## ESSAY IV. APPLICATION OF THE PREDICTION PROCESS TO MARTINGALES

### 0. INTRODUCTION.

Let $X(t)$, $t \geq 0$, be a right-continuous supermartingale relative to an increasing family of $\sigma$-fields $G_t^*$ on some probability space $(\Omega^*, F^*, P^*)$. We assume that the $G_t^*$ are countably generated for each $t$. It is then easy, by using indicator functions of generators of $G_t^*$, to construct a sequence $X_{2(n+1)}(t)$, $1 \leq n$, of real-valued processes such that $\{X(s), (X_{2(n+1)}(s)), s \leq t\}$ generates $G_t^*$ for each rational $t$. We can now transfer both process and probability to the canonical space $\Omega$ of Essay 1. We simply set $P\{w_{2n-1}(s) = 0,$ all $s \geq 0$ and $n \geq 1\} = 1$, and for $S \in X_{t \geq 0} \bar{B}_\infty$ (see Essay 1, Section 1 for notation)

$$P\{(w_{2n}(\cdot)) \in S\} = P^*\{(X(\cdot), X_{2(n+1)}(\cdot)) \in S\} .$$

Then we obtain a canonically defined process $X_t((w_n)) = w_2(t)$ which is a supermartingale with respect to $P$ and the $\sigma$-fields $G_t^o$ of Essay 1. In the present work, we let $X_t$ denote this process (rather than the sequential process $(w_{2n}))$, and we drop the odd coordinates from the notation (i.e., we discard the set of probability 0 where they are non-zero). Thus we do not allow any "hidden information": $F_t^o = G_t^o$. By a well-known convergence theorem we have

$$\begin{aligned} E(X_{s+t} | F_{t+}^o) &= \lim_{r \to t+} E(X_{s+t} | F_r^o) \\ &\leq \lim_{r \to t+} X_r \\ &= X_t . \end{aligned}$$

Hence $X_t$ is a supermartingale relative to $F_{t+}^o$, and we can connect it with its prediction process $Z_t^P$.

As in Essay 1, the method requires that $P$ be treated as a variable. In the present work we are concerned initially with three familiar classes of $P$ on $(\Omega, F^o)$, as follows.

DEFINITION 0.1. Let $\mathcal{M} = \{P: X_t$ is an $\mathcal{F}^o_{t+}$-martingale and $\sup E\, X^2(t) < \infty\}$, $\mathcal{U} = \{P: X_t$ is a uniformly integrable martingale, i.e., $X_t = E(X_\infty | \mathcal{F}^o_{t+}\}$, and $\mathcal{D} = \{P: X_t$ is a non-negative supermartingale of class D with $\lim_{t\to\infty} EX_t = 0\}$. The classes $\mathcal{M}$ and $\mathcal{D}$ are called respectively the square-integrable martingales and the potentials of class D, or simply the potentials (see [4, VI, Part 1, 9]).

Of course, we have $\mathcal{M} \subset \mathcal{U}$, and most of the attention will be on $\mathcal{M}$ and $\mathcal{D}$. For $P \in \mathcal{M}$ we have a decomposition

$$X_t - X_0 = X_1^P(t) + X_2^P(t) \ , \tag{0.1}$$

where $X_1^P$ is a continuous $\mathcal{F}^o_{t+}$-martingale and $X_2^P$ is a "compensated sum of jumps" with $E(X_1^P X_2^P) = 0$. This decomposition is due to P. A. Meyer [11], but it will be obtained here as a consequence of a result on additive functionals of a Markov process (Theorem 1.6), more in the spirit of H. Kunita and S. Watanabe [10]. Given such a decomposition (for fixed P) it is clear that $Z_t^P$ contains the distributions of both processes $X_i^P(s) \circ \theta_t$ given $\mathcal{F}^o_{t+}$, but this approach is not useful because one does <u>not</u> have $X_i^P(s+t) - X_i^P(t) = X_i^P(s) \circ \theta_t$. Rather, one has at least in principle

$$X_i^P(s+t) - X_i^P(t) = X_i^{Z_t^P}(s) \circ \theta_t \ ,$$

so that the $X_i(t)$ become additive functionals of the prediction process.

To make this approach rigorous, it is very convenient (and probably necessary) to transfer the setting once more to the prediction spaces of Essay 1, Section 2. Here the $Z_t^P$ are given a single definition not depending on P, and for example the above $X_i(t)$ become actual additive functionals of $Z_t$. In the setting of $\mathcal{D}$, this enables us to avoid the technical difficulties encountered in [7] with a similar question.

This approach permits the application of general Markovian methods to the analysis of the $X_i(t)$, and to other decompositions in $\mathcal{U}$ and $\mathcal{D}$. In particular, we obtain the celebrated Doob-Meyer decomposition in $\mathcal{D}$ as a theorem on Markov additive functionals (Theorem 1.8). Further investigation of the discontinuities is based on the theory of Lévy systems ([1]). Thanks to the use of a suitably weak topology for $Z_t$, it is possible to transfer directly the known components of the Lévy system of a Ray process to $Z_t$, including separate terms for the compensation of totally inaccessible jumps and previsible jumps. Rather surprisingly, this operation is in no way restricted to martingales. By returning the components to the original probability space $(\Omega, F^o)$, we obtain (what is termed) the Levy system of an arbitrary r.c.l.l. process (Definition 2.2, Theorem 2.3).

Treatment of the continuous components, unlike that of the jumps, is restricted to the case of martingales. The continuous local martingales comprise a single prediction process (a "packet," as in Essay 1). By the means of a time change inverse to an additive functional, they are all reduced to a single Brownian motion (but it is a Brownian motion for many different probabilities). We then specialize to the case of autonomous germ-Markov probabilities, which generalizes the one-dimensional diffusion processes in the natural scale on $(-\infty,\infty)$. Even in this case the variety of possible behavior is large, and we do not obtain anything like as comprehensive a theory as is available for ordinary diffusion.

A significant feature of the prediction process approach to the present material is thus its generality. It is sometimes possible to restrict the process to a subset which is especially chosen to fit a given P, but for the present purposes there is usually no advantage in doing so. Instead, by considering as a single packet all P such that $(X_t, P)$ has some abstract defining property, we obtain at once the results which are implied by that property. On the other hand, since the definition of $X_t$ is fixed, this approach is not as flexible as the usual one for treating all processes adapted to $F^o_{t+}$, relative to a fixed P.

## 1. THE MARTINGALE PREDICTION SPACES.

In this section we study the classes $M$, $U$, and $D$ by transplanting them to prediction space, as in Section 2 of Essay 1. In the following Section 2, it is shown how these results can be interpreted in the original setting of processes on $(\Omega, G^o)$, at least if we only deal with one process at a time. Some familiarity with the terminology and results of Essay 1, Section 2, is assumed for the present section. One new basic method is introduced which is in no way tied to the martingale setting, although it is perhaps especially well suited to martingales. This is the application of the Lévy system of a Ray process to a packet of the prediction process. In view of Corollary 2.12 of Essay 1, this is a natural step to take. Here we do not propose to exhaust its implications even for martingales, but only to use it for the limited purpose of obtaining certain well-known decomposition theorems in the prediction space setting. It is hardly surprising that these appear as results on Markov additive functionals of the prediction process, since on prediction space we have a richer structure than on the original space. A key ingredient is the fact that on prediction space the prediction process behaves well under the translation operators $\theta_t^Z$, while on the original space there is no corresponding operation. Throughout the present section, we make one significant change in the notation of Definitions 1.8 and 2.1 of Essay 1. We let $\varphi(z)$ denote only the first coordinate from its previous definition. Thus with the present restricted definition of $X_t$ we retain the fact that for each $h$, $(\varphi(Z_t), Z_t)$ is $P^h$-equivalent to $(X_t, Z_t^h)$.

We first establish that the martingale prediction spaces have the nicest possible general properties.

THEOREM 1.1. The sets $M$, $U$, and $D$ are complete Borel packets of the prediction process (in the sense of Definition 2.1, 3) of Essay 1).

PROOF. Since for every $h \in H$ the processes $X_t$ and $\varphi(Z_t)$ are $P^h$-equivalent we may verify that the three sets are in $H$ by using either $X_t$ or $\varphi(Z_t)$. We choose to use $\varphi(Z_t)$. The three sets have in common the uniform integrability of $\varphi(Z_t)$, $0 \le t$. By Fatou's lemma, we have for any $N > 0$, $h \in H$, and $t \ge 0$,

$$\liminf_{r \to t+} E^h(|\varphi(Z_r)| \; ; \; |\varphi(Z_r)| > N)$$
$$\ge E^h(|\varphi(Z_t)| ; \; |\varphi(Z_t)| > N) .$$

Hence uniform integrability is equivalently expressed, using the rationals $Q$, by the condition

$$\lim_{N \to \infty} \sup_{0 \le r \in Q} E^h(|\varphi(Z_r)| \; ; \; |\varphi(Z_r)| > N) = 0 ,$$

and the set of $h$ satisfying this is clearly in $H$. We now further restrict this set by the martingale or supermartingale conditions. By the Markov property of $Z_t$, these are respectively

$$E^{Z_t} \varphi(Z_s) \begin{cases} = \varphi(Z_t) , \\ \le \varphi(Z_t) \end{cases} P^h\text{-a.s.}$$

But a simple application of Hunt's lemma shows that when $\varphi(Z_t)$ is uniformly integrable, if these are assumed only for $0 \le s, t \in Q$, we can take right-limits to extend them to all $s,t$. Thus the class of uniformly integrable martingales (respectively, supermartingales) is in $H$.

REMARK. For the martingale case, one can proceed more simply by writing the condition as $E^{Z_t} \varphi(Z_s) = E^{Z_t} \varphi_\infty$, $P^h$-a.s., where $\varphi_\infty = \lim_{n \to \infty} \varphi(Z_n)$

exists $P^h$-a.s., which is also a Borel condition on $h$.

Now to obtain $\mathcal{M}$ we have only to append the Borel condition $E^h \varphi_\infty^2 < \infty$, so it remains to obtain $\mathcal{D}$ from the uniformly integrable supermartingales. Clearly the conditions $P^h\{\varphi(Z_t) \geq 0 \text{ all } t \geq 0\} = 1$, and $\lim_{n\to\infty} E^h(\varphi(Z_n)) = 0$, are Borel in $h$, so we need only check the class D requirement. According to a criterion of Johnson and Helms (see [4, IV, Section 1, Theorem 25]), for a positive, right-continuous supermartingale $\varphi(Z_t)$ to be of class D it is necessary and sufficient, for any increasing sequence $0 < C_k \longrightarrow \infty$, that $\lim_{k\to\infty} E^h(\varphi(Z_{T_K}); \ T_k < \infty) = 0$, where $T_k = \inf\{t: \varphi(Z_t) > C_k\}$. Obviously $\varphi(T_k) \in \mathcal{Z}^o$, hence this is again a Borel condition on $h$, as required.

Since $Z_t$ is $\mathcal{Z}_t$-optional and our three sets are Borel, to show that they are prediction packets it suffices to show that for $\mathcal{Z}_t$-optional $T < \infty$, $Z_T$ is $P^h$-a.s. in each set along with $h$. For $s \geq 0$, we introduce the sets $S_s = \{z \in H\colon E^z \varphi(Z_s) = \varphi(z)\}$, and $S_s^< = \{z \in H\colon E^z \varphi(Z_s) \leq \varphi(z)\}$. Clearly these are in $\mathcal{H}$, and the martingale (respectively supermartingale) condition on $h$ becomes $P^h\{Z_t \in S_s\} = 1$ (respectively, $P^h\{Z_t \in S_s^<\} = 1$) for all $s,t \geq 0$. By the classical optional sampling theorems of Doob (Neveu [13, Proposition IV, 5.5]) we have respectively, for $h$ in the corresponding set, $P^h\{Z_{T+t} \in S_s\} = 1$ and $P^h\{Z_{T+t} \in S_s^<\} = 1$. Therefore, we have in the first case

$$
\begin{aligned}
(1.1) \qquad 1 &= E^h P^h(Z_{T+t} \in S_s | \mathcal{Z}_T) \\
&= E^h P^{Z_T}(Z_t \in S_s) ,
\end{aligned}
$$

and therefore $P^{Z_T}\{Z_t \in S_s\} = 1$, $P^h$-a.s. for each $(s,t)$ with the corresponding result using $S_s^<$ in the second case. Therefore, they hold, for $P^h$-a.s. $Z_T$, for $0 \leq s,t \in Q$ with $P^{Z_T}$-probability one. But this means, in the martingale cases, that $\varphi(Z_t)$ is a $P^{Z_T}$-martingale, at least for rational $t$. Consequently, for fixed $K$ the $\varphi(Z_t)$ are $P^{Z_T}$-uniformly integrable for rational $t \leq K$, $P^{Z_T}$-a.s. Then as in the first part of the proof we use Hunt's Lemma to extend the martingale property to all $(s,t)$ with $s + t < K$, and then let $K \to \infty$. The case of positive supermartingales is a little different. One first observes that simply by martingale convergence of conditional expectations, if $\varphi(Z_r)$ is a $P^{Z_T}$-positive-supermartingale as $r$ varies in $Q$, then for any $0 \leq t < r$, $r \in Q$, one has

$$E^{Z_t}\varphi(Z_{r-t}) = \lim_{\substack{s\to t+\\ s\in Q}} E^{Z_T}(\varphi(Z_r)|\mathcal{Z}_s)$$

$$\leq \lim_{s\to t+} \varphi(Z_s)$$

$$= \varphi(Z_t), \quad P^{Z_T}\text{-a.s.}$$

Then by positivity and Fatou's Lemma, for $0 \leq t < s$ one has

$$E^{Z_t}\varphi(Z_{s-t}) \leq \liminf_{r\to s+} E^{Z_t}\varphi(Z_{r-t})$$

$$\leq \varphi(Z_t), \quad P^{Z_T}\text{-a.s.}$$

Thus $\varphi(Z_t)$ is a $P^{Z_T}$-positive supermartingale, $P^h$-a.s.

The uniform integrability, or square integrability, of the $P^{Z_T}$-martingales now follows easily from the fact that convergence of $\varphi(Z_{T+t})$ to $\varphi_\infty$ in $L^1$ or $L^2$ for $P^h$ as $t\to\infty$ implies the same convergence conditional on $\mathcal{Z}_T$, at least for a sequence $t_k\to\infty$ sufficiently fast. This suffices to identify $\varphi_\infty$ as the value at $t=\infty$ of the $P^{Z_T}$-martingales for $P^h$-a.e. $Z_T$ (see Neveu [13, IV, 5.6]).

