

# CONSTRUCTING STRONG MARKOV PROCESSES

ROBERT J. VANDERBEI

*Dedicated to the memory of Lynda Singhsinsuk.*

ABSTRACT. The construction presented in this paper can be briefly described as follows: starting from any “finite-dimensional” Markov transition function  $p_t$ , on a measurable state space  $(E, \mathcal{B})$ , we construct a strong Markov process on a certain “intrinsic” state space that is, in fact, a closed subset of a finite dimensional Euclidean space  $\mathbb{R}^d$ . Of course we must explain the meaning of finite-dimensionality and intrinsicity. Starting with  $p_t$ , we consider the range of the nonnegative bounded measurable functions under the action of the resolvent. This class of functions induces a uniform structure on  $E$ . We say that  $E$  is finite-dimensional if this uniformity is finitely generated. In such cases we then map  $E$  into  $\mathbb{R}^d$ . The intrinsic state space is the closure of the range of this mapping. On this enlarged state space we construct a strong Markov process, which corresponds quite naturally to  $p_t$ . We give several examples including the usual examples of nonstrong Markov process.

## 1. INTRODUCTION AND NOTATION

Every student of probability learns very early that not every Markov process possesses the strong Markov property. And, as is often the case, the simplest pathological examples are very simple indeed. For instance, the process that moves deterministically to the left on the real line with unit speed except that when it reaches zero it pauses there for an exponential amount of time before continuing, is perhaps the best such example. The fact that it is a Markov process follows from the memoryless property of the exponential random variable. To see that it is not a strong Markov process, consider the first hitting time of the open left half-line. Starting from this time, the process proceeds immediately into the left half-line. However, at this random time the process is situated at the origin and consequently if it had the strong Markov property it would have to remain there for

---

*Date:* Original draft: August 1984. Converted to LaTeX: March 11, 2013.

*1991 Mathematics Subject Classification.* 60J25, 47D07 .

*Key words and phrases.* strong Markov property, Ray-Knight compactification, completion of state space.

Original version is archived at

<http://vanderbei.princeton.edu/tex/StrongMarkovProcesses/CSMP.pdf> .

an exponential amount of time. The fact that it does not shows that it is not strong Markov.

Of course the above pathology can be remedied by splitting 0 into two points  $0^+$  and  $0^-$ . If we put the exponential alarm clock at  $0^+$  and have the process immediately appear at  $0^-$  when the clock rings (i.e. a right continuous trajectory) then in this enlarged state space the process is strong Markov.

This idea of enlarging the state space first appeared in the paper [15] by Yushkevich. Shortly after that, D. Ray published a fundamental paper [14] in which he introduced a certain compactification of the state space on which the Markov process becomes what is today called a Ray process (which has the strong Markov property). Ray's methods were clarified and improved by F. Knight [13]. The short monograph [11] by Gettoor has become the standard reference on Ray processes and the Ray-Knight compactification. However, the most general results appeared later in the papers [3, 4, 5, 6, 7, 8, 9] of H. J. Engelbert. His basic assumption is that there exists a certain class of nonnegative bounded measurable functions, on a countably generated measurable state space, that forms a separable Ray cone (see Definition 12 in [5]).

The starting point of this paper is a Markov transition function on a measurable state space. We introduce in a canonical way a certain class  $R_+$  of functions that satisfy most of the requisite properties of a Ray cone. The class  $R_+$  is generally not separable (even for the above example) but is usually "finite dimensional." This means that there exists a finite collection  $\mathcal{E} = \{e_1, e_2, \dots, e_d\}$  of real-valued functions (not necessarily bounded) such that, for every sequence of points  $x_n$  in the state space  $E$ ,  $f(x_n)$  is Cauchy for every  $f \in R_+$  if and only if  $e_j(x_n)$  is Cauchy for  $i = 1, 2, \dots, d$ . By thinking of the functions  $e_j$  as coordinate functions, the map  $x \rightarrow (e_1(x), e_2(x), \dots, e_d(x))$  embeds the state space in  $\mathbb{R}^d$ . We then take the closure of the range of this mapping to get our intrinsic state space. Thinking in terms of our original state space  $E$ , this closure corresponds to a completion of the state space relative to the uniformity generated by  $R_+$  (for a general discussion of uniformization see e.g. [12] Chapter VI). On the resulting closed space we construct a strong Markov process that corresponds naturally to the original Markov transition function.

