ almost surely P^x, then $E^x f(X_{T_n}) \to 0$ as $n \to \infty$. It follows from the comments in Section 3.1 that if A is an (arbitrary) additive functional with finite potential then U_A is a potential of class (D). The next result is due to Meyer [9].

3.2.2. THEOREM. *Let f be a finite potential of class* (D), *then there exists a unique* (*up to equivalence*) *natural additive functional A such that $f = U_A$.*

This result implies that there is a mapping $A \to A^*$ from the set of all additive functionals with finite potential onto the set of all *natural* additive functionals with finite potential such that $U_A = U_{A^*}$, and $A^* = B^*$ if and only if $U_A = U_B$.

Here and in the sequel equality between additive functionals means equivalence. Roughly speaking one obtains A^* from A by the following construction. One takes all the jumps of $t \to A_t$ which occur at times when $t \to X_t$ is continuous and lumps them together to form a "pure jump" natural additive functional A^j. It is then possible to show that $U_A - U_{A^j}$ is a regular potential, and hence the potential of a CAF, A^c. Consequently $A^* = A^j + A^c$. Thus among all additive functionals which have the same finite potential, f, the unique natural one, A^*, has the least number of discontinuities. In fact one can show that the jumps of A^* are precisely those jumps of $t \to f(X_t)$ which occur at times when $t \to X_t$ is continuous.

3.3. *Classification of excessive functions.* An excessive function, f, is *regular* provided $t \to f(X_t)$ is continuous wherever $t \to X_t$ is continuous almost surely. Thus the only possible discontinuities of $t \to f(X_t)$ when f is a regular excessive function are the discontinuities of the path function $t \to X_t$ almost surely. It is implicit in the discussion at the end of Section 3.2 that a regular potential is regular. An excessive function f is of class (D) provided the family $\{f(X_T); T$ a stopping time$\}$ is uniformly integrable with respect to P^x for each x. One can show that a potential of class (D), and hence a regular potential, is of class (D), so that the terminology is consistent.

Thus we are led to consider the following two conditions on X. (The above discussion shows that these conditions are closely related, and, in fact, it is not difficult to see that they are equivalent.)

(K) Every (finite) excessive function is regular.

(M) Every NAF (with finite potential) is continuous.

Hunt has shown [8] that, under the set up in Part III of [8], (K), and hence (M), is equivalent to the following condition.

(H) Every semi-polar set is polar.

In the set up of Hunt's Part III there is a positive Radon measure ξ on E such that the potential kernel $U(x, dy)$ has the form $U(x, dy) = u(x, y)\xi(dy)$, and Hunt showed that under these assumptions a sufficient condition that (H), and hence (K) and (M), hold is that $u(x, y)$ be symmetric in x and y. Thus (H), (K), and (M) hold for Brownian motion, the symmetric stable processes, and many other familiar processes.

The typical example in which (H) fails to hold is translation to the right along the real line with uniform velocity. In this example a set consisting of a single point is thin, but not polar. More generally the space-time Brownian motion process, i.e., the process whose generator is the heat operator, does not satisfy (H). In a certain rough sense (H) should correspond to the "ellipticity" of the generator of X. Thus it should hold for processes corresponding to Brelot's axiomatic potential theory, but not for those corresponding to Bauer's axiomatic theory. We should remark that the two processes mentioned above, translation and space-time Brownian motion, do satisfy the assumptions of Hunt's Part III.

It is tempting to conjecture that the equivalence of (H), (K), and (M) holds for general Hunt processes. However, this is not the case as the following example shows. Let E be the closed left half-plane, $E = \{(x, y); x \leqq 0\}$, x and y real. Let X be two dimensional Brownian motion until it first hits the y-axis at which time it becomes one dimensional Brownian motion along the y-axis. This process has continuous paths and satisfies (H). It also satisfies Hypotheses (B) and (C) of [8]. However, it doesn't satisfy (K). Let f be one in the open left half-plane and zero on the y-axis, then f is excessive but $t \to f(X_t)$ is discontinuous at the time X first reaches the y-axis. It would be very interesting to formulate more general conditions than those of Hunt's Part III under which (H) and (K) are equivalent.