It remains to verify that $Z_t$ is of class D for $P^{Z_T}$ in the supermartingale case, and to show that the three packets are complete. With $C_k$ as in the first part of the proof, and optional $T<\infty$, let $T_k' = \inf\{t > T: \varphi(Z_t) > C_k\}$. Then with the previous $T_k$ we have $T_k \leq T_k'$ and, for $h \in \mathcal{D}$, $\lim_{k\to\infty} E^h(\varphi(Z_{T_k});\ T_k<\infty) = 0$. It follows that $E^h(\varphi(Z_{T_k});\ T_k < \infty|\mathcal{Z}_T)$ tends to zero in probability as $k\to\infty$. Now we have

$$E^h(\varphi(Z_{T_k'});\ T_k' < \infty|\mathcal{Z}_{T\vee T_k}) \tag{1.2}$$

$$\leq \varphi(Z_{(T\vee T_k)})\, I_{\{T_k<\infty\}}, \quad P^h\text{-a.s.},$$

by [4, VI, Section 1, [10]], hence by conditioning on $\mathcal{Z}_T$ we have

$$E^h(\varphi(Z_{T_k'});\ T_k' < \infty|\mathcal{Z}_T) \tag{1.3}$$

$$\leq E^h(\varphi(Z_{T_k\vee T});\ T_k < \infty|\mathcal{Z}_T)\ .$$

Therefore, the left side also tends to 0 in probability, and so there is a subsequence $k_j$ for which it converges to 0 $P^h$-a.s. Clearly, then, the sequence $C_{k_j}$ satisfies the Johnson-Helms criterion for $Z_T$, $P^h$-a.s., proving that $Z_T$ is a.s. of class D .

Turning, finally, to the completeness, we must show that for $\mathcal{Z}_t$-previsible $0 < T < \infty$, $Z_{T-}$ is in the corresponding packet, $P^h$-a.s. First, we note that $E^h E^{Z_{T-}} |\varphi(Z_t)| = E^h|\varphi(Z_{T+t})| < \infty$, hence $\varphi(Z_t)$ has finite expectation for $P^{Z_{T-}}$, $P^h$-a.s. We can now verify the martingale property (resp. positive supermartingale property) just as before, except that in (1.1) the conditioning is on $\mathcal{Z}_{T-}$ . The uniform or square integrability then follows in the martingale cases, as before, by convergence of $\varphi(Z_{T+t_k})$ to $\varphi_\infty$, conditional on $\mathcal{Z}_{T-}$, and it only remains to check the class D restriction in the supermartingale case. It is clear as before that $E^h(\varphi(Z_{T_k \vee T}); T_k < \infty | \mathcal{Z}_{T-})$ converges in probability to 0 as $k \to \infty$ . Conditioning both sides of (1.3) by $\mathcal{Z}_{T-}$, we then obtain the same criterion, and the proof of Theorem 1.1 is complete.

REMARK. The supermartingale case could also have been handled by means of the Doob-Meyer decomposition of $\varphi(Z_t)$, but since our intention is to obtain this decomposition from the prediction space, this would lead to a circular reasoning.

COROLLARY 1.1. The prediction process is a right-process on $\mathcal{M} \cap H_0$, $\mathcal{U} \cap H_0$, or $\mathcal{D} \cap H_0$, and in each case $\varphi(z)$ is an excessive function. On $\mathcal{M} \cap H_0$ and $\mathcal{U} \cap H_0$, $\varphi(z)$ is an invariant function. On $\mathcal{D} \cap H_0$ it is a potential of class D.

REMARKS. Since the literature of excessive functions is usually confined to standard processes, this terminology is not quite orthodox. For standard processes, such $\varphi$ are considered under (1) and (2) of the Notes and Comments to Chapter IV in [2].

PROOF. In all three cases, for $z \in H_0$ we have $E^z\varphi(Z_t) \le \varphi(z)$, and $\lim_{t\to 0} E^z\varphi(Z_t) = \varphi(z)$ by right-continuity of $\varphi(Z_t)$ and the fact that $P^z\{\varphi(Z_0) = \varphi(z)\} = 1$ for $z \in H_0$ . Invariance, by definition, becomes the martingale property $E^z\varphi(Z_t) = \varphi(z)$ . For the last assertion, which is again true by definition, we observe that for any increasing sequence $T_n \to \infty$ of stopping times one has $\lim_{n\to\infty} E^z\varphi(T_n) = 0$ for $z \in \mathcal{D} \cap H_0$, since the $\varphi(T_n)$ are $P^z$-uniformly integrable and, by supermartingale convergence, $\lim_{t\to\infty} \varphi(Z_t) = 0$ a.s.

We next take up the discontinuities of $\varphi(Z_t)$. Here our chief tool is the Lévy system of $Z_t$ on the corresponding packet. The theory of Levy systems was initiated by M. Motoo and S. Watanabe under the hypothesis of absolute continuity [12], and developed further by J. Walsh and M. Weil [16]. The final touches, and also the simplest proofs, are provided by A. Benveniste and J. Jacod [1], whose formulation applies to all the discontinuity times of any Ray process. Since we know by Corollary 2.12 of Essay 1 that on any Borel prediction packet the prediction process is (in a sense) a Ray process restricted to a suitable Borel set, it is natural to use the result of Corollary 5.2 of [1] which we now describe.

Continuing the notation (however unwieldy) of Essay 1, let $H_A$ be any <u>Borel</u> prediction packet, and let $(H_A \cap H_0)^+$ be the Ray compactification of its "non-branching" points. We denote the canonical Ray process by $\bar{X}_t$, with probabilities $\bar{P}^x$, transition function $\bar{p}$, and resolvent $\bar{R}_\lambda$. Then there exists a Levy system of $\bar{X}_t$, which consists of four parts, N, M, H, and L. Here $N = N(x,dy)$ and $M = M(x,dy)$ are Borel measure kernels on $(H_A \cap H_0)^+$, while $H = H_t(\bar{w})$ and $L = L_t(\bar{w})$ are additive functionals on the probability space of $\bar{X}_t$. In more detail, $N(x,dy)$ is defined for $x \in D$, where $D = (H_A \cap H_0)^+ - B$ (with B as before denoting the Ray branching points), while $M(x,dy)$ is defined for $x \in B$. Both N and M yield finite measures on the Borel sets of D for each x where defined, and are Borel measurable in x (B is a Ray-Borel set). Further, $N(x,\{x\}) = 0$, $x \in D$, and $M(x,\{x\}) = 0$, $x \in B$. Next, we have $E^x H_t < \infty$ and $E^x L_t < \infty$ for $x \in D$ and all t. Finally, $H_t$ is a continuous additive functional while L is a purely discontinuous additive functional which is previsible with respect to the usual $\sigma$-fields $\bar{F}_t$ of the Ray process $\bar{X}_t$. In fact, we have $L_t = \sum_{0<s\le t} f(\bar{X}_{s-})\, I_{\{\bar{X}_{s-} \in B\}}$ for a Borel $f \ge 0$.

These four objects have the property that (N,H) "compensates" the totally inaccessible discontinuities of $Z_t$ while (M,L) compensates the previsible discontinuities. In more detail, let $f(x,y)$ be any non-negative, jointly-Ray-Borel function with $f(x,x) = 0$. Then

(1.4)(a)
$$E^x \sum_{0<s\le t} f(\bar{X}_{s-},\bar{X}_s)\, I_{\{\bar{X}_{s-} \in D\}}$$

$$= E^x \int_0^t dH_s \int_D N(\bar{X}_{s-},dy)\, f(\bar{X}_{s-},y) ,$$

and

(1.4)(b) $$E^x \sum_{0<s\leq t} f(\bar{X}_{s-},\bar{X}_s)\, I_{\{\bar{X}_{s-} \in B\}}\ ,$$

$$= E^x \int_0^t dL_s \int_D M(\bar{X}_{s-},dy)\, f(\bar{X}_{s-},y)\ ,$$

for all $x$ and $t > 0$. We note that on the right side of (a), $\bar{X}_{s-}$ may be replaced by $\bar{X}_s$ since $H_s$ is continuous.

We combine the Lévy system by setting $\bar{N} = N + M$, where $N$ and $M$ are extended to be the zero measure for $x$ at which they were undefined, and $\bar{H} = H + L$. Thus we have finally

(1.4)(c) $$E^x \sum_{0<s\leq t} f(\bar{X}_{s-},\ \bar{X}_s)$$

$$= E^x \int_0^t d\bar{H}_s \int_D \bar{N}(\bar{X}_{s-},dy)\, f(\bar{X}_{s-},y)\ ,$$

for $f$, $x$, and $t$ as before.

In order to apply this to the packet $H_A$, since left limits are involved and we do not assume $H_A$ is complete, we use the complete Borel packet $h(D_A) \cap H$ of Essay 1, Corollary 2.12. We recall that $H_A \subset (h(D_A) \cap H)$, and for $h \in (h(D_A) \cap H)$ we have $P^h\{Z_t \in H_A \cap H_0$ for all $t > 0\} = 1$. Thus $h(D_A) \cap H$ can be regarded as a packet of entrance laws for $Z_t$ on $H_A \cap H_0$ which, in addition, are given by elements of $H$ (and therefore, by distributions on $h(D_A) \cap H_0$). Since such an entrance law is determined by $\bar{P}^z$ for at most one point of $(H_A \cap H_0)^+$, and in the present case by at least one, we see that the mapping $h(z)$ is one-to-one on $D_A$. Thus we can introduce the inverse $h^{-1}$: $h(D_A) \to D_A$, and for any initial distribution $\mu$ for $Z_t$ on $h(D_A) \cap H_0$ we can identify $\bar{X}_t = h^{-1}(Z_t)$ and $\bar{X}_{z-} = h^{-1}(Z_{t-})$ for $t > 0$ as a realization of a Ray process with initial distribution $\bar{\mu}(S) = \mu(h(S))$ on $h^{-1}(h(D_A) \cap H_0)$. Furthermore, we showed in the proof of Essay 1, Theorem 2.13, that $h^{-1}(h(D_A) \cap H_0) = h^{-1}(h(D_A) \cap H) - B$, or in other words, for the right process $Z_t$ on $h(D_A) \cap H_0$ with left limits in $h(D_A) \cap H$, the elements of $(h(D_A) \cap H) - H_0$ correspond under $h^{-1}$ to the Ray-branching points in $h^{-1}(h(D_A) \cap H)$. Thus we can transfer the Lévy system of $\bar{X}_t$ to obtain a Lévy system of $Z_t$ on $h(D_A) \cap H$.

In detail, let $\Omega_{Z,A} = \{w_Z \in \Omega_Z$ such that $w_Z(t) \in h(D_A) \cap H_0$, $t \geq 0$, and $w_Z(t-) \in h(D_A) \cap H$, $t > 0\}$. Then $\Omega_{Z,A}$, with the $\sigma$-fields $\{A \cap \Omega_{Z,A};\ A \in Z_t^o\}$ on $\Omega_{Z,A}$, is canonical sample space for $Z_t$ as a right process on $h(D_A) \cap H_0$. Using this sample space, we define the four elements of a Lévy system by

(1.5) $N_Z(h_1, dh_2) = N(h^{-1}(h_1), h^{-1}(dh_2));\ h_1 \in h(D_A) \cap H_0$

$M_Z(h_1, dh_2) = M(h^{-1}(h_1), h^{-1}(dh_2));\ h_1 \in h(D_A) - H_0$ ,

$H_{Z,t}(w_Z) = H_t(\bar{w});\ w_z \in \Omega_{Z,A}$ ,

$L_{Z,t}(w_Z) = L(\bar{w});\ w_Z \in \Omega_{Z,A}$ ,

where $\bar{w}(t) = h^{-1}(w_Z(t))$ for $t \geq 0$ . Then since $\theta_t^Z w_Z$ corresponds to $\bar{\theta}_t \bar{w}$ as $w_Z$ does to $\bar{w}$ (where $\bar{\theta}_t$ is the Ray-translation operator), we see that $H_Z$ and $L_Z$ are additive functionals of $Z_t$ on $\Omega_{Z,A}$ . Also $H_Z$ is continuous, while since

$$L_{Z,t}(w_Z) = \sum_{0<s\leq t} f(\bar{w}_{s-})\, I_{\{\bar{w}(s-) \in B\}}$$

$$= \sum_{0<s\leq t} f(h^{-1}(w_Z(s-))I_{\{w_Z(s-)\in H-H_0\}} ,$$

where $f\,h^{-1}$ is Borel on $h(D_A) \cap H$, $L_Z$ is $\mathcal{Z}_t$-previsible and purely discontinuous. Of course, the local integrability of $H_Z$ and $L_Z$ carries over to $z \in h(D_A) \cap H_0$ .

It is now just a matter of transferring (1.4) to the present context, and setting $\bar{N}_Z = N_Z + M_Z$ and $\bar{H}_Z = H_Z + L_Z$, to obtain

THEOREM 1.2. For $0 \leq f(z_1, z_2) \in \mathcal{H} \times \mathcal{H}$, $f(z,z) = 0$, the objects (1.5) satisfy

(1.6) (a) $$E^z \sum_{0<s\leq t} f(Z_{s-}, Z_s)\, I_{\{Z_{s-} \in H_0\}}$$

$$= E^z \int_0^t dH_{Z,s} \int_{H_0} N_Z(Z_{s-}, dz)\, f(Z_{s-}, z)$$

(b) $$E^z \sum_{0<s\leq t} f(Z_{s-}, Z_s)\, I_{\{Z_{s-} \in H-H_0\}}$$

$$= E^z \int_0^t dL_{Z,s} \int_{H_0} M_Z(Z_{s-}, dz)\, f(Z_{s-}, z)$$

(c) $$E^z \sum_{0<s\leq t} f(Z_{s-}, Z_s)$$

$$= E^z \int_0^t d\bar{H}_{Z,s} \int_{H_0} \bar{N}_Z(Z_{s-}, dz)\, f(Z_{s-}, z)$$

for all $z \in h(D_A) \cap H_0$ and $t > 0$ .

Of course, the kernels $N_Z$, $M_Z$, $\bar{N}_Z$ are Borel in $z$ and the measures are concentrated on $h(D_A) \cap H_0$ . Also, we may as well assume $f(z_1, z_2) = 0$

except for $z_1 \in h(D_A) \cap H$ and $z_2 \in h(D_A) \cap H_0$. On the other hand, our use of subscript Z instead of A for the elements of the Lévy system is quite appropriate for the following reason. We could just as well begin with the case $H_A = H$, $h(D_A) \cap H_0 = H_0$. Then we obtain the four components of the Levy system in a form which applies, except for negligible changes, to <u>any</u> Borel prediction packet $H_A \subset H$. In fact, the only objection to identifying the restrictions of these components to $h(D_A) \cap H$ and $\Omega_{Z,A}$ with the components of Theorem 1.2, in general, is that the measure kernels might not be suitably restricted to the corresponding $h(D_A) \cap H_0$. However, for any fixed $H_A$ one may redefine these measures to be 0 outside $h(D_A) \cap H_0$ without losing property (1.6). This follows by substituting $f(z_1,z_2) = 1 - I_{h(D_A) \cap H_0}(z_2)$ in (1.6) and noting that for $z \in h(D_A) \cap H_0$ the result is 0.