The advantage of our approach is that the construction is canonical. The Ray-Knight compactification, on the other hand, depends on the separable Ray cone with which one starts. In fact a substantial part of the theory of the Ray-Knight compactification consists in studying to what extent the new state space depends on the Ray cone (see e.g. [11], Chapter 15).

The main goal of this paper is the construction of a process with the strong Markov property. Along the way we get some other results; e.g. right continuity and the existence of left limits (recently this has been called *cadlaguity*—a *gauche* term indeed). Since the state space we construct is a closed subset of  $\mathbb{R}^d$ , the further sample path properties are for the most part immediate consequences of already well known results and so they will not be pursued here.

The remainder of this section is devoted to establishing notations that will be used throughout the rest of the paper.

DEFINITION. A *kernel*  $k$  on a measurable space  $(E, \mathcal{B})$  is a real-valued function of  $E \times \mathcal{B}$  satisfying:

- 1.1  $x \rightarrow k(x, B)$  is  $\mathcal{B}$ -measurable;
- 1.2  $B \rightarrow k(x, B)$  is a signed measure having finite total variation measure  $|k|(x, \cdot)$ .

The formula

$$1.3 \quad \|k\| = \sup_x |k|(x, E)$$

defines a norm on the space  $K$  of all kernels for which (1.3) is finite. It is easy to check that  $K$  is a Banach space. The formula

$$k * k'(x, B) = \int k(x, dy)k'(y, B)$$

makes  $K$  into an algebra. We denote by  $k^{(n)}$  the  $n^{\text{th}}$  power of the kernel  $k$ . We also put

$$k^{(0)}(x, B) = 1(x, B) = \begin{cases} 1 & x \in B \\ 0 & x \notin B. \end{cases}$$

It is easy to see that  $\|k * k'\| \leq \|k\|\|k'\|$  and hence that  $\|k^{(n)}\| \leq \|k\|^n$ . A kernel  $k$  is called a *stochastic kernel* if, for every  $x \in E$ ,  $k(x, \cdot)$  is a probability measure.

If  $(E, \mathcal{B})$  is a measurable space, we also denote by  $\mathcal{B}$  the Banach space of bounded  $\mathcal{B}$ -measurable functions on  $E$  and we denote by  $\mathcal{B}^+$  those  $f \in \mathcal{B}$  that are nonnegative. The norm on  $\mathcal{B}$  is given by  $\|f\| = \sup_{x \in E} |f(x)|$ . For every kernel (say  $k$ ) on  $(E, \mathcal{B})$ , we denote by the corresponding upper case letter ( $K$  in this case) the linear operator on  $\mathcal{B}$  defined by

$$Kf(x) = \int k(x, dy)f(y).$$

If  $B \rightarrow k(x, B)$  is a positive measure then  $K$  is a positive operator and if  $k$  is a stochastic kernel then  $K1 = 1$  and  $\|Kf\| \leq \|f\|$ .

If  $E$  is a subset of  $\mathbb{R}^d$ , we denote by  $C^k(E)$  the collection of all real-valued bounded continuous functions on  $E$  that have limits on the closure of  $E$  and that are  $k$  times continuously differentiable with each derivative being bounded. We will write  $C(E)$  instead of  $C^0(E)$ . The notation  $f \in C^k(E_1, E_2, \dots, E_n)$  indicates that, for each  $j$ , the function  $f$  restricted to  $E_j$  is of class  $C^k(E_j)$ . The notation  $C_0(E)$  denotes those functions in  $C(E)$  that tend to zero at infinity. Hence if  $E$  is bounded,  $C_0(E) = C(E)$ . Finally  $\bar{E}$  denotes the closure of  $E$ .

## 2. THE INTRINSIC STATE SPACE

DEFINITION. A family of kernels  $(p_t)_{t>0}$  on a measurable space  $(E, \mathcal{B})$  is called a *Markov transition function* if:

- 2.1  $(t, x) \rightarrow p_t(x, B)$  is  $\mathcal{B}(\mathbb{R}^+) \times \mathcal{B}$  measurable;
- 2.2 every  $p_t$  is a stochastic kernel;
- 2.3  $p_{s+t} = p_s * p_t$ .

The space  $(E, \mathcal{B})$  is called the *state space* for  $p_t$ .