In general one can decompose a NAF into a CAF and a "pure jump" NAF, and under the assumptions of Hunt's Part III one has the following very nice description of the "pure jump" component. Let A be a "pure jump" NAF, then there exists a semi-polar set Q and a function $\varphi \geqq 0$ on E such that

$$A(t) = \sum_{T_\alpha \leqq t} \varphi(X_{T_\alpha})$$

where the T_α (α a countable ordinal) are the successive times at which the process X hits Q. Again it would be of interest to formulate more general conditions under which this representation is valid. The example of the previous paragraph shows that it fails to hold for Hunt processes in general. (If T is the hitting time of the y-axis for the process in the previous paragraph, then $A(t) = 0$ for $t < T$ and $A(t) = 1$ for $t \geqq T$ if the process starts in the open left half-plane and $A(t) \equiv 0$ if the process starts on the y-axis is a counterexample.)

4. The fine support of a CAF. Throughout this section A will be an additive functional with finite potential. The results of this section are valid if one only assumes that A has a finite λ-potential for some fixed $\lambda \geqq 0$, but we will assume $\lambda = 0$ for simplicity. We say that A *vanishes* on a nearly Borel set D provided $U_A(x, D) = U_A I_D(x) = 0$ for all x. We can then define the *support* of A as the smallest closed set on whole complement A vanishes. A standard argument yields the existence and uniqueness of the support of A. However, the support of A is a very crude notion and it is natural to try and define the *fine support* of A as the smallest finely closed set on whose complement A vanishes. But the fine topology is not locally compact in general and so one can not establish the existence of the fine support of A in the usual manner. Nevertheless one can establish the existence of the fine support of a *continuous* additive functional by a direct argument.

In the remainder of this section A will be a CAF with finite potential. Define

$$R = \inf\{t: A_t > 0\} =' \sup\{t: A_t = 0\},$$

and note that $A_R = 0$ almost surely on $\{R < \infty\}$ since A is continuous. It is easy to see that R is a stopping time and that $\varphi(x) = E^x(e^{-R})$ is 1-excessive. We now define

$$F = \{x: P^x(R = 0) = 1\} = \{x: \varphi(x) = 1\}.$$

Intuitively, x is in F if and only if $t \to A_t$ begins to increase immediately with probability one when the process starts at x. One can now establish, [7], the following properties of F:

(i) F is a finely closed nearly Borel set.

(ii) F is empty $\Leftrightarrow A_t \equiv 0$ a.s. $\Leftrightarrow R = \infty$ a.s.

(iii) $T_F = R$ a.s.

(iv) Each x in F is regular for F.

If A is not continuous, F may be empty even though A is not identically zero.

Next define the following sets, each of which depends on ω.

$$I = \{t: A(t + \epsilon) - A(t) > 0 \quad \text{for all} \quad \epsilon > 0\}$$

$$\bar{I} = \{t: A(t + \epsilon) - A(t - \epsilon) > 0 \quad \text{for all} \quad \epsilon > 0\}$$

$$Z = \{t: X_t \varepsilon F\}.$$

Clearly $\bar{I}$ is the closure of I almost surely. I is the set of points of *right increase* of A, while $\bar{I}$ is the set of points of *increase* of A. Moreover $\bar{I}(\omega)$ is the support of the measure $dA(t, \omega)$ on $[0, \infty)$, and since $\bar{I} \setminus I$ is almost surely countable, the continuity of A implies that $dA(t)$ is carried by I almost surely. The following theorem is the main result of this development. See [7].