Another form of the same observation is useful in treating $\mathcal{M}$, $\mathcal{U}$, and $\mathcal{D}$. We may and do take as sample spaces the canonical prediction spaces of all elements of $\Omega_Z$ with values and left limits in the respective (complete) packet. An inspection of (1.6) shows that we may just as well restrict the components of the Lévy system from $H_A = H$ to any such complete Borel packet, instead of just $h(D_A) \cap H$. (This can also be seen by intersecting the packet first with the corresponding $h(D_A)$, but the step is unnecessary.) We may state

THEOREM 1.3. There exists a Lévy system for any complete Borel prediction packet and corresponding r.c.l.l. process $Z_t$. In fact, the components of Theorem 1.2 with $H_A = H$ may be restricted to yield such a system.

For application to $\varphi(Z_t)$, we need to restate the properties of the Lévy system in a somewhat different form.

COROLLARY 1.3. For any complete Borel prediction packet, the properties (1.6) of the Levy system imply that for any $0 \le f(z_1,z_2) \in \bar{\mathcal{B}} \times \mathcal{H}$, and any $\mathcal{Z}_t$-previsible process $y_t \in \bar{R}$, (1.6) holds with $f(Z_{t-},z)$ replaced by $f(y_t,z)I_{(Z_{t-} \neq z)}$.

PROOF. We justify the substitution in (1.6)(c), the other two cases being analogous. First we consider f and $y_t$ of the special form $f(y,z) = k(y)g(z)$, $k \in \bar{\mathcal{B}}$, $g \in \mathcal{H}$, and $y_t = I_A(w_Z)I_A(w_Z)I_{(t_1,t_2]}(t)$ for $A \in \mathcal{Z}_{t_1-}$. Now by (1.6)(a) and the Markov property of $Z_t$, for z in the packet we have

$$(1.7)\quad E^z\left[\sum_{t_1<s\le t_2} g(Z_s)\, I_{\{Z_s\ne Z_{s-}\}} \;\middle|\; Z_{t_1-}\right]$$
$$= \lim_{r\to t_1-} E^z\left[\sum_{t_1<s\le t_2} g(Z_s)\, I_{\{Z_s\ne Z_{s-}\}} \;\middle|\; Z_r\right]$$
$$= \lim_{r\to t_1-} E^{Z_r}\int_{t_1-r}^{t_2-r} d\bar{H}_{Z,s}\int_{H_0} N_Z(Z_{s-},dz) I_{\{Z_{s-}\ne z\}}\, g(z)$$
$$= E^z\left[\int_{t_1}^{t_2} d\bar{H}_{Z,s}\int_{H_0} \bar{N}_Z(Z_{s-},dz) I_{\{Z_{s-}\ne z\}}\, g(z) \;\middle|\; Z_{t_1-}\right].$$

Multiplying both sides of (1.7) by $I_A$ and taking the expectations $E^z$ to remove the conditioning, we obtain the assertion for this $y_t$ and $k(y) = |y|$. By [3, Chapter IV, Theorem 67], the previsible $\sigma$-field is generated by such $y_t$ and sets of the form $\{0\} \times A$, $A \in \mathcal{Z}_{0-}$, whose indicators satisfy the assertion trivially. Moreover, the class of finite sums of disjoint indicators of the form $y_t$ is closed under multiplication. Since the class of indicators $y_t$ satisfying the assertion is monotone, it follows by [3, Chapter 1, 19] that it contains all $\mathcal{Z}_t$-previsible indicator functions, and hence all previsible $y_t \ge 0$. Since $k(y_t)$ is again prevísbly for $0 \le k \in \bar{\mathcal{B}}$, we obtain the result for $f(z_1,z_2) = k(z_1)g(z_2)$ immediately. Therefore, it holds for finite positive linear combinations of such, and by [3, Chapter 1, 21] it holds for $0 \le f \in \bar{\mathcal{B}} \times \mathcal{H}$ as asserted.

We first apply this method to obtain a well-known decomposition of square integrable martingales due to P. A. Meyer [11] and Kunita, Watanabe [10]. We recall that two square integrable martingales are called orthogonal if their product is a martingale. If they are additive functionals of $Z_t$, then one requires this to hold for every $P^h$, $h \in \mathcal{M}$. The following notation will be used for $h \in \mathcal{M}$, $\mathcal{U}$, or $\mathcal{D}$, according to context.

NOTATION 1.4. For $t > 0$, let $\varphi_-(t) = \limsup_{Q\ni r\to t-} \varphi(Z_t)$. We recall that for any $h \in H$, $\varphi_-(t)$ exists $P^h$-a.s. for all $t > 0$ as an ordinary left limit, and $\varphi_-(t)$ is $P^h$-equivalent in distribution to $X_{t-}$ (Definition 2.1, 2), of Essay 1), as $\varphi(Z_t)$ is to $X_t$.

Our object is to disengage the jumps of $\varphi(Z_t)$ into a separate martingale, called the "discontinuous part," by means of the Lévy system. For this we need

LEMMA 1.5. For either $h \in \mathcal{M}$ or $h \in \mathcal{U}$, $P^h\{\varphi_-(t) \ne \varphi(Z_t)$ implies $Z_{t-} \ne Z_t$, all $t > 0\} = 1$. In words, the discontinuity times of $\varphi(Z_t)$ are contined in those of $Z_t$.

PROOF. For $\varepsilon > 0$, let $T_\varepsilon = \inf\{t > 0: |\varphi(Z_t) - \varphi_-(t)| > \varepsilon\}$. Clearly $T_\varepsilon$ is a $\mathcal{Z}_t$-stopping time, and by right-continuity of $\varphi(Z_t)$ we have $P^h\{T_\varepsilon > 0\} = 1$ for all $h$. By the strong Markov property and the existence of $\varphi_-(t)$ for all $t > 0$, the iterates $T_\varepsilon^{n+1} = T_\varepsilon \circ \theta^Z_{T_\varepsilon^n}$, with $T_\varepsilon^1 = T_\varepsilon$, tend to $\infty$ along with $n$, $P^h$-a.s. Hence by the strong Markov property it suffices to show that

$$P^h\{Z_{T_\varepsilon^-} \neq Z_{T_\varepsilon}\} = P^h\{T_\varepsilon < \infty\} .$$

By [3, IV, Theorem 81c)] there is a decomposition $T = T_A \wedge T_{A^c}$, where $T_A$ is the totally inaccessible part of $T$ for $(\mathcal{Z}_t^h, P^h)$, and $T_{A^c} = T_{\Omega_Z - A}$ is the accessible part. According to Theorem 2.13(ii) of Essay 1, we always have $Z_{T_\varepsilon^-} \neq Z_{T_\varepsilon}$ $P^h$-a.s. on $\{T_\varepsilon = T_A\}$. On the other hand, since $Z_{T_\varepsilon^-}$ is $\mathcal{Z}^h_{T_\varepsilon^-}$-measurable on $A^c$ ([3, IV, 57]), and by the moderate Markov property we have

$$P^h(\varphi(Z_{T_\varepsilon}) \in B \mid \mathcal{Z}^h_{T_\varepsilon^-}) = q(0, Z_{T_\varepsilon^-}, \{\varphi(z) - B\}) \tag{1.8}$$

$P^h$-a.s. on $A^c \cap \{T_\varepsilon < \infty\}$ for $B \in \mathcal{B}$, it follows that

$$P^h(\varphi(Z_{T_\varepsilon}) = \varphi(Z_{T_\varepsilon^-}) \mid \mathcal{Z}^h_{T_\varepsilon^-}) = 1$$

on $\{Z_{T_\varepsilon^-} \in H_0\} \cap A^c \cap \{T_\varepsilon < \infty\}$, $P^h$-a.s. Therefore, we have $Z_{T_\varepsilon^-} \notin H_0$, $P^h$-a.s. on $A^c \cap \{T_\varepsilon < \infty\}$, and so $Z_{T_\varepsilon -} \neq Z_{T_\varepsilon}$ as required.

We now state and prove the decomposition theorem in $M$. Let $\Omega_M = \{w_z: w_z(t) \in M$ for all $t \geq 0$ and $w_z(t-) \in M$ for all $t > 0\}$.

THEOREM 1.6. There is a decomposition $\varphi(Z_t) - \varphi(Z_0) = M_c(t) + M_d(t)$, where for $h \in M$, $M_c$ is a continuous, $P^h$-square-integrable, martingale additive functional on $\Omega_M$, $M_d$ is an $E^h$-mean-square limit of martingale additive functionals of bounded variation, and $M_c$ is orthogonal to $M_d$. The decomposition is unique up to a $P^h$-null set for all $h$.

In the course of the proof, and also later, we need

NOTATION 1.7. For any r.c.l.l. process $M(t)$, let $\Delta M(t) = M(t) - M(t-)$, $t > 0$, where $M(t-)$ denotes the left limit at time $t$, and we use $\infty - \infty = 0$. In particular, let $\Delta\varphi(t) = \varphi(Z_t) - \varphi_-(t)$.

PROOF. If $M(t)$ is any $P^h$-square-integrable martingale (in particular $E^h M^2(\infty) < \infty$) then it is a familiar fact that $M(t)$ has orthogonal increments. Thus if $\{t_{n,i}, 1 \le i \le n\}$ is a sequence of partitions of $[0,t]$ with maximum separation tending to 0, then by Fatou's Lemma, with $t_{n,0} = 0$,

$$E^h M^2(t) = \lim_{n\to\infty} E^h \sum_{i=1}^{n} (M(t_{n,i}) - M(t_{n,i-1}))^2$$

$$\ge \lim_{\varepsilon\to 0+} E^h \sum_{t_{j,\varepsilon}} (\Delta M(t_{j,\varepsilon}))^2$$

where the last sum is over all $t_{j,\varepsilon} \le t$ such that $\Delta^2 M(t_{j,\varepsilon}) > \varepsilon$. Letting $t \to \infty$, it follows that

$$E^h M^2(\infty) \ge E^h \sum_{t_j} (\Delta M(t_j))^2 , \tag{1.9}$$

where $t_j$ enumerate the discontinuity times of $M(t)$.

We now fix $0 < a < b$, and apply Corollary 1.3 with $y_t = \varphi_-(t)$ (which is $\mathcal{Z}_t$-previsible by [3, IV, Theorem 92]), and $f(y,z) = (\varphi(z) - y) I_{(a<\varphi(z)-y\le b)}$. Letting $\varphi_\infty = \limsup_{t\to\infty} \varphi(Z_t)$, (1.9) implies that

$$E^h \Big| \sum_{0<s\le t} f(\varphi_-(s), Z_s) \Big| \le a^{-1} E^h (\varphi_\infty - \varphi(Z_0))^2 < \infty .$$

Then we may subtract the right side in Corollary 1.3, and by Lemma 1.5 and the Markov property of $Z_t$ we obtain that the process

$$M_{(a,b)}(t) = \Big[ \sum_{0<s\le t} \Delta\varphi(s) I_{(a<\Delta\varphi(s)\le b)} - \int_0^t d\bar{H}_{Z,s} \int_{H_0} \bar{N}_Z(Z_{s-},dz) f(\varphi_-(s),z) \Big]$$

is a martingale additive function of $Z_t$ on $\Omega_{\mathcal{M}}$ (here $f(y,z)$ was substituted explicitly only in the sum).

The martingale $M_{(a,b)}(t)$ is clearly of bounded variation, and we now evaluate its mean square precisely. Denoting the above difference by $M^+_{(a,b)}(t) - M^-_{(a,b)}(t)$, let $T_N = \inf\{t: M^+_{(a,b)}(t) + M^-_{(a,b)}(t) \ge N\}$. Then $T_N$ is a $\mathcal{Z}_t$-stopping time, $M^+_{(a,b)}(T_N) \le N + b$, and $T_N \to \infty$, $P^h$-a.s., as $N \to \infty$. Also, as in (1.4), $M^-_{(a,b)}(t)$ has at most only accessible times of discontinuity, and for previsible $T < \infty$ it follows by (1.8) of Lemma 1.5 and Jensen's inequality for conditional expectations that

$$E^h (\Delta M^-_{(a,b)}(T))^2 \le E^h (\Delta M^+_{(a,b)}(T))^2 .$$

Then by decomposition of $T_N$ we have easily

$$(1.10) \qquad E^h(M^-_{(a,b)}(T_N))^2 \le E^h(N+\Delta M^-_{(a,b)}(T_N))^2$$
$$\le E^h(N+\Delta M^+_{(a,b)}(T_N))^2$$
$$\le (N+b)^2 .$$

Next, for $t > 0$ fixed, let $t_{n,j} = jt2^{-n}$, and reapply the argument beginning the proof to the martingale $M_{a,b}(t \wedge T_N)$. Simply by decomposing paths of bounded variation into continuous and jump components, we see that the sums of the squared increments of $M_{a,b}(t \wedge T_N)$ along the partitions $t_{n,j}$ converge $P^h$-a.s. as $n \to \infty$ to $\sum_{t_i \le t} (\Delta M_{(a,b)}(t_i \wedge T_N))^2$, where the sum is over the jump times less than $t$. Also, the sums of squares of increments of $M^+$ alone are decreasing with $n$, as are those of $M^-$ alone. Using $(c-d)^2 \le (c^2+d^2)$ to bound the squared increments of $M$, the sums are dominated by $(M^+_{(a,b)}(t \wedge T_N))^2 + (M^-_{(a,b)}(t \wedge T_N))^2$, which has finite expectation. Hence by the dominated convergence theorem,

$$E^h M^2_{(a,b)}(t \wedge T_N) = E^h \sum_{t_i \le t} (\Delta M_{a,b}(t_i \wedge T_N))^2 .$$

Letting $N \to \infty$, it follows readily that

$$(1.11) \qquad E^h M^2_{(a,b)}(t) = E^h \sum_{t_i \le t} (\Delta M_{a,b}(t_i))^2 .$$

In particular, by (1.9) it follows that $E^h M^2_{(a,b)}(t) \le E^h M^2(\infty)$, hence $M_{(a,b)}$ is square-integrable.