DEFINITION. A family of kernels  $(u^\alpha)_{\alpha>0}$  on a measurable space  $(E, \mathcal{B})$  is called a *Markov resolvent* if

- 2.4 for every  $\alpha > 0$ ,  $\frac{1}{\alpha}u^\alpha$  is a stochastic kernel;
- 2.5  $u^\beta - u^\alpha = (\alpha - \beta)u^\alpha * u^\beta$ .

If  $p_t$  is a Markov transition function, the formula

$$u^\alpha = \int_0^\infty e^{-\alpha t} p_t dt$$

defines a Markov resolvent.

For the rest of this section, let  $p_t$  denote a fixed Markov transition function on a fixed state space  $(E, \mathcal{B})$  and let  $u^\alpha$  be the corresponding resolvent. We put

$$\mathcal{C} = U^\alpha \mathcal{B}.$$

To see that the right hand side does not depend on  $\alpha$  (as long as  $\alpha > 0$ ) rewrite property (2.5) as  $u^\beta = u^\alpha * [1 + (\alpha - \beta)u^\beta]$ . Hence, if  $f = U^\beta g$  and  $g \in \mathcal{B}$ , then  $f = U^\alpha h$  where  $h = g + (\alpha - \beta)U^\beta g$  clearly also belongs to  $\mathcal{B}$ . Put

$$\mathcal{C}^+ = \bigcup_{\alpha>0} U^\alpha \mathcal{B}^+.$$

The following properties of  $\mathcal{C}$  and  $\mathcal{C}^+$  are easily verified:

- 2.6  $U^\alpha \mathcal{C}^+ \subset \mathcal{C}^+$ ,  $U^\alpha \mathcal{C} \subset \mathcal{C}$ ;
- 2.7  $\mathcal{C}^+$  is a convex cone,  $\mathcal{C}$  is a linear space;
- 2.8  $\mathcal{C}^+$  contains the nonnegative constants;
- 2.9  $\mathcal{C} = \mathcal{C}^+ - \mathcal{C}^+$ .

DEFINITION. A function  $f$  is called  $\alpha$ -supermedian if  $f \in \mathcal{B}^+$  and  $f \geq (\beta - \alpha)U^\beta f$  for all  $\beta > 0$ .

2.10 PROPOSITION. The following properties hold:

- (i) every  $f \in \mathcal{C}^+$  is  $\alpha$ -supermedian for some  $\alpha > 0$ ;
- (ii) for every  $f \in \mathcal{C}$ ,  $\lim_{t \downarrow 0} \|P_t f - f\| = 0$ .

*Proof.* Property (i) is an immediate consequence of the definition of  $\mathcal{C}^+$  and the resolvent property (2.5). To prove (ii), put  $g = U^\alpha f$ . It is not hard to show that

$$e^{-\alpha t} P_t g - g = \int_0^t e^{-\alpha s} P_s f ds.$$

Hence

$$0 \leq \|e^{-\alpha t} P_t g - g\| \leq t \|f\|,$$

and so

$$\begin{aligned} 0 &\leq \|P_t g - g\| \\ &\leq \|P_t g - e^{-\alpha t} P_t g\| + \|e^{-\alpha t} P_t g - g\| \\ &\leq (1 - e^{-\alpha t}) \|g\| + t \|f\|. \end{aligned}$$

Letting  $t$  tend to zero establishes (ii). □

It is easy to see that the minimum of two  $\alpha$ -supermedian functions is again  $\alpha$ -supermedian. Also convex combinations of  $\alpha$ -supermedian functions are again  $\alpha$ -supermedian.

2.11 THEOREM. There exists a unique minimal convex cone  $\mathcal{R}_+$  such that:

- (i)  $\mathcal{C}^+ \subset \mathcal{R}_+ \subset \mathcal{B}^+$ ;
- (ii) for  $\alpha > 0$ ,  $U^\alpha \mathcal{R}_+ \subset \mathcal{R}_+$ ;
- (iii)  $f, g \in \mathcal{R}_+$  implies that  $f \wedge g \in \mathcal{R}_+$ .

In addition, every  $f \in \mathcal{R}_+$  is  $\alpha$ -supermedian for some  $\alpha > 0$ .