4.1. THEOREM. *Almost surely* $I \subset Z \subset \bar{I}$.

A corollary of this is the fact that F is the fine support of A. Also A and $I_F A$ are equivalent. Here $I_F A$ is the additive functional $t \to \int_0^t I_F(X_s)\, dA_s$. (More generally if f is any non-negative nearly Borel function we will write fA for the family of random variables $t \to \int_0^t f(X_s)\, dA_s$. Of course fA is an additive functional under some finiteness assumption, for example, if f is bounded and A has a finite potential, then fA is an AF with a finite potential, $U_{fA} = U_A f$.) Two final remarks:

(i) A is strictly increasing if and only if $E = F$, and

(ii) the closure of F is the support of A.

5. An analog of the Radon-Nikodym theorem. Again in this section A will be an AF with finite potential, although the results are valid with obvious modifications if one only assumes that A has a finite λ-potential. The following hypothesis will be assumed throughout this section.

(L) There exists a (non-negative) Radon measure ξ on E such that an excessive function which vanishes a.e. (ξ) is identically zero.

Note that (L) is certainly satisfied if excessive functions are lower semicontinuous. In particular (L) holds in the set up of Hunt's Part III.

The following is the key result in this section and it is due to Motoo [10].

5.1. THEOREM. *Let A be a* CAF *with a finite potential and assume* (L) *holds. Then an excessive function f has the representation $f = U_A h$ with h a Borel measurable function satisfying $0 \leqq h \leqq 1$ if and only if $f \leqq U_A$ and*

$$f(x) - P_t f(x) \leqq E^x A(t)$$

for all x and t.

Let **A** be the class of all *continuous* additive functionals with finite potential. If A and B are in **A** an $\alpha \geqq 0$, define $A + B$ and αA in the obvious manner.

If A and B are any two additive functionals we write $A \leqq B$ provided there exists an additive functional C such that $A + C = B$. It is immediate that "$\leqq$" is a partial ordering and that if B is in **A** and $A \leqq B$, then A is in **A**. We now consider **A** with this partial ordering.

5.2. THEOREM.

(i) **A** *is a boundedly complete lattice.*

(ii) *The following three statements are equivalent*:

(a) $A \leqq B$

(b) $E^x A(t) \leqq E^x B(t)$ *for all* t, x

(c) $A = fB$ *for a Borel measurable* f *with* $0 \leqq f \leqq 1$.

Clearly **A** is a convex cone and it is easy to see that "local times" at points (when they exist) determine extremal rays in **A**. Most likely these are the only extremal rays.

The following result is what we call the analog of the Radon-Nikodym theorem.

5.3. THEOREM. *If* A *and* B *are in* **A**, *then a necessary and sufficient condition that* $A = fB$ *with* f *a non-negative Borel measurable function is that for each* x *the measure* $U_A(x, \cdot)$ *be absolutely continuous with respect to* $U_B(x, \cdot)$.

It would seem that the non-trivial part of Theorem 5.3 could be proved in the following manner. By the Radon-Nikodym theorem, for each x there exists a function f_x such that $U_A(x, dy) = f_x(y) U_B(x, dy)$. One should then be able to show that f_x is independent of x, and consequently it would follow that the resulting function f does the required job. However, it seems to be difficult to carry out this argument. In fact, Theorem 5.3 and Theorem 5.1 are *not* valid for natural additive functionals as one can see from simple examples. See [7], p. 63. The proofs of Theorems 5.2 and 5.3 are given in [7], although they are relatively simple consequences of Theorem 5.1.

There are two interesting questions that one can ask about the results of this section. First of all, (L) is not needed in the statement of these results and so it would be interesting to know if they are valid without assuming (L). Secondly, and perhaps of more importance, Theorems 5.1 and 5.3 are valid for natural additive functionals under the assumptions of Hunt's Part III, and hence it would be very interesting to formulate the "most general" conditions under which they are valid for natural additive functionals.