Furthermore, for $0 < a < b < c$ we have by (1.11) that

$$E^h M^2_{(a,c)}(t) = E^h(M_{(a,b)}(t) + M_{(b,c)}(t))^2$$
$$= E^h M^2_{(a,b)}(t) + E^h M^2_{(b,c)}(t) .$$

Hence $E^h(M_{(a,b)}(t)\, M_{(b,c)}(t)) = 0$, and by the Markov property of $Z_t$ it follows that $M_{(a,b)}$ and $M_{(b,c)}$ are orthogonal. Similarly, for $c < d < 0$ we can define $M_{(c,d)}$ to compensate the negative jumps $c \le \varphi(Z_s) - \varphi_-(s) < d$, and (1.11) applies. Finally, since $-M_{(c,d)}$ has the same form as $M_{(a,b)}$, it is seen that

$$E^h(M_{(a,b)}(t) - M_{(c,d)}(t))^2$$

$$= E^h M^2_{(a,b)}(t) + E^h M^2_{(c,d)}(t) .$$

Hence $M_{(a,b)}$ and $M_{(c,d)}$ are orthogonal, and

$$E^h(M_{(a,b)}(t) + M_{(c,d)}(t))^2 \leq E^h M^2(\infty) .$$

We next choose a sequence $a_n \to 0+$, $b_n \to \infty$, $c_n \to -\infty$, $d_n \to 0-$. It follows directly from the above that for $h \in \mathcal{M}$ and $0 \leq t \leq \infty$ there exist $E^h$-mean-square limits of $M_{(a,b)}(t) + M_{(c,d)}(t)$ along this sequence. Furthermore, it is known from general theorems of analysis that such limits always may be chosen so as to be valid for all $h$ (see [14, Theorem 3]). Accordingly, we denote such a choice by $M^*_d(t)$,, and define

$$M_d(t) = \begin{cases} \lim\limits_{Q \ni r \to t+} M^*_d(r) & \text{if this exists for all } t < \infty \text{ and} \\ & \text{equals } 0 \text{ for } t = 0 , \\ 0 & \text{elsewhere.} \end{cases}$$

For each $h$, we have easily

$$M^*_d(r) = E^h(M^*_d(\infty) | Z_r) , \quad P^h\text{-a.s.} ,$$

from which it follows that $M_d(t)$ is a right-continuous version of $E^h(M^*_d(\infty)|Z_t)$ for each $h$, and thus it is a square-integrable martingale. To see that it is an additive functional of $Z_t$, we note that for fixed $s,t$, and $h \in \mathcal{M}$ we can choose $\alpha_k$, $\beta_k$, $\gamma_k$, and $\delta_k$ such that $P^h(S_{s+t}) = 1$, $P^h(S_t) = 1$, and $P^{Z_t}(S_s) = 1$ for $P^h$-a.e. $Z_t$, where $S_u$, $u \geq 0$, is given by

$$S_u = \{M_d(u) = \lim_{k\to\infty} (M_{(\alpha_k,\beta_k)}(u) + M_{(\gamma_k,\delta_k)}(u))\} .$$

Since then

$$P^h(\theta_t^{-1} S_s) = E^h(P^{Z_t}(S_s)) = 1 ,$$

the property $M_d(s+t) = M_d(t) + M_d(s) \circ \theta_t$, $P^h$-a.s., follows from the corresponding fact for $M_{(a,b)}(t) + M_{(c,d)}(t)$.

Similarly, it follows from a classical martingale theorem of Doob ([5, (Theorem 5.1), p. 363]) that for each $h$ we can choose a subsequence $n_k$ for which $M_d$ is the limit of $M_{(a,b)} + M_{(c,d)}$ uniformly in $t$,

$P^h$-a.s. for $a = a_{n_k}$, etc. Clearly, then, $M_d(t)$ contains all the totally inaccessible jumps of $\varphi(Z_t)$. But for previsible $T < \infty$, we have $\Delta M_d(T) = \Delta\varphi(T)$ plus a quantity which is $\mathcal{Z}_{T-}$-measurable along with the $\Delta M^-_{(a,b)}(T)$, and since $E^h(\Delta M_d(T)\,|\,\mathcal{Z}_{T-}) = 0$, this quantity must be $0$. Hence we see that $M_c(t) = \varphi(Z_t) - \varphi(Z_0) - M_d(t)$ defines a continuous martingale additive functional of $Z_t$. It remains only to show that $M_c$ and $M_d$ are orthogonal, or again that $M_c$ and $M_{(a,b)}$ are orthogonal. To this effect we have only to apply some of the argument for (1.11) to $M_c(t) + M_{(a,b)}(t)$ with $T_N$ redefined by $\inf\{t: (M^+_{(a,b)}(t) + M^-_{(a,b)}(t) + |M_c(t)|) \geq N\}$. It follows readily that in computing $E^h(M_c(t \wedge T_N) + M_{(a,b)}(t \wedge T_N))^2$ along $\{t_{n,j}\}$ the sum in $j$ of the cross-products of increments of $M_c$ and $M_{(a,b)}$ in $(t_{n,j-1}, t_{n,j})$ is bounded by $2N(M^+_{(a,b)}(T_N) + M^-_{(a,b)}(T_N))$, which has finite expectation and the sum tends to $0$ along with the partition size $2^{-n}$, $P^h$-a.s. By dominated convergence we obtain $E^h(M_c(t \wedge T_N)\, M_{(a,b)}(t \wedge T_N)) = 0$, and to conclude the existence proof it suffices to observe that, by using Fatou's Lemma,

$$\lim_{N\to\infty} E^h(M(t) - M(t \wedge T_N))^2$$

$$= E^h M^2(t) - \lim_{N\to\infty} E^h(M^2(t \wedge T_N))$$

$$= 0 ,$$

for any $P^h$-square-integrable martingale $M(t)$ with respect to $\mathcal{Z}_t$.

As to the uniqueness, since any two choices for $M_d$ differ by a continuous martingale additive functional, it needs only be shown that such cannot be the $E^h$-mean-square limit of martingales of integrable total variation unless it is $E^h$-a.s. identically $0$. The reader will readily check that the proof of orthogonality of $M_c$ and $M_{(a,b)}$ just given applies without change to any square integrable martingales which are respectively continuous and of integrable total variation (where $M^+(t) + M^-(t)$ is defined to be the total variation at time $t$). Hence the former cannot be approximated in the mean square by the latter, and the uniqueness is proved.

In the present section we make no further use of the packet $\mathcal{U}$, except to remark that any class D right-continuous submartingale $X_t$

may be decomposed in the form $X_t - E(X_\infty|G^o_{t+}) - (E(X_\infty|G^o_{t+}) - X_t)$ where the first process on the right is in $\mathcal{U}$ and the second is in $\mathcal{D}$. For the elements of the packet $\mathcal{D}$, we will derive the celebrated Doob-Meyer decomposition theorem as a theorem on Markov processes. It will be seen that this yields the corresponding decomposition result for $X_t$ by expressing it in the above form.

Many proofs of the Doob-Meyer decomposition are known, and some are perhaps easier than ours. Nevertheless, ours seems worthwhile because it connects the decomposition with the prediction process, and provides additive functionals where the decomposition alone only provides unrelated pairs of processes. Besides, it does not use the theory of Levy systems, and most of the work needed for the proof has already been done in [2, Chapter 4, Section 3] and therefore need not be repeated here. We let

$$\Omega = \{w_Z : w_Z(t) \in \mathcal{D} \text{ for } t \geq 0 \text{ and } w_Z(t-) \in \mathcal{D} \text{ for } t > 0\} .$$

The result to be proved is as follows.

THEOREM 1.8. There is a decomposition

$$\varphi(Z_t) - \varphi(Z_0) = M(t) - A(t) \text{ on } \Omega_{\mathcal{D}} ,$$

where $A(t)$ is a (non-decreasing) additive functional of $Z_t$, $Z^h_t$-previsible for every $h \in \mathcal{D}$, and $M(t)$ is a uniformly integrable martingale additive functional. The decomposition is unique up to equivalence (i.e., $P^h$-a.s. for all $h \in \mathcal{D}$) .

PROOF. The method of the proof is to write $\varphi = \varphi_{1,d} + \varphi_{2,d} + \varphi_r$, where the three terms on the right are class D potentials of $Z_t$ on $\mathcal{D}$, and moreover $\varphi_{1,d}$ corresponds to discontinuities of $\varphi(Z_t)$ at which $Z_t$ is continuous, $\varphi_{2,d}$ corresponds to discontinuities of $\varphi(Z_t)$ with $Z_{t-} \in H - H_0$, and $\varphi_r$ is a regular potential. The asserted decomposition is obtained separately for each of the three terms.

Recalling from Notation 1.4 that $\varphi_-(t)$ is a $Z_t$-previsible process indistinguishable from the left-limit process of $\varphi(Z_t)$, for fixed $\varepsilon > 0$ let

$$T = \inf\{t > 0: (|\Delta\varphi(t)| I_{(Z_{t-} = Z_t)}) \geq \varepsilon\} .$$

Since $\varphi(Z_t)$ is r.c.l.l. except on a null set, its jumps of size $\varepsilon$ do not accumulate, and hence we see that on $\{T < \infty\}$ $\varphi(Z_t)$ has a jump of size at least $\varepsilon$ at $t = T$ where $Z_t$ is continuous. Also, since $T$ is a $Z^o_{t+}$ (hence $Z_t$) stopping time, (and a terminal time), it follows by Theorem 2.13 of Essay 1 that $T$ is $Z^\mu_t$-previsible for each initial distribution $\mu$ . Then by the moderate Markov property

$$(1.12)\qquad E^h(\varphi(Z_T)\,|\,\mathcal{Z}_{T-}) = E^{Z_{T-}}\varphi(Z_0) = \varphi(Z_T)\ ,\quad P^h\text{-a.s. on } \{T<\infty\}\ .$$

Since $\varphi(Z_t)$ is a supermartingale, the optional sampling theorem implies that $\varphi(Z_T) \le \varphi_-(T)$, $P^h$-a.s. on $\{T<\infty\}$ (see [4, VI, Part 1, Theorem 14]).

Letting $T = T_1$ and $T_{n+1} = T_1 \circ \theta^Z_{T_n}$, $1 \le n$, it follows in the same way that $\Delta\varphi(T_n) \le 0$, $P^h$-a.s. on $\{T_n < \infty\}$ for all $n$. Next, by the same supermartingale property we see that

$$E^h\left|\sum_{n=1}\Delta\varphi(T_n)\right| \le E^h(\varphi(h) - \lim_{t\to\infty}\varphi(Z_t)) = \varphi(h)\ .$$

As $\varepsilon \to 0+$, the same facts are seen to hold for all $\varepsilon$. Hence we may introduce the process $A_{1,d}(t) = -\sum_{t_i\le t}\Delta\varphi(t_i)$, the sum being over all $t_i \le t$ with $Z_{t_i} = Z_{t_{i-}}$ and $\Delta\varphi(t_i) < 0$ (in case this yields $A_{1,d}(0+) = \infty$, we set $A_{1,d}(t) = 0$ for all $t$). It is then clear that $A_{1,d}(t)$ is an additive functional, and we have $E^hA_{1,d}(\infty) \le \varphi(h)$ for all $h \in \mathcal{D}$. Moreover, for each $\varepsilon$ the process

$$(1.13)\qquad A_\varepsilon(t) = -\sum_{n:T_n\le t}\Delta\varphi(T_n)$$

is $\mathcal{Z}^h_t$-previsible and equivalent to an additive functional. Therefore $A_{1,d}$ is $\mathcal{Z}^h_t$-previsible (since it is $P^h$-indistinguishable from such, and $\mathcal{Z}^h_t$ contains all $P^h$-null sets).

We now set $\varphi_{1,d}(h) = E^h(A_{1,d}(\infty))$. It is immediately clear that $\varphi_{1,d}$ is a potential of class D, with $\varphi_{1,d} \le \varphi$. Of course, the Doob-Meyer decomposition of $\varphi_{1,d}(Z_t)$ for each $P^h$ is given by

$$\varphi_{1,d}(Z_t) = E^h(A_{1,d}(\infty)\,|\,\mathcal{Z}^h_t) - A_{1,d}(t)\ .$$

According to our construction, $A_{1,d}(\infty)$ is even $\mathcal{Z}^o$-measurable, hence $\varphi_{1,d}$ is Borel. Finally, let $0 < T$ be $\mathcal{Z}_t$-previsible with $Z_{T-} = Z_T$ on $\{T<\infty\}$. Then we have

(1.14) $$\Delta\varphi_{1,d}(T) = E^h(A_{1,d}(\infty)|Z_T) - E^h(A_{1,d}(\infty)|Z_{T-}) - \Delta A_{1,d}(T)$$

$$= -\Delta A_{1,d}(T)\ ,\quad P^h\text{-a.s. on } \{T < \infty\}\ .$$

Consequently, $\varphi_{1,d}(Z_t)$ contains all of the discontinuities at times $T_n$ of $\varphi(Z_t)$, and does not introduce any others of the same kind. Setting $\varphi_2 = \varphi - \varphi_{1,d}$, it follows that $\varphi_2(Z_t)$ has its discontinuity time set contained a.s. in that of $Z_t$.

It is next to be shown that $\varphi_2$ is excessive, hence a potential of class D . Setting $\varphi_\varepsilon(h) = E^h(A_\varepsilon(\infty))$, since $\varphi_\varepsilon(h)$ increases to $\varphi_{1,d}(h)$ as $\varepsilon \to 0+$ it suffices to show that for each $\varepsilon$ and $t > 0$

$$E^h(\varphi(Z_t) - \varphi_\varepsilon(Z_t)) \le \varphi(h) - \varphi_\varepsilon(h)\ ,$$

or again that this holds with $t \wedge T_n$ in place of $t$, for every $n$ . Starting with $n = 1$, we have

(1.15) $$E^h\varphi_\varepsilon(Z_{t\wedge T_1}) - \varphi_\varepsilon(h) =$$

$$E^h(E^h(A_\varepsilon(\infty) - A_\varepsilon(t\wedge T_1)|Z_{t\wedge T_1})) - E^hA_\varepsilon(\infty)$$

$$= -E^hA_\varepsilon(t\wedge T_1)$$

$$= -E^h(\Delta A_\varepsilon(T_1)\ ;\ T_1 \le t)\ .$$

On the other hand, by the previsibility of $t \wedge T_1$ and optional sampling for supermartingales

$$\varphi(h) - E^h\varphi(Z_{(t\wedge T_1)}) = (\varphi(h) - E^h\varphi_-(t\wedge T_1)) - E^h(\Delta\varphi(t\wedge T_1))$$

$$\ge -E^h(\Delta\varphi(t\wedge T_1)\ ;\ T_1 \le t)$$

$$= E^h(\Delta A_\varepsilon(T_1)\ ;\ T_1 \le t)\ .$$

This finishes the case $n = 1$ . Assuming the case $n$ and writing $\varphi_\varepsilon(Z_{t\wedge T_n}) = E^h(A_\varepsilon(\infty) - A_\varepsilon(t\wedge T_n)|Z_{t\wedge T_n})$ it follows similarly that

(1.16)(a) $$E^h(\varphi_\varepsilon(Z_{t\wedge T_{n+1}}) - \varphi_\varepsilon(Z_{t\wedge T_n})\ ;\ t\wedge T_n < t\wedge T_{n+1})$$

$$= E^h(A_\varepsilon(t\wedge T_n) - A_\varepsilon(t\wedge T_{n+1}))$$

$$= -E^h(\Delta A_\varepsilon(T_{n+1})\ ;\ T_{n+1} \le t\ ,\ \text{and}$$

(b) $$\varphi(h) - E^h\varphi(Z_{(t\wedge T_{n+1})})$$

$$= (\varphi(h) - E^h\varphi(Z_{t\wedge T_n})) + E^h(\varphi(Z_{t\wedge T_n}) - \varphi(Z_{t\wedge T_{n+1}}))$$

$$\geq (\varphi_\varepsilon(h) - E^h\varphi_\varepsilon(Z_{t\wedge T_n})) + E^h(\Delta A_\varepsilon(T_{n+1})\ ;\ T_{n+1} \leq t)\ .$$

This proves the case $n + 1$, and hence the assertion. We note that only the previsibility of the $T_{n+1}$ is used here, not the continuity of $Z(t)$ at $t = T_{n+1}$.