*Proof.* If  $\mathcal{H}$  is any convex cone in  $\mathcal{B}^+$ , we define the usual enlargements  $U(\mathcal{H})$  and  $\Lambda(\mathcal{H})$  as follows:

$$\begin{aligned} U(\mathcal{H}) &= \{U^{\alpha_1} f_1 + \cdots + U^{\alpha_n} f_n \mid \alpha_j > 0, f_j \in \mathcal{H}, 1 \leq j \leq n, n \geq 1\} \\ \Lambda(\mathcal{H}) &= \{f_1 \wedge f_2 \wedge \cdots \wedge f_n \mid f_j \in \mathcal{H}, 1 \leq j \leq n, n \geq 1\}. \end{aligned}$$

It is easy to check that  $U(\mathcal{H})$  and  $\Lambda(\mathcal{H})$  are again convex cones in  $\mathcal{B}^+$ . Also the sum of two convex cones is again a convex cone.

Let  $\mathcal{R}_0 = \mathcal{C}^+$  and define  $\mathcal{R}_n$  recursively by  $\mathcal{R}_n = \Lambda(\mathcal{R}_{n-1} + U(\mathcal{R}_{n-1}))$ . Put  $\mathcal{R}_+ = \bigcup_n \mathcal{R}_n$ . Note that  $\mathcal{R}_n \subset \mathcal{R}_{n+1}$  for each  $n \geq 0$ . By induction on  $n$ , it follows that each  $\mathcal{R}_n$  is a convex cone in  $\mathcal{B}^+$  and that every  $f \in \mathcal{R}_n$  is  $\alpha$ -supermedian for some  $\alpha > 0$ . Hence  $\mathcal{R}_+$  also has these properties. If  $f, g \in \mathcal{R}_+$  then there exists an  $n \geq 0$  with  $f, g \in \mathcal{R}_n$ . Then  $f \wedge g \in \mathcal{R}_{n+1} \subset \mathcal{R}_+$ . Also, if  $f \in \mathcal{R}_n$  then  $U^\alpha f \in \mathcal{R}_{n+1} \subset \mathcal{R}_+$ . It is obvious from the construction that  $\mathcal{R}_+$  is the unique minimal convex cone that satisfies (i)-(iii).  $\square$

**2.12 DEFINITION.** Two classes  $\mathcal{H}$  and  $\mathcal{G}$  of real-valued functions defined on a set  $E$  are said to generate the same uniformity on  $E$  if, for every sequence  $x_n \in E$ , the following are equivalent:

- (i)  $h(x_n)$  is a Cauchy sequence for every  $h \in \mathcal{H}$ ,
- (ii)  $g(x_n)$  is a Cauchy sequence for every  $g \in \mathcal{G}$ .

We write  $\mathcal{H} \sim \mathcal{G}$  to indicate that  $\mathcal{H}$  and  $\mathcal{G}$  generate the same uniformity. We say that *the uniformity generated* by a class  $\mathcal{G}$  is *finitely generated* if there exists a finite collection  $\mathcal{H}$  such that  $\mathcal{H} \sim \mathcal{G}$ . We will only consider finitely generated uniformities (which is why we do not need to introduce nets).

The basic example to keep in mind is the following: let  $E = \mathbb{R}^d$ ,  $\mathcal{H}$  be the set of coordinate functions and  $\mathcal{G}$  be the set of bounded continuous functions. Then  $\mathcal{H} \sim \mathcal{G}$  and since  $\mathcal{H}$  is finite,  $\mathcal{G}$  is finitely generated. In this case  $h(x_n)$  is Cauchy for all  $h \in \mathcal{H}$  if and only if  $x_n$  is a Cauchy sequence relative to the Euclidean metric.

It is easy to see that the relation  $\sim$  is transitive:  $\mathcal{H} \sim \mathcal{H}'$  and  $\mathcal{H}' \sim \mathcal{H}''$  implies that  $\mathcal{H} \sim \mathcal{H}''$ . We also have

**2.13 PROPOSITION.** If  $\mathcal{H} \sim \mathcal{G}$ , then  $\mathcal{H}$  separates points  $x$  and  $y$  if and only if  $\mathcal{G}$  does.

*Proof.* If  $h(x) = h(y)$  for all  $h \in \mathcal{H}$  then  $(h(x), h(y), h(x), h(y), \dots)$  is a Cauchy sequence for every  $h \in \mathcal{H}$ . Hence  $(g(x), g(y), g(x), g(y), \dots)$  is a Cauchy sequence for every  $g \in \mathcal{G}$  and so  $g(x) = g(y)$  for all  $g \in \mathcal{G}$ .  $\square$

Put  $\mathcal{R} = \mathcal{R}_+ - \mathcal{R}_+$ . We are now ready to introduce our basic assumption on  $p_t$ :

**2.14 ASSUMPTION.** There exists a finite collection  $\mathcal{E} = \{e_1, e_2, \dots, e_d\}$  of real-valued functions on  $E$  (not necessarily bounded) such that  $\mathcal{E} \sim \mathcal{R}$ .