6. Potentials of measures. In this section we will mention briefly the connection between additive functionals and measures under the assumptions of Hunt's Part III. We will not go into detail on these assumptions. The interested reader may consult [3] or [9], Part II, Section 6, for a summary or, of course, [8], Part III. The main thing to remember is that the potential kernel $U(x, dy)$ is of the form $u(x, y)\xi(dy)$, and that most familiar processes (Brownian motion, processes with independent increments, etc.) satisfy these assumptions. We are assuming that we can take $\lambda = 0$ as usual.

Since the potential kernel is given by a point function one can define the potential of a measure $\mu \geqq 0$ by

$$U\mu(x) = \int u(x, y)\mu(dy).$$

If $U\mu$ is finite, then it is excessive and determines μ. The following result of Meyer [9] establishes the relationship between μ and the concepts developed above.

6.1. THEOREM. *Let μ be a non-negative measure with $U\mu$ finite, then*

(i) *$U\mu$ is a potential of class* (D) *if and only if μ charges no polar set,*

(ii) *$U\mu$ is a regular potential if and only if μ charges no semi-polar set.*

The next result is also due to Meyer [8], but a much simpler proof is given in [3].

6.2. THEOREM. *Let A be a* NAF *with a finite potential and μ a non-negative measure such that $U_A = U\mu$, then $U_A(x, dy) = u(x, y)\mu(dy)$.*

Finally let us mention that it follows from results of Hunt that if f is a potential of class (D) then $f = U\mu$ for an appropriate measure μ. Hence to each NAF, A, with finite potential there corresponds a unique measure μ such that $U_A = U\mu$. Theorem 6.2 states that if A corresponds to μ, then fA corresponds to $f\mu$.

REFERENCES

[1] BAUER, H. (1963). Markoffsche Prozesse. Vorlesung an der Universität Hamburg.

[2] BLUMENTHAL, R. M., GETOOR, R. K. and MCKEAN, H. P., JR. (1962). Markov processes with identical hitting distributions. *Illinois J. Math.* **6** 402–420.

[3] BLUMENTHAL, R. M. and GETOOR, R. K. (1964). Additive functionals of Markov processes in duality. *Trans. Amer. Math. Soc.* **112** 131–163.

[4] BLUMENTHAL, R. M. and GETOOR, R. K. (1964). Local times for Markov processes. *Z. Wahrsheinlichkeitstheorie und Verw. Gebiete.* **3** 50–74.

[5] BLUMENTHAL, R. M. and GETOOR, R. K. (1964). A theorem on stopping times. *Ann. Math. Statist.* **35** 1348–1350.

[6] DYNKIN, E. B. (1961). *Theory of Markov Processes* (English translation). Prentice-Hall, Englewood Cliffs, N. J.

[7] GETOOR, R. K. (1964). Additive functionals of a Markov process. Lecture notes. University of Hamburg.

[8] HUNT, G. A. (1957), (1958). Markoff processes and potentials. *Illinois J. Math.* **1** 42–93, 316–369, and **2** 151–213.

[9] MEYER, P. A. (1962). Fonctionelles multiplicatives et additives de Markov. *Ann. Inst. Fourier, Grenoble.* **12** 125–230.

[10] MOTOO, M. (1962). Representations of a certain class of excessive functions and a generator of Markov process. *Sci. Papers Coll. Gen. Ed. Univ. Tokyo.* **12** 143–159.

[11] Séminaire de théorie du potentiel (Séminaire Brelot-Choquet-Deny), Inst. Henri Poincaré, Paris, Fifth year, 1960–61.

[12] ŠUR, M. G. (1961). Continuous additive functionals of Markov processes and excessive functions. *Dokl. Akad. Nauk SSSR* **137** 800–803 (in Russian). Translated in *Soviet Mathematics* (1961) **2** 365–368.

[13] VOLKONSKI, V. A. (1960). Additive functionals of a Markov process. *Trudy Moskov. Mat. Obšč.* **9** 143–189 (in Russian).