We next compensate the accessible jump times of $\varphi_2(Z_t)$. Since these are contained in those of $Z_t$, it follows by Theorem 2.13(ii) of Essay 1 that these are contained in the set of times where $Z_{t-} \in H - H_0$. By taking accessible parts of all the discontinuity times of $Z_t$, it is easy to see that $\{(t,w_Z): Z_{t-} \in H - H_0\}$ is contained in a countable union of graphs of $Z_t^h$-previsible times, $P^h$-a.s. for each $h$. Then by [3, IV, 88 b)] this set is equal to a countable disjoint union of graphs of such times. Let $(T_n)$ denote such a set, and for each $n$ let $(T_{k,n},\ 1 \leq k \leq n)$ be defined by $T_{k,n} = T_j$ on the set where exactly $k$ among $(T_1, \ldots, T_n)$ are less than or equal to $T_j$. Then the $T_{k,n}$ are $Z_t^h$-previsible, and define a natural ordering of $T_1, \ldots, T_n$.

We now set $\varphi_{2-}(t) = \limsup_{Q\ni r\to t-} \varphi_2(Z_r)$, $t > 0$ and (letting $\infty - \infty = 0$) define

$$A_n(t) = \sum_{T_{k,n}\leq t} (\varphi_{2-}(T_{k,n}) - E^{Z_{T_{k,n}-}}\varphi_2(Z_0))\ .$$

Since the last term on the right is a version of $E(\varphi_2(Z_{T_{k,n}})|Z^h_{T_{k,n}-})$, it follows by the supermartingale property that $0 \leq E^h(A_n(t)) \leq E^h\varphi_2(0) - E^h(\varphi_2(t))$, and the $A_n(t)$ are increasing in $n$ for all $t$, and in $t$ for each $n$, $P^h$-a.s. Thus we may define

$$A_{2,d}(t) = \sum_{0<t_i\leq t} (\varphi_{2-}(t) - E^{Z_{t_i-}}\varphi_2(Z_0)) \tag{1.17}$$

where the sum is over all $t_i$ with $Z_{t_i-} \in H - H_0$ and $0 < \varphi_{2-}(t_i) - E^{Z_{t_i-}}\varphi_2(Z_0)$ if this gives $A_{2,d}(0+) = 0$, and $A_{2,d}(t) = 0$ elsewhere. Then $A_{2,d}$ is an additive functional and, setting $\varphi_{2,d}(h) = E^hA_{2,d}(\infty)$, we have $0 \leq \varphi_{2,d}(h) \leq \varphi_2(h)$. Moreover, since $\varphi_2$ is Borel it is easily seen that $A_{2,d}$ is $Z_t^h$-previsible for every $h$.

Then $\varphi_{2,d}$ is a potential of class D, and for each $P^h$ the Doob-Meyer decomposition of $\varphi_{2,d}(Z_t)$ is

$$\varphi_{2,d}(Z_t) = E^h(A_{2,d}(\infty)|\mathcal{Z}_t^h) - A_{2,d}(t) .$$

We recall, now, that a potential $\varphi_r$ of class D is called a regular potential if, for any increasing sequence $T_n$ of $\mathcal{Z}_t$-stopping times, and any $h \in \mathcal{D}$,

$$E^h \lim_{n\to\infty} \varphi_r(Z_{T_n}) = E^h \varphi_r(Z_T)$$

where $T = \lim_{n\to\infty} T_n$ and $\varphi_r(Z_\infty) = 0$. The key fact needed to reduce Theorem 1.8 to standard methods of Markov processes is

LEMMA 1.8. Set $\varphi_r = \varphi_2 - \varphi_{2,d} = \varphi - \varphi_{1,d} - \varphi_{2,d}$. Then $\varphi_r$ is a regular class D potential.

PROOF. We have seen that $\varphi_r \geq 0$, and clearly $\lim_{t\to\infty} E^h \varphi_r(Z_t) = 0$ and $\lim_{t\to 0} \varphi_r(Z_t) = \varphi_r(h)$, $P^h$-a.s. If we show that $E^h \varphi_r(Z_t) \leq \varphi_r(h)$, then $\varphi_r$ is a potential of class D. To this effect, we need to repeat the argument used for $\varphi_2$, and for this we require the analogues of $A_\varepsilon$ and $\varphi_\varepsilon$. Using the same symbols as before, we introduce

$$T = \inf\{t > 0: (\varphi_{2-}(t) - E^{Z_{t-}} \varphi_2(Z_0))I_{(Z_{t-} \in H-H_0)} \geq \varepsilon\} .$$

Since $A_{2,d}(\infty) < \infty$ holds a.s., it is now easy to see that $T > 0$ a.s., and setting $T_1 = T$, $T_{n+1} = T \circ \theta^Z_{T_n}$, we see that $\lim_{n\to\infty} T_n = \infty$ a.s. Thus we may define as before an $A_\varepsilon$ and $\varphi_\varepsilon$ by

$$A_\varepsilon(t) = \sum_{n:T_n \leq t} (\varphi_{2-}(T_n) - E^{Z_{T_n}} \varphi_2(Z_0)) \; ; \quad \varphi_\varepsilon(h) = E^h(A_\varepsilon(\infty)) ,$$

and observe that $A_\varepsilon$ is equivalent to a $\mathcal{Z}_t^h$-previsible additive functional, while $\varphi_\varepsilon(h)$ increases to $\varphi_{2,d}(h)$ as $\varepsilon \to 0+$. It is now easy to check that the proof of (1.15) and (1.16) applies here with $\varphi$ replaced by $\varphi_2$, showing that $E^h\varphi_r(Z_t) \leq \varphi_r(h)$.

Finally, let us prove the regularity. Let $T_n$ be any sequence of $\mathcal{Z}_t$-stopping times increasing to $T \leq \infty$. Then clearly $\lim_{n\to\infty} E^h\{\varphi_r(Z_{T_n}); T = \infty\} = 0$. On the other hand, over $\{T < \infty\}$ there is no difficulty in passing to the limit on $\{T_n = T$ for large $n\}$. Then setting $S = \{T_n < T < \infty,$ all $n\}$ we have

$$(1.18)\quad \lim_{n\to\infty} E^h(\varphi_r(Z_{T_n}) - \varphi_r(Z_T);\ S)$$

$$= \lim_{n\to\infty} E^h(\varphi_r(Z_{T_n}) - \varphi_r(Z_T);\ \{Z_{T-} = Z_T\} \cap S) +$$

$$+ \lim_{n\to\infty} E^h(\varphi_r(Z_{T_n}) - \varphi_r(Z_T);\ \{Z_{T-} \in H - H_0\} \cap S)$$

$$= E^h(\Delta\varphi(T) - \Delta\varphi_{1,d}(T) - \Delta\varphi_{2,d}(T);\ \{Z_{T-} = Z_T\} \cap S)$$

$$+ E^h(\Delta\varphi(T) - \Delta\varphi_{1,d}(T) - \Delta\varphi_{2,d}(T);\ \{Z_{T-} \in H - H_0\} \cap S)\ .$$

But on $\{Z_{T-} = Z_T\} \cap S$ we have by (1.14) $\Delta\varphi(T) = -\Delta A_{1,d}(T) = \Delta\varphi_{1,d}(T)$, while by (1.17) and the martingale property of $E^h(A_{2,d}(\infty)\,|\,Z_t^h)$ we have $E^h(\Delta\varphi_{2,d}(T);\ \{Z_{T-} = Z_T\} \cap S) = 0$ . Hence the first summand on the right vanishes. As for the second, on $\{Z_{T-} \in H - H_0\}$ we have $\Delta A_{1,d}(T) = 0$, and therefore $E^h(\Delta\varphi_{1,d}(T);\ \{Z_{T-} \in H - H_0\} \cap S) = 0$, while by (1.17) and the moderate Markov property

$$E^h(\Delta\varphi\ (T) - \Delta\varphi_{2,d}(T);\ \{Z_{T-} \in H - H_0\} \cap S)$$

$$= E^h(\Delta\varphi\ (T) + \Delta A_{2,d}(T);\ \{Z_{T-} \in H - H_0\} \cap S)$$

$$= E^h(\varphi_2(Z_T) - E^{Z_{T-}}\varphi_2(Z_0);\ \{Z_{T-} \in H - H_0\} \cap S)$$

$$= 0\ .$$

This completes the proof of Lemma 1.8.

As mentioned at the beginning, the rest of the proof has already been done in a somewhat different context in [2, IV, Section 3]. It follows that there is a continuous increasing additive functional $Z_c(t)$ with $\varphi_r(z) = E^z A_c(\infty)$ . The method used is that of Sur [15], together with refinements which reduce the problem to a bounded regular potential. The proof is unfortunately not short. Some simplifications can be made because the multiplicative functionals $M_t$ of [2] are not present, and hence S and the $S_p$'s of [2] are absent, but it does not seem merited to rewrite the proof. In the present case there may be branching points, but it can be checked that the proof in [2] makes no use of quasi-left continuity of $X_t$ and so applies also to Borel right processes (see Getoor [6, 9.] for the relevant information on hitting times and excessive functions.)

In the extension argument of [2, p. 168] from bounded to unbounded potentials, use is made of the uniqueness theorem to the effect that $Z_c(t)$ is uniquely determined by its potential ([3, IV, (2.13)]). However, this is easy to see directly from martingale arguments. Thus if $A_c^1$ and $A_c^2$ are continuous additive functionals with $\varphi_r(z) = E^z A_c^1(\infty) = E^z A_c^2(\infty)$, then for each $h$ the identity

$$\varphi_r(Z_t) = E^h(A_c^1(\infty)|\mathcal{Z}_t) - A_c^1(t)$$

$$= E^h(A_c^2(\infty)|\mathcal{Z}_t) - A_c^2(t)$$

implies that $A_c^1(t) - A_c^2(t)$ is a continuous, $(P^h, \mathcal{Z}_t)$-martingale of bounded variation. By arguments given in the proof of Theorem 1.6 (since the martingale may be stopped at any $N$) this implies that $A_c^1 - A_c^2$ is indistinguishable from the zero martingale. But the same reasoning applies if $A_c^1$ and $A_c^2$ are only assumed to be $\mathcal{Z}_t^h$-previsible. Indeed, a previsible right-continuous martingale $M_t$ is continuous: otherwise, since $M_t - M_{t-}$ is a previsible process, a bounded previsible stopping time $T$ could be found with $P^h\{M_T \neq M_{T-}\} > 0$, leading to immediate contradiction with the fact that $M_T$ is $\mathcal{Z}_{T-}^h$-measurable ([3, IV, 67]). It follows that any decomposition of the type asserted by Theorem 1.8 is unique up to equivalence. Hence we must have

$$A(t) = A_{1,d}(t) + A_{2,d}(t) + A_c(t) ,$$

$$M(t) = \varphi(Z_t) - \varphi(Z_0) + A(t) ,$$

and the proof is complete.

## 2. TRANSITION TO THE INITIAL SETTING: THE LÉVY SYSTEM OF A PROCESS.

In order to translate results back and forth between the prediction process setting and their original setting, it is useful to examine more carefully the connection of $(\Omega, G^o)$ and $(\Omega_Z, \mathcal{Z}^o)$. Since the connections we have in mind are completely general, not requiring any restriction on the probabilities, we return <u>temporarily</u> to the notations of Essay 1, Section 2. Thus $\Omega_Z$ is the space of all paths $w_Z(t) \in H_0$ which are right-continuous with left limits in $H$ for all $t > 0$, $\varphi(h)$ is the function of Definition 1.8 (rather than only its first coordinate) and $\varphi_-(t)$ denotes coordinatewise $\limsup_{Q \ni r \to t-} \varphi(Z_r)$. As remarked at the end of Essay 1, $(\Omega_Z, \mathcal{Z}^o)$ may be topologized as a coanalytic subset of a Lusin space. While neither this nor the following assertion is essential to the development here, it is worthwhile to have them on record.

PROPOSITION 2.1. Let $\Omega_\varphi = \{w_Z\colon \varphi w_Z(t)$ is r.c.l.l. in $\bar{R}_\infty$ with the product topology$\}$, where $\varphi w_Z(t) = \varphi(w_Z(t))$, $t \geq 0$. Then we have $\Omega_\varphi \in Z^o$, and $\varphi(\Omega_\varphi) = \{$all r.c.l.l. paths in $\bar{R}_\infty\}$.

PROOF. Since $\varphi$ is a Borel function, the components of $\varphi w_Z(t)$ are $Z^o_t$-progressively measurable. We now apply the results of [3, IV, 17], according to which the two processes defined as the right limsup and liminf of $\varphi(w_Z(r))$ along rational $r \geq t$, $r \downarrow t$, are $Z^o_{t+}$-progressively-measurable, and the two processes defined as the left limsup and liminf along rational $r < t$, $r \uparrow t$, are $Z^o_t$-progressively-measurable in $t > 0$. The condition $w_Z \in \Omega_\varphi$ is simply that the two right-limit processes should equal $\varphi(w_Z(t))$ and the two left-limit processes should equal each other. Since $\varphi(w_Z(t))$ is also $Z^o_{t+}$-progressively-measurable, these conditions define a set in $Z^o$.

To see that $\varphi(\Omega_\varphi) = \{$all r.c.l.l. paths$\}$, we fix $w \in \Omega$ and let $h_w$ be the unit probability at $w$. Then $h_w \in H$, and so we can define its prediction process $Z_t^{h_w}$ as in Essay 1, Section 1. By Theorem 1.9 there, we have $P^{h_w}\{\varphi(Z_t^{h_w}) = X_t$ for all $t\} = 1$, meaning in the present case that the even coordinates $w_{2n}(t)$ of $w$ are identical with those of $\varphi(Z_t^{h_w})$ at $w$. Since $Z_{(\cdot)}^{h_w}$ at $w$ defines an element of $\Omega_Z$, and any r.c.l.l. path $X_t$ is obtained as $X_t = (w_{2n}(t))$ from a $w \in \Omega$, the assertion follows.