The examples in Section 4 show that this condition is often readily verifiable. In analogy with the example given above, it is good to think of functions in  $\mathcal{R}$  as bounded continuous functions and those in  $\mathcal{E}$  as coordinate functions. With this in mind we see that 2.14 is the assumption that the state space is finite dimensional.

In assumption 2.14, it is clear that we could replace  $\mathcal{R}$  by  $\mathcal{R}_+$ . Define a map  $\psi$  from  $E$  into  $\mathbb{R}^d$  by

$$\psi(x) = (e_1(x), e_2(x), \dots, e_d(x)).$$

The mapping  $\psi$  may fail to be injective (see Example 4.8). In any case, we will abuse language and refer to  $\psi$  as an embedding of  $E$  into  $\mathbb{R}^d$ . Even if  $\psi$  is not injective, we have the following result.

**2.15 PROPOSITION.** If  $\psi(x) = \psi(y)$  then  $f(x) = f(y)$  for all  $f \in \mathcal{R}$ .

*Proof.* Immediate consequence of Proposition 2.13.  $\square$

It follows that  $\psi$  is injective if and only if  $\mathcal{R}$  separates points.

**2.16 PROPOSITION.** Suppose that  $\mathcal{E} = \{e_1, e_2, \dots, e_d\}$  and  $\mathcal{E}' = \{e'_1, e'_2, \dots, e'_{d'}\}$  satisfy 2.14 and let  $\psi$  and  $\psi'$  be the corresponding embeddings. Then there exists a continuous bijection between  $\overline{\psi(\mathcal{E})} \subset \mathbb{R}^d$  and  $\overline{\psi'(\mathcal{E}')} \subset \mathbb{R}^{d'}$  whose inverse is also continuous.

*Proof.* Since  $\sim$  is transitive we see that  $\mathcal{E} \sim \mathcal{E}'$ . Hence, it follows from proposition 2.13 that  $\psi' \circ \psi^{-1}$  and  $\psi \circ (\psi')^{-1}$  are well defined. It is also clear that  $x_n$  is a Cauchy sequence in  $\psi(\mathcal{E})$  if and only if  $x'_n = \psi' \circ \psi^{-1}(x_n)$  is Cauchy in  $\psi'(\mathcal{E}')$ . Hence  $\psi' \circ \psi^{-1}$  and  $\psi \circ (\psi')^{-1}$  can both be extended to the closures of their domains and give us the desired bijections.  $\square$

Proposition 2.16 shows that our construction is canonical and that the specific choice of  $\mathcal{E}$  is just a matter of choosing coordinate functions.

Proposition 2.15 shows that the following formula defines unambiguously a mapping  $\Psi$ , from  $\mathcal{R}$  into the space of real-valued functions on  $\psi(\mathcal{E})$ :

$$(\Psi f)(\xi) = f(\psi^{-1}(\xi)), \quad \xi \in \psi(\mathcal{E}), \quad f \in \mathcal{R}.$$

The mapping  $\Psi$  is linear and injective. Also note that  $\Psi e_i$  makes sense and is the  $i^{\text{th}}$  coordinate function on  $\psi(\mathcal{E})$ .

**2.17 PROPOSITION.** For every  $f \in \mathcal{R}$ ,  $\Psi f$  is a continuous function on  $\psi(\mathcal{E})$ .

*Proof.* Suppose that  $\xi_m \rightarrow \xi$ . Choose  $x_m$ , and  $x$  in  $E$  such that  $x_m \in \psi^{-1}(\xi_m)$  and  $x \in \psi^{-1}(\xi)$ . Then  $(e_i(x_1), e_i(x), e_i(x_2), e_i(x), \dots)$  is a Cauchy sequence for each  $i$ . By assumption 2.14,  $(f(x_1), f(x), f(x_2), f(x), \dots)$  is a Cauchy sequence for every  $f \in \mathcal{R}$ . Hence  $f(x_m) \rightarrow f(x)$  for every  $f \in \mathcal{R}$  and so  $\Psi f(\xi_m) \rightarrow \Psi f(\xi)$  for every  $f \in \mathcal{R}$ .  $\square$