The mapping $\varphi$ on $\Omega_\varphi$ is not one-to-one. In fact, since $P^h(\Omega_\varphi) = 1$ for every $h \in H$ (as usual, we use the same notation $P^h$ for measures on $\Omega_Z$ or on $\Omega$) if $\varphi$ were one-to-one then the prediction process $Z^h$ on $\Omega$ would not depend on $h$ except for null sets, which is absurd. Thus we cannot use $\varphi$ on $\Omega_\varphi$ to transfer a process on $\Omega$ to one on $\Omega_Z$. Instead, we must reduce $\Omega_\varphi$ to a subset depending on $h$. Thus, given $h$ and a particular choice of $Z^h_{(\cdot)}$ on $(\Omega, G^o)$, we can regard $Z^h_{(\cdot)}$ as a measurable mapping of $(\Omega, F^o) \longrightarrow (\Omega_Z, Z^o)$. Then the set $\Omega_h = \{w \in \Omega\colon \varphi Z^h_{(\cdot)}(w) = w\}$ is in $F^o$, and we have $Z^h_{(\cdot)}(\Omega_h) \subset \Omega$, $Z^h_{(\cdot)}(\Omega_h) \in Z^o$, and $\varphi$ is one-to-one on $Z^h_{(\cdot)}(\Omega_h)$. Also $P^h(\Omega_h) = 1$, and hence we can use $\varphi$ to transfer objects from $\Omega$ to $\Omega_Z$, except for an $h$-null set.

In the cases of interest here the problem is to go in the other direction, from $\Omega_Z$ to $\Omega$, and this presents almost no difficulty. Thus we now define the concept of a Lévy system for any $h \in H$, and obtain its existence and properties from Theorem and Corollary 1.3.

DEFINITION 2.2 A Lévy system for $h$ consists of kernels $N_Z$ and $M_Z$ on $H_0$, (as in (1.5) with $H(D_A) \cap H_0 = H_0$), and $F_t^h$-previsible, increasing processes $H_{h,s}$ and $L_{h,s}$ on $\Omega$, with $H_{h,0} = L_{h,0} = 0$, where $H_{h,s}$ is continuous, $L_{h,s}$ is pure-jump, $E^h H_{h,s} < \infty$, and $E^h L_{h,s} < \infty$, such that for $0 \le f(z_1, z_2) \in H \times H$, $f(z,z) = 0$, we have

(2.1)(a)
$$E^h \sum_{0<s\le t} f(Z_{s-}^h, Z_s^h) I_{\{Z_s^h \in H_0\}}$$
$$= E^h \int_0^t dH_{h,s} \int_{H_0} N_Z(Z_{s-}^h, dz) f(Z_{s-}^h, z)$$

(b)
$$E^h \sum_{0<s\le t} f(Z_{s-}^h, Z_s^h) I_{\{Z_{s-}^h \in H - H_0\}}$$
$$= E^h \int_0^t dL_{h,s} \int_{H_0} M_Z(Z_{s-}^h, dz) f(Z_{s-}^h, z)$$

(c)
$$E^h \sum_{0<s\le t} f(Z_{s-}^h, Z_s^h)$$
$$= E^h \int_0^t d\bar{H}_{h,s} \int_{H_0} \bar{N}_Z(Z_{s-}^h, dz) f(Z_{s-}^h, z)$$

where $\bar{H}_h = H_h + L_h$, and $\bar{N}_Z = N_Z + M_Z$ .

DISCUSSION. Obviously a Lévy system for $h$ does not depend on the particular choice of $Z^h$ . Furthermore, when a Lévy system exists then the proof of Corollary 1.3 carries over without difficulty to show that, for any $F_t^h$-previsible process $y_t$ with values in $(\bar{R}_\infty, \bar{B}_\infty)$, (2.1) remains true if $f(Z_{t-}^h, z)$ is replaced by $f(y_t, z) I_{(Z_{t-}^h \neq z)}$ for $0 \le f \in \bar{B}_\infty \times H$ . In particular, for $y_t = X_{t-}$ we may replace this $f(y_t, z)$ by $f(y_t, \varphi(z)) I_{(y_t \neq \varphi(z))}$ to obtain the following.

(2.1)(d) For $0 \le f(x_1, x_2) \in \bar{B}_\infty \times \bar{B}_\infty$, with $f(x,x) = 0$, we have

$$E^h \sum_{0<s\le t} f(X_{s-}, X_s) I_{(Z_{s-}^h \neq Z_s^h)}$$
$$= E^h \int_0^t d\bar{H}_{h,s} \int_{H_0} \bar{N}_Z (Z_{s-}^h, dz) f(X_{s-}, \varphi(z)) .$$

Analogous statements also hold corresponding to (2.1)(a) and (b). In other words, the Lévy system compensates the jumps of $X_t$ which coincide in time with jumps of $Z_t^h$ . On the other hand, this is the most that could be expected. By Essay I, Theorem 2.13(i), a time of the form

$$T = \inf\{t > 0: |\Delta\varphi_t| I_{(Z_{t-} = Z_t)} \geq \varepsilon\}$$

on $\Omega_Z$ is $\mathcal{Z}_t^h$-previsible, where $|\Delta\varphi_t|$ denotes the magnitude of $\varphi(Z_t) - \varphi_-(t)$ in any convenient Borel metric on $\mathsf{X}_{n=1}^{\infty} \bar{R}$ . By the moderate Markov property, we then have $\varphi(Z_T) \in \mathcal{Z}_{T-}^h$ over $\{0 < T < \infty\}$ . By iteration of $T$, and letting $\varepsilon \to 0$, it follows in the usual way that the processes $\sum_{0<s\leq t} f(\varphi_-(s), \varphi(Z_s)) I_{(Z_{s-} = Z_s)}$ are $\mathcal{Z}_t^h$-previsible (where $f(x,x) = 0$ as before). Therefore, these processes are their own compensators, and the Lévy system is irrelevant. It will be seen easily from the proof to follow that this fact translates into the $\mathcal{F}_t^h$-previsibility of $\sum_{0<s\leq t} f(X_{s-}, X_s) I_{(Z_{s-}^h = Z_s^h)}$, hence there is no need to compensate these discontinuities.

THEOREM 2.3. A Lévy system exists for any probability $h$ on $(\Omega, \mathcal{G}^o)$ .

PROOF. All that needs to be done is to define from Theorem 1.2 (with $H_A = H$)

$$H_{h,t}(w) = H_{Z,t}(Z_{(\cdot)}^h(w)) \ ,$$

$$L_{h,t}(w) = L_{Z,t}(Z_{(\cdot)}^h(w)) \ ,$$

and to show that they are $\mathcal{F}_t^h$-previsible processes. More generally, let us show that if $Y_t$ is any real-valued, $\mathcal{Z}_t^h$-previsible process with $P^h\{Y_0 = 0\} = -$, then $Y_t(Z_{(\cdot)}^h(w))$ is $\mathcal{F}_t^h$-previsible. Since, by definition the previsible $\sigma$-field is generated by the left-continuous adapted processes (if we take $\mathcal{F}_{0-}^h = \mathcal{F}_0^h$; see [3, IV, 61]), and this class is closed under linear and lattice operations, it will suffice to consider the case of left-continuous $Y_t$ (without assuming $P^h\{Y_0 = 0\} = 1$) . Then $Y_t(Z_{(\cdot)}^h(w))$ is left-continuous, and we need only show it is $\mathcal{F}_t^h$-adapted, or again, that for $S \in \mathcal{Z}_t^h$ we have $\{w: Z_{(\cdot)}^h(w) \in S\} \in \mathcal{F}_t^h$ . Let $S_0 \in \mathcal{Z}_t$ be such that $P^h(S_0 \Delta S) = 0$ . Then because $Z_t^h$ is $\mathcal{F}_{t+}^o$-progressively measurable, it follows that $\{w: Z_{(\cdot)}^h(w) \in S_0\} \in \mathcal{F}_{t+}^o$ . Also, by definition of $P^h$ on $\mathcal{Z}$ we have

$$\begin{aligned} &P^h\ w:\{Z_{(\cdot)}^h(w) \in S_0 \Delta S\} \\ &= P^h\{S_0 \ \Delta S\} \\ &= 0 \ . \end{aligned}$$

Consequently, $\{w: Z^h_{(\cdot)}(w) \in S\} \in F^h_{t+}$, and so $Y_t(Z^h_{(\cdot)}(w))$ is $F^h_{t+}$-adapted. But by left-continuity this is equivalent to $F^h_t$-adaptedness for $t > 0$, and at $t = 0$ we have the extra hypothesis needed to complete the proof.

In the more specialized contexts of martingales or class D supermartingales, we have the decomposition results corresponding to Theorems 1.6 and 1.8. Here we return to the notation of Section 1: $\varphi(z)$ and $X_t$ denote only their first coordinates, while $F^o_t$ and $G^o_t$ coincide.

THEOREM 2.4. Let P be any $G^o_{t+}$-square-integrable martingale probability for $X_t = w_2(t)$ on $(\Omega, G^o)$. Then there is a decomposition $X_t = X^c_t + X^d_t$ where $X^c_t$ is a continuous, $F^P_{t+}$-square-integrable martingale with $X^c_0 = 0$, $X^d_t$ is a square-integrable martingale which is a P-mean-square limit of martingales of bounded variation, and $(X^c_t X^d_t)$ is a martingale. The decomposition is unique up to a P-null set.

PROOF. Choosing $Z^P_t$ as any version of the prediction process of P on $\Omega$, we first write $X^c_t(w) = M_c(t, Z^P_{(\cdot)}(w))$, and $X^d_t(w) = X_0 + M_d(t, Z^P_{(\cdot)}(w))$, where $M_c$ and $M_d$ are the martingales of Theorem 1.6 on $\Omega_M$. These definitions are meaningful except for a P-null set since $P\{Z^P_{(\cdot)} \in \Omega_M\} = 1$. Moreover, $X^c_t$ and $X^d_t$ are right-continuous, and by arguing just as in the proof of Theorem 2.3 it follows that they are $F^h_{t+}$-adapted. Further, we recall from the Remark to Theorem 1.9 of Essay 1 that $F^P_{t+}$ is P-equivalent to the $\sigma$-field generated by $\{Z^P_s, s \le t\}$. Thus, since $M_c$ is a P-martingale,

$$(2.2) \qquad \begin{aligned} & E^P(X^c_{s+t} | F^P_{t+}) \\ &= E^P(M_c(s+t, Z^P_{(\cdot)}(w)) | \sigma\{Z^P_s, s \le t\}) \\ &= M_c(t, Z^P_{(\cdot)}(w)) \\ &= X^c_t, \qquad \text{P-a.s.} \end{aligned}$$

The same reasoning applies to $X^d_t$, and to $X^c_t X^d_t$. Finally, the mean-square approximation obviously transfers from Theorem 1.6 in the same way, and the uniqueness proof was already formulated for fixed P. We have only to apply it to $X_t - X_0 = X^c_t + (X^d_t - X^d_0)$ to complete the derivation.

Turning to the specialization of Theorem 1.8, we obtain the Doob-Meyer decomposition of a class D, right-continuous submartingale $(Y_t, G^o_{t+})$ by writing $Y_t = E(Y_\infty | G^o_{t+}) - (E(Y_\infty | G^o_{t+}) - Y_t)$, and noting that the last term is a class D potential. Then we have only to decompose the last term, as follows.

THEOREM 2.5. Let $P$ on $(\Omega, G^o)$ be such that $X_t$ is a class D potential. Then there is a P-a.s. unique decomposition $X_t = M_t - A_t$, where $M_t$ is a $G^P_{t+}$-uniformly integrable, right-continuous martingale, and $A_t$ is a $G^P_t$-previsible increasing process with $A_0 = 0$.

NOTE. Unlike the decomposition of Theorem 2.4, the present components depend on the choice of $\sigma$-fields, and may be altered if one replaces them by the $\sigma$-fields generated by $X_t$. Hence we denote them by $G^P_{t+}$ instead of $F^P_{t+}$, although in the present notation these are actually identical.

PROOF. We recall again that the uniqueness part of the proof of Theorem 1.8 was in no way Markovian, and applies here without change. For the existence, we set

$$M_t(w) = M(t, Z^P_{(\cdot)}(w))$$

and

$$A_t(w) = A(t, Z^P_{(\cdot)}(w)) ,$$

where $A$ and $M$ on the right are from Theorem 1.8, and $Z^P_{(\cdot)}$ is any fixed choice of the prediction process of $P$ on $\Omega$. Since $A_0(w) = 0$, the proof of Theorem 2.3 shows that $A_t$ is $G^P_{t+}$-previsible. Of course, since $E^P A_\infty < \infty$ it is clear that $M_t$ is uniformly integrable, hence it need only be shown that it is a $G^P_{t+}$-martingale. However, this follows by the same reasoning as (2.2), completing the proof.

## 3. ON CONTINUOUS LOCAL MARTINGALES.

We pass over any examination of general local martingales or semi-martingales. These are treated at length in [4], and it is not clear whether our Markovian approach has anything to add. The continuous local martingales provide not only a simpler application, but also one in which the method of time changes can be aptly illustrated in a prediction process setting. In point of detail, we avoid the "adjoined Brownian motions" of the usual time-change result (as in H. Kunita, S. Watanabe [10], for example). In the last part of the section, we specialize further to the continuous local martingales which are autonomous germ-Markov processes, as defined in [9]. These generalize the one-dimensional diffusions in natural scale, and perhaps should be called germ-diffusion processes in natural scale. However, this would be misleading in that no reduction of the general germ-diffusion to a scale in which it is a local martingale is possible.

We continue with the notations of Section 1: $X_t(w) = w_2(t) \equiv \varphi(Z^P_t)$, etc., but our starting point is the prediction space of all continuous local martingales. We recall that $A_t$ is a "continuous local martingale" means that if $T_N = \inf\{t: |A_t| > N\}$ then $A_{t \wedge T_N}$ is a continuous martingale for every $N > 0$. Actually, we will deal somewhat more generally with processes which are continuous local martingales given their initial values. Thus we do not require that the initial value have finite expectation.

PROPOSITION 3.1. Let $L_c \subset H$ be the set of all probabilities $h$ on $(\Omega, G^o)$ such that $(\varphi(Z_t) - \varphi(Z_0), Z_t)$ is a $P^h$-continuous local martingale. Then $L_c$ is a complete Borel packet of the prediction process.

NOTE: It is assumed that, $P^h$ = a.s., $\varphi(Z_0) \neq \pm\infty$.

PROOF. In the first place, since $\varphi(Z_t)$ is r.c.l.l. for any $h \in H$, $P^h$-a.s. the condition that $\varphi(Z_t)$ be continuous is the same as $\varphi(Z_t) - \varphi_-(t) = 0$. As seen in the proof of Proposition 2.1, this is a Borel condition on $h$. By the usual optional section theorem argument this implies that $\{h: \varphi(Z_t)$ is $P^h$-a.s. continuous$\}$ is a Borel packet, and the moderate Markov property plus the previsible section theorem show that this packet is complete. In the definition of local martingale, we can redefine $T_N$ by $T_N = \inf\{r \in Q: |\varphi(Z_r)| > N\}$, which is $Z^o$-measurable. Then we see that $\varphi(Z_{t \wedge T_N}) - \varphi(Z_0)$ is $Z_t^o$-adapted, and in conjunction with the continuity and boundedness of $\varphi(Z_{t \wedge T_N})$ the martingale condition becomes $E^{Z_{r_1}} \varphi(Z_{r_2 \wedge T_N}) = \varphi(Z_{r_1})$, $P^h$-a.s. over $\{r_1 < T_N\}$, for $0 < r_1, r_2 \in Q$. Hence the set of continuous local martingale probabilities for $\varphi(Z_t) - \varphi(Z_0)$ is in $H$.

To show that it is a packet, for $Z_t^o$-optional $T < \infty$ we can replace the martingale condition by

$$E^{Z_{r_1}} \varphi(Z_{r_2 \wedge T_N \wedge T}) = \varphi(Z_{r_1})$$

on $\{r_1 < T_N \wedge T\}$, $0 < r_1 \in Q$, together with the conditions $E^{Z_{T+r_3}} \varphi(Z_{r_2 \wedge T_N}) = \varphi(Z_{T+r_3})$, $0 < r_3 \in Q$, on $\{T + r_3 < T_N\}$, since these imply that over $\{T \le r_1 < T_N\}$ one has

$$E^{Z_{r_1}} \varphi(Z_{r_2 \wedge T_N}) = \lim_{r_3 \to (T-r_1)+} E^{Z_{T+r_3}} \varphi(Z_{r_2 \wedge T_N})$$

$$= \lim_{r_3 \to (T-r_1)+} \varphi(Z_{T+r_3})$$

$$= \varphi(Z_{r_1}) ,$$

using Hunt's Lemma for conditional expectations in the first equality. Then it follows immediately that $Z_T$ is $P^h$-a.s. in $L_c$ along with $h$. By the optional section theorem this implies that $L_c$ is a Borel packet. The completeness then follows because, for previsible $0 < T < \infty$,

$$P^h\{Z_{T-} \in L_c\}$$

$$= E^h(P^{Z_{T-}}\{Z_0 \in L_c\})$$

$$= P^h\{Z_T \in L_c\}$$

$$= 1 .$$

We require the following lemma, which was first proved for Hunt processes and square-integrable martingales by H. Kunita and S. Watanabe [10]. But our situation is different, and we prefer to use again the argument of M. G. Sur (see [2, Chapter 4, Section 3]). From now on, we denote $\varphi(Z_t) - \varphi(Z_0)$ by $A_t$.

LEMMA 3.2. There is a unique continuous (non-decreasing) additive functional $\tau(t)$ of $Z_t$ on $\Omega_{L_c}$ such that, for $h \in L_c$, $A_t^2 - \tau(t)$ is a continuous local martingale.

PROOF. For each $h \in L_c$ and $N, A^2_{t \wedge T_N}$ is a bounded submartingale, and by Theorem 2.5 it has a Doob-Meyer decomposition $A^2_{t \wedge T_N} = E^h(A^2_{T_N} | Z_t) - E^h(\tau_N(\infty) | Z_t) + \tau_N(t)$, for an increasing previsible process $\tau_N(t)$, depending on $h$, with $\tau_N(0) = 0$. Thus $A^2_{t \wedge T_N} - \tau_N(t)$ is a uniformly integrable martingale, and since $A^2_{t \wedge T_N}$ is continuous while $\tau_N(t)$ is previsible, it follows easily that $\tau_N(t)$ is continuous, and of course $\tau_N$ is a.s. constant for $t \geq T_N$ along with $A_{t \wedge T_N}$. Furthermore, $\tau_N(\cdot)$ is unique up to a $P^h$-null set, hence we can define $\tau_h(t) = \limsup_{N \to \infty} \tau_N(t)$ to obtain a $Z_t$-adapted, $P^h$-a.s. continuous, non-decreasing process such that $A_t^2 - \tau_h(t)$ is a continuous, $P^h$-local-martingale.

It now remains only to modify the $\tau_h$ to obtain an additive functional $\tau(t)$, but the argument we follow would also suffice to define $\tau(t)$ from scratch. We first observe that $T_N$ is a terminal time of $Z_t$

on $\Omega_{L_c}$. Indeed, since $\varphi(Z_t)$ is a.s. continuous and $\varphi$ is Borel, we may set $T_N = \inf\{r \in Q: \varphi(Z_r) \in B_N\}$ where $B_N$ is the Borel set $B_N = \{z \in L_c: |\varphi(z)| > N\}$. Then $T_N$ is $Z^o$-measurable, and the process on $E_N = L_c \cap H_0 \cap \{h: P^h(T_N > 0) = 1\}$ defined by killing $Z_t$ at $T_N$ is a Borel right process. Thus we define

$$Z_t^N = \begin{cases} Z_t & \text{for } t \le T_N, \text{ and} \\ \Delta & \text{for } t > T_N \end{cases}$$

$P^\Delta(Z_t^N = \Delta \text{ for all } t) = 1$. Then with the Borel transition function derived from that of $Z_t$, $Z_t^N$ becomes a right process on $E_N \cup \Delta$.

We show next that if $e_N(h) = E^h \tau_h(T_N)$, $h \in E_N$, and $e_N(\Delta) = 0$, then $e_N$ is a bounded regular potential for $Z_t^N$. Indeed, by the optional stopping theorem we have $e_N(h) = E^h A^2_{T_N}$, $h \in E_N$. Next we note that there is no difficulty extending the additive functional property of $A_t$ to stopping times. Then for any stopping time $S \le T_N$ we have for $h \in E_N$,

$$E^h(A^2_{T_N}) = E^h(A_S + A_{T_N} \circ \theta^Z_S)^2 \tag{3.1}$$

$$= E^h(A_S^2; S<T_N) + 2E^h(A_S E^{Z_S} A_{T_N}; S<T_N)$$

$$+ E^h(E^{Z_S} A^2_{T_N}; S<T_N) + E^h(A^2_{T_N}; S \ge T_N)$$

$$= E^H(A^2_{S \wedge T_N}) + E^h(E^{Z_S} A^2_{T_N}; S<T_N) .$$

Setting first $S = t \wedge T_N$, we obtain

$$E^h(E^{Z_t^N} A^2_{T_N}) = E^h(E^{Z_t} A^2_{T_N}; t<T_N)$$

$$= E^h(A^2_{T_N}) - E^h(A^2_{t \wedge T_N}) .$$

Since $E^h A^2_{t \wedge T_N}$ decreases to $0$ as $t \to 0+$, it follows that $e_N(h)$ is an excessive function for $Z_t^N$. On the other hand, since $A^2_{t \wedge T_N}$ is a bounded submartingale we have $\lim_{t \to \infty} A^2_{t \wedge T_N} = A^2_{T_N}$, $P^h$-a.s. (or more precisely, the right side is defined by the left on $\{T_N = \infty\})$ . Thus by dominated convergence we have $\lim_{t\to\infty} E^h(E^{Z_t^N} A^2_{T_N}) = 0$, and so $e_N$ is a bounded potential. Finally, let $S_n \le T_N$ increase to a limit $T$ . Then by (3.1) with $S = S_n$ we have

$$\lim_{n\to\infty} E^h e_N(Z^N_{S_n}) - E^h e_N(Z^N_T)$$

$$= \lim_{n\to\infty} E^h(A^2_{T \wedge T_N} - A^2_{S_n \wedge T_N})$$

$$= 0 ,$$

proving that $e_N$ is a regular potential.

It follows by the argument of Sur [15] that there exists a continuous additive functional $\tau_N(t)$ of $Z_t^N$ with potential $e_N$ . Of course, for each $h \in E_N$ this $\tau_N(t)$ is $P^h$-equivalent to the one obtained by the Doob-Meyer decomposition since their difference is a continuous martingale of bounded variation. Now we have $\tau_N(s+t) = \tau_N(t) + \tau_N(s) \circ \theta_t^Z$, $P^h$-a.s. on $\{s+t<T_N\}$ . To complete the proof we need only let $N \to \infty$ and note that $T_N \to \infty$, $P^h$-a.s. for $h \in L_c$ .

It is well-known that $A(t)$ may be reduced in some sense to a Brownian motion by the time change inverse to $\tau(t)$ (see for example H. Kunita and S. Watanabe [10], Theorem 3.1). We wish to formulate a result of this type in the present context, and it will be useful to make a slight enlargement of the $\sigma$-fields $Z_t$ to cope with the case when $P^h\{\lim_{t\to\infty} \tau(t) < \infty\} > 0$ .

LEMMA 3.3. Let $M = \sup_t \tau(t)$, and $T = \inf\{t: \tau(t) = M\}$, where $M$ and $T$ are permitted the value $+\infty$ . Let $Z_t^*$ denote $Z_t \vee \sigma(TI_{\{T\le t\}})$, that is, the $\sigma$-field generated jointly by $Z_t$, the atom $\{T>t\}$, and the trace of $\sigma(T)$ on $\{T\le t\}$ . The family $Z_t^*$ is non-decreasing, and for $h \in L_c$ both $A(t)$ and $A^2(t) - \tau(t)$ are $P^h$-continuous local martingales relative to $Z_t^*$ .

NOTE: $T$ is not a stopping time of $Z_t$, but it is not hard to see that $Z_t^*$ is right-continuous in $t$ .

PROOF. A familiar argument using Jensen's Inequality (as for (1.10)) shows that for $0 < r_1 < r_2 \in Q$,

$$P^h\{\tau(r_1) = \tau(r_2)\} = P^h\{\tau(r_1) = \tau(r_2);\ A(r_1) = A(r_2)\}$$

whence we obtain without difficulty that $P^h\{A(t) = A(T)$ for all $t \ge T\} = 1$ .

Next, observe that any $S \in \mathcal{Z}_t^*$ may be written in the form $S = (S_1 \cap \{T \le t\}) \cup (S_2 \cap \{T > t\})$ with $S_i \in \mathcal{Z}_t$, $i = 1$ or $2$. Then for $s > 0$, with $T_N = \inf t\colon |\varphi(Z_t)| > N\}$ we have trivially

$$E^z(A((s+t) \wedge T_N);\ S)$$

$$= E^z(A(t \wedge T_N);\ S_1 \cap \{T \le t\}) + E^z(A((s+t) \wedge T_N);\ S_2 \cap \{T > t\}) .$$

But the last term on the right becomes

$$E^z(A((t+s) \wedge T_N);\ S_2) - E^z(A((t+s) \wedge T_N);\ S_2 \cap \{T \le t\})$$

$$= E^z(A(t \wedge T_N);\ S_2 \cap \{T > t\}) ,$$

by the martingale property and the same reasoning as before. Adding the two terms yields the local martingale property of $A(t)$ relative to $\mathcal{Z}_t^*$. The case of $A^2(t) - \tau(t)$ is clearly analogous.

This is the key step; the rest is somewhat routine and we will omit some details. Set $\tau^{-1}(t) = \inf\{s\colon \tau(s) > t\}$ with $\inf(\phi) = \infty$. A routine check shows that $\tau^{-1}(t) \wedge T$ is a stopping time of $\mathcal{Z}_t^*$. Let $\mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$ denote the usual indicated $\sigma$-fields, thus $S \in \mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$ means that for $c \le \infty$ $S \cap \{\tau^{-1}(t) \wedge T \le c\} \in \mathcal{Z}_c^*$. Then we have

$$\{M \le d\} = \{\tau^{-1}(d) = \infty\} = \{\tau^{-1}(d) \wedge T = T\} ,$$

from which it follows easily that $M$ is a stopping time $\mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$.

The theorem we wish to prove is as follows.

THEOREM 3.4. For $h \in L_c$, the process $B(t \wedge M) = A(\tau^{-1}(t) \wedge T)$ is a Brownian motion adapted to $\mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$, stopped at time $M$. The times $\tau(s)$ are stopping times of $\mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$, and $A(t) = B(\tau(t) \wedge M)$ for all $t$, $P^h$-a.s.

REMARK. It is a simple matter to see that $B(t \wedge M)$ remains constant for $t \ge M$, so that our notation is consistent. It also would not be difficult to adjoin an auxiliary independent Brownian motion and continue $B(t \wedge M)$ beyond time $M$ as an unstopped Brownian motion (as in [10]) but since $M$ is a stopping time the meaning is clear without this step.

PROOF. The adaptedness and measurability assertions are again routine, and left to the reader. Since $A(\tau^{-1}(\tau(t))) = A(t) = A(t \wedge T)$, where $\tau^{-1}(\tau(t)) = \infty$ for $t \ge T$, the last assertion is clear.

By a characterization theorem of J. L. Doob ([5, VII, Theorem 11.9]) to show that $B(t \wedge M)$ is stopped Brownian motion relative to $\mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$ becomes equivalent to showing that both $B(t \wedge M)$ and $B^2(t \wedge M) - (t \wedge M)$ are martingales relative to $\mathcal{Z}^*_{\tau^{-1}(t) \wedge T}$. By Lemma 3.3 and the optional sampling theorem, they are plainly local martingales. Since $t \wedge M$ is bounded by $t$, the second is then clearly a martingale. Thus $E^h B^2(t \wedge M)$ is finite, hence so is $E^h \sup_{s \le t} |B(s \wedge M)|$. Then by dominated convergence as $N \to \infty$ we obtain the martingale property of the local martingale $B(t \wedge M)$.

While Theorem 3.4 provides a rough outline of the process $A(t)$, it conceals a variety of possibilities which emerges only when we introduce further assumptions. For convenience, we let $\mathcal{L}_d$ denote a complete Borel $H_0$-subpacket of $\mathcal{L}_c$ such that, for $z \in \mathcal{L}_d$, $P^z\{M = \infty\} = 1$. Then it is clear from the theorem that $A(t)$ is unbounded above and below, $P^z$-a.s. for $z \in \mathcal{L}_d$. If we assume that $(\varphi(Z_t), \mathcal{Z}_t)$ is a homogeneous strong-Markov process, then it follows from well-known facts that it must be for each $z$ a regular diffusion in the natural scale on $(-\infty,\infty)$. Then the representation of Theorem 3.4 becomes $B(t) = A(\tau^{-1}(t))$, and there is a unique measure $m(dx)$, positive on open intervals, such that we have

$$t = \int_{-\infty}^{\infty} \underset{\sim}{s}(\tau(t),y)\, m(dy) \tag{3.2}$$

where $\underset{\sim}{s}(t,y) = \frac{1}{2}\frac{d}{dy}\int_0^t I_{(-\infty,y)}(B(s))ds$ is the local time of $B(t)$, jointly continuous in $(t,y)$ outside a $P^z$-null set for each $z$. Here $m(dx)$ is the "speed measure," and does not depend on $z$. The theory of such processes is highly developed, going back to W. Feller in the 1950's, and is well-represented in the book of Ito and McKean [8]. It will not concern us here, except as a starting point for the discussion.

Suppose, indeed, that instead of the Markov property we assume only the autonomous germ-Markov property as in [9, Definition 2.2]. By Definition 2.2 and Theorem 2.4 of that work, this means that there is a packet $(K \cap K_0) \subset H_0$ such that the trace of $\mathcal{H}$ on $K \cap K_0$ is generated by the functions $z(S)$, $S \in \mathcal{G}^o_{0+}$ (and hence, since $\mathcal{H}$ is countably generated, by $z(S_n)$ for a sequence $S_n \in \mathcal{G}^o_{0+}$). The germ-Markov property, as well as homogeneity in time, then follow for $z \in K \cap K_0$ as described below. But first we replace $K \cap K_0$ by $\mathcal{L}_c \cap K \cap K_0$ assuming $\mathcal{L}_c \cap K \neq \phi$, to obtain the packet of a continuous martingale autonomous germ-Markov process.

As discussed in [9], for such a process $Z_t$ the future and past of $\varphi(Z_t)$ are conditionally independent given the present germ $\sigma(Z_t(S), S \in \mathcal{G}^o_{0+})$, and this is equivalent to conditional independence of the future and past of $\varphi(Z_t)$ given the germ $\bigcap_{\varepsilon>0} \sigma(\varphi(Z_{t+s}), 0 < s < \varepsilon))$. Moreover, such a process $\varphi(Z_t)$ has a stationary transition mechanism (a function of the "germ-state") hence from an empirical viewpoint it perhaps may be said to arise from essentially the same conditions as if it were a homogeneous Markov process, i.e., in this case a regular diffusion.

However, most of the resemblance ends here, as we now indicate. In the first place unlike the case of diffusion where $\varphi(Z_t)$ and $Z_t$ may be identified, here the process $Z_t$ may be discontinuous, and the discontinuities of $Z_t$ may greatly effect the behavior of $\varphi(Z_t)$. We give two such examples: in the first the discontinuities of $Z_t$ are totally inaccessible, while in the second they are previsible.

EXAMPLE 3.5 Let $P^x$ be the probabilities of a Brownian motion $B(t)$, $B(0) = x$, and let $e_1, e_2, \ldots$ be independent, exponential random variables with parameter 1, independent of $B(t)$ for each $x$. Define a process $X(t)$ on the same probability space by

$$X(t) = \begin{cases} B(t+T_{2n}); & T_{2n} \leq t < T_{2n+1} \\ B(T_{2n+1}); & T_{2n+1} \leq t < T_{2(n+1)} \end{cases}$$

where $T_0 = 0$ and $T_n = \Sigma_{i=1}^n e_i$, $n \geq 1$. Clearly (for each $x$) $X(t)$ is a $P^x$-continuous martingale, and we may introduce corresponding probabilities $P^x$ on $(\Omega, \mathcal{G}^o_{t+}, \mathcal{F}^o_{t+}, X(t), \theta_t)$ in such a way that $w_k(t) \equiv 0$ for $k \neq 2$, $P^x$-a.s. so that we may assume $\mathcal{G}^o_{t+} = \mathcal{F}^o_{t+}$ as before. Then we introduce the prediction space and process $Z_t$, where the probabilities $P^z$ are of two kinds: either

$$P^z\{Z_t = Z_0 \text{ for all small } t\} = 1, \text{ or}$$

$$P^z\{Z_t \neq Z_0\} = 1 \text{ for each } t, \text{ respectively,}$$

according as the process $Z_{t+s}$, $s \geq 0$, starts during a "level stretch" of $\varphi(Z_t)$, or during a "Brownian stretch" of $\varphi(Z_t)$. In each case there corresponds a distinct $P^z$ for each initial point $x$, so that we can write $\mathcal{L}_c \cap K \cap K_0 = \{z(1,x), z(2,x); -\infty < x < \infty\}$ for the prediction state space. It is not hard to see that this does define a packet for which $\varphi(Z_t)$ is an autonomous germ-Markov process and $\varphi(Z_t)$ is a continuous martingale for each $P^z$. The times $T_n$ become totally inaccessible stopping times, and the character of $\varphi(Z_t)$ changes abruptly at each $T_n$. Of course, such exponential holding times are impossible for diffusions because of the strong-Markov property. Here, the strong-Markov property holds for $Z_t$ because $Z_t$ contains the "information" that a level stretch has just ended or begun, but it does not hold for $\varphi(Z_t)$.

EXAMPLE 3.6. Let $P^x$ be Brownian probabilities for $B(t)$ as before, and let $B_1(t)$ be an independent "instantaneous return" Brownian motion on $[0,1)$ for the same probability space, so that when $B_1(t-) = 1$, then $B_1(t) = 0$, and we assume that 0 is a reflecting boundary for $B_1$ in the usual sense. Let $P^{x,y}$ be the joint probabilities for $(B,B_1)$ with $B(0) = x$ and $B_1(0) = y$, and assume further a sequence $Q_1, Q_2, \ldots$ of independent Bernoulli random variables with $P^{x,y}\{Q_k = a \text{ or } b\} = \frac{1}{2}$ for all $(x,y)$, where $a \neq b$ are two strictly positive constants. We consider a process

$$X(t) = B(\tau(t)) ,$$

where $\tau(t)$ is defined as follows. Let $T_1, T_2, \ldots$ be the successive instantaneous return times of $B_1(t)$ to 0 from $1-$. Then we set $\tau(t) = Q_1 \int_0^t B_1(s)\,ds$ for $0 \leq t < T_1$, and for $n \geq 1$ we define inductively

$$\tau(t) = \tau(T_n) + Q_{n+1} \int_{T_n}^t B_1(s)ds \quad \text{for} \quad T_n \leq t < T_{n+1} .$$

Here the corresponding prediction state space is identified by triples $z = (x,y,c)$ in $R \times [0,1) \times \{a,b\}$ where $x = B(0)$, $y = B_1(0)$, and $c = Q_1$. It is not hard to recognize that this leads to a Borel packet of the prediction process for which $\varphi(Z_t)$ is autonomous germ-Markov and a continuous martingale for each $P^z$. Here the times $T_n$ are previsible stopping times, since they occur when the "rate" $d\tau(t) = Q_n B_1(t)dt$ reaches its maximum $Q_n$ on each cycle. Also, $Z_t$ has a previsible jump at each $T_n$ since the value of $Q_{n+1}$ is not determined by $Z_{T_n-}$, but is determined by $Z_{T_n}$ ($Z_{T_n-}$ is thus a branching point). Since $\varphi(Z_{T_n})$ is arbitrary, $\varphi(Z_t)$ is not a strong-Markov process in the usual sense. But $\varphi(Z_t)$ is always a strong-germ-Markov process (as defined and proved in Theorem 2.3 of [9]).

From these examples it is clear that germ-Markov processes exhibit much more variety of behavior than Markov processes, even under quite restrictive assumptions. The situation is not much simpler even if we require $Z_t$ to be continuous along with $\varphi(Z_t)$ . Thus if we set $a = b$ in Example 3.6, so that the $Q_n$'s are constant, $Z_t$ becomes continuous but $X(t)$ still has predictable but sudden changes of behavior at the times $T_n$ .

In this example, the time scale $\tau(t)$ is independent of $B(t)$ . If we permit dependence, then two general types of process (with continuous $Z_t$) still may be distinguished. The first may be called processes in which the speed measure developes independently of position. Here we may begin with any fixed speed measure $m(dy)$, and any autonomous germ-Markov process $g(t)$ with a continuous prediction process and such that the process $\psi(t) = \int_0^t g(s)ds$ is strictly increasing (in the last example, $\psi(t) = \int_0^t B_1(s)\,ds)$ . Now let $B(t)$, $B(0) = 0$, be a Brownian motion independent of $\psi(t)$, with local times $\underset{\sim}{s}(t,y)$ as in (3.2). We may then define a random time $\tau(t)$ by

$$\psi(t) = \int_{-\infty}^{\infty} \underset{\sim}{s}(\tau(t),y)m(dy) \ , \tag{3.3}$$

and then set $X(t) = B(\tau(t))$ . It is to be shown that $X(t)$ is an autonomous germ-Markov process, with a continuous prediction process, which is a continuous local martingale. In fact, if $X_m(t)$ denotes the regular diffusion with speed measure $m(dy)$ based on $B(t)$ as in (3.2) and if $\tau_m(t)$ is the corresponding additive functional of $X_m(t)$, then we have

$$X(t) = X_m(\psi(t)) = B(\tau_m\psi(t)) \ . \tag{3.4}$$

Now since $\psi(t)$ is independent of $X_m$ it can be seen that

$$\tau_m\psi(t) = \lim_{n\to\infty} \sum_{k=1}^{n} (X(\frac{kt}{n}) - X(\frac{(k-1)}{n}))^2 \ ,$$

at least in the sense of convergence in probability. This is enough to see that $\tau_m\psi$ is an additive functional of $X$ . Next, we will obtain

$$\int_0^t g^{-1}(\underset{\sim}{s})\,ds(\tau_m\psi(s),y) \tag{3.5}$$

as an expression for the local time of $X$ at $y$ with respect to $m(dy)$ . Setting $u = \psi(s)$, we have from (3.4)

$$\int_0^t I_{(-\infty,y)}(X(s))ds = \int_0^{\psi(t)} I_{(-\infty,y)}(X_m(u))d\psi^{(-1)}(u) \ ,$$

where $d\psi^{(-1)}(u) = (g(\psi^{(-1)}(u)))^{-1}\,du$ . But for bounded step functions $f(u)$ we have $P^z$-a.s.

$$\int_0^{\psi(t)} I_{(-\infty,y)}(X_m(u))\, f(u)du$$

$$= \int_{-\infty}^{y} \left[\int_0^{\psi(t)} f(u)d\underset{\sim}{s}(\tau_m(u),y\right] m(dy) ,$$

since $\underset{\sim}{s}(\tau_m(u),y)$ is the local time of $X_m$ . Since this holds $P^z$-a.s. for a countable family of step functions generating $b(\mathcal{B})$, it follows by monotone extension that it holds for all Borel $f \geq 0$ . Substituting $f(u) = (g(\psi^{(-1)}(u)))^{-1}$, differentiating with respect to $m(dy)$, and finally returning to the variable $s$, yield (3.5).

Integrating (3.5) with respect to $dy$ gives $\int_0^t g^{-1}(s)d(\tau_m\psi(s))$ which is therefore also an additive functional of $X$ . We denote it by $C(t)$, and observe that

$$\psi(t) = \int_0^t \frac{d\tau_m\psi(s)}{dC(s)} ds ,$$

where the integrand is the Lebesgue density as indicated. Thus $\psi(t)$ is an additive functional of $X$ . Then the germ of $\psi(t)$ is contained in that of $X(t)$, and hence the germ of $g(t)$ is also. But this together with $X(t)$ determines the prediction process of $X(t)$ autonomously, in view of our assumptions on $g$ and $B$ . It is clear that $X(t)$ is a continuous local martingale, and that its prediction process is continuous along with that of $g(t)$ .

It is quite apparent how to extend this type of example to germ-Markov functionals $\psi(t)$ other than those which have a density $g(t)$ with respect to Lebesgue measure. The analogue of the speed measure of the process at time $t$ is given formally by $\frac{dt}{d\psi} m(dy)$, or $(1/g(t))m(dy)$ in our special case, and it evolves independently of the position $X(t)$ .

Not surprisingly, this is not the only type of continuous local martingale which is an autonomous germ-Markov process. There are also cases in which the evolution of $g(t)$ depends on $B(t)$ . One such example is the solution of the stochastic integral equation

$$X(t) = x_0 + \int_0^t \left[\int_0^s X(\tau)d\tau\right] dB(s) \ ; \quad x_0 \neq 0 \ .$$

The existence and uniqueness of the pathwise solution, given any (continuous) Brownian motion $B(t)$, is proved in Section 3.4 of [9]. Here the additive functional $\tau(t)$ is clearly

$$\tau(t) = \int_0^t \left[\int_0^s (X(u))du\right]^2 ds \ .$$

Thus if we write formally

$$dt = d \int_{-\infty}^{\infty} \underset{\sim}{s}(\tau(t),y)m(dy)$$

we find that this is satisfied at time $t$ if

$$m(dy) = m_t(dy) = 2\ dy\ dt/d\tau$$
$$= 2\left[\int_0^t X(s)ds\right]^{-2} dy\ .$$

On the other hand, if we fix $m(dy) = 2dy$ as in (3.3) the analogue of $\psi(t)$ is just $\tau(t)$, and clearly it depends on $X(t)$. It might be of interest to look for further examples of this type in which $m(dy) \neq c\ dy$.

As examples of continuous martingales, such processes are rather specialized. However, in view of the significance of the martingale property (or natural scale) for diffusion, it seems a natural first step to consider it also for a germ diffusion. But perhaps the chief significance of the examples is only to call attention to the fact that germ-diffusion processes are very much less limited in behavior than ordinary diffusions. Since they both give expression to essentially the same underlying physical hypotheses, it would seem necessary to use some caution before assuming the validity of a diffusion model of a real phenomenon.

## REFERENCES


1. Beveniste, A. and Jacod, J. "Systemès de Lévy des processus de Markov," Inventiones math., 21, 1973, 183-198.

2. Blumenthal, R. M. and Getoor, R. K. Markov Processes and Potential Theory. Academic Press, New York, 1968.

3. Dellacherie, C. and Meyer, P.-A. Probabilités et Potentiel, Chapter I - IV. Hermann, Paris, 1975.

4. Dellacherie, C. and Meyer, P.-A. Ibid., Chapter V - VIII, Hermann, Paris, 1980.

5. Doob, J. L. Stochastic Processes. Wiley and Sons, New York, 1953.

6. Getoor, R. K. Markov processes: Ray processes and right processes. Lecture Notes in Math. 440, Springer, Berlin, 1975.

7. Getoor, R. Homogeneous potentials, Seminaire de Prob. XII, 398-410. Lecture Notes in Math. 649, Springer, Berlin, 1978.

8. Ito, K. and McKean, H. P., Jr. Diffusion Processes and their Sample Paths. Academic Press, Inc., New York, 1965.

9. Knight, F. B. "Prediction processes and an autonomous germ-Markov property," The Annals of Probability, 7, 1979, 385-405.

10. Kunita, H. and Watanabe, S. "On square integrable martingales," Nagoya Math. J., 30, 1967, 209-245.

11. Meyer, P.-A. Probability and Potentials. Blaisdell, Waltham, Mass., 1966.

12. Motoo, M. and Watanabe, S. "On a class of additive functionals of Markov processes," J. Math. Kyoto Univ., 4, 1965, 429-469.

13. Neveu, J. Mathematical Foundations of the Calculus of Probability. Holden-Day, Inc., San Francisco, 1965.

14. Seminaire de Probabilités VII, Univ. de Strasbourg. P.-A. Meyer. Limites mediales, d'apres Mokobodzki. Lecture Notes in Math. 321, Springer, Berlin, 198-204, 1973.

15. Sur, M. G. Continuous additive functionals of a Markov process, English translation: Soviet Math. 2 (Dokl. Akad. Nauk SSSR), 1961, 365-368.

16. Walsh, J. B. and Weil, M. "Representation des temps terminaux et application aux fonctionelles additives et aux systemes de Lévy," Ann. Sci. Ec. Norm., Sup. 5, 1972, 121-155.