For each  $f \in \mathcal{R}$  we can extend the definition  $\Psi f$  to  $\overline{\psi(\mathcal{E})}$  by continuity. To see this, suppose that  $\xi_m \rightarrow \xi$  and  $\xi_m \in \psi(\mathcal{E})$ . Then  $e_i(x_m)$  is Cauchy for each  $i = 1, \dots, n$  where  $x_m$  is chosen from  $\psi^{-1}(\xi_m)$ . Assumption 2.14 then implies that  $f(x_m)$  is Cauchy for every  $f \in \mathcal{R}$ . Denote this new function by  $\overline{\Psi f}$ . Put  $\mathcal{R}' = \{\overline{\Psi f} : f \in \mathcal{R}\}$ ,  $\mathcal{R}'_+ = \{\overline{\Psi f} : f \in \mathcal{R}_+\}$  and  $E' = \overline{\Psi(\mathcal{E})}$ . The space  $E'$  is called the *intrinsic state space* for  $p_t$ . The class  $\mathcal{R}'_+$  is a convex subcone of  $\mathcal{C}^+(E')$  that contains the nonnegative constant functions and is closed under pointwise minimization. Also  $\mathcal{R}' = \mathcal{R}'_+ - \mathcal{R}'_+$  so it follows that  $\mathcal{R}'$  is a linear subspace of  $C(E')$  which contains the constant functions and is closed under pointwise minimization and maximization. We will require one final property of  $\mathcal{R}'$  which we state as a Lemma.

**2.18 LEMMA.** The class  $\mathcal{R}'$  separates points in  $E'$ .

*Proof.* Suppose that  $\xi, \eta \in E'$  and  $\overline{\Psi f}(\xi) = \overline{\Psi f}(\eta)$  for all  $f \in \mathcal{R}$ . Let  $\xi_n$  be a sequence of points in  $\psi(\mathcal{E})$  that converges to  $\xi$  and let  $\eta_n$  be a sequence converging to  $\eta$ . Then  $\lim_n \Psi f(\xi_n) = \lim_n \Psi f(\eta_n)$ . Let  $x_n$  be a point in  $\psi^{-1}(\xi_n)$  and let  $y_n$  be a point in  $\psi^{-1}(\eta_n)$ . Then we have  $\lim_n f(x_n) = \lim_n f(y_n)$ . Hence, the sequence  $(f(x_1), f(y_1), f(x_2), f(y_2), \dots)$  is Cauchy for every  $f \in \mathcal{R}$ . By assumption 2.14,  $(e_i(x_1), e_i(y_1), e_i(x_2), e_i(y_2), \dots)$  is



Cauchy for  $i = 1, \dots, d$ . Hence  $\lim_n e_i(x_n) = \lim_n e_i(y_n)$  for all  $i$ . This means that  $\lim \xi_n = \lim \eta_n$  and so  $\xi = \eta$ .  $\square$

For each  $\alpha > 0$ , define a positive linear operator  $V^\alpha$  on  $\mathcal{R}'$  as follows:

$$V^\alpha : \overline{\Psi f} \rightarrow \Psi f \rightarrow f \rightarrow U^\alpha f \rightarrow \Psi(U^\alpha f) \rightarrow \overline{\Psi(U^\alpha f)}.$$

The first map is simply the restriction map. The second map is  $\Psi^{-1}$ , which exists since  $\Psi$  is an injection. The fourth and fifth maps make sense since  $U^\alpha \mathcal{R} \subset \mathcal{R}$ . It is easy to see that  $V^\alpha$  inherits the following properties from  $U^\alpha$ :

- (i)  $\alpha V^\alpha 1 = 1$ ;
- (ii)  $V^\alpha - V^\beta = (\beta - \alpha)V^\alpha V^\beta$ ;
- (iii) if  $f_n \downarrow 0$  pointwise and  $f_n \in \mathcal{R}'$ , then  $V^\alpha f_n \downarrow 0$  pointwise.

By Daniel's theorem (see e.g. [2] Section III-35), for each  $x \in E'$ , and  $\alpha > 0$ , there exists a unique measure  $\nu^\alpha(x)$  on  $\sigma(\mathcal{R}')$  such that

$$V^\alpha f(x) = \int \nu^\alpha(x, dy) f(y), \quad f \in \mathcal{R}'.$$

Since  $\mathcal{R}'$  separates points in  $E'$  and  $E'$  is closed, it follows from Proposition 5.1 that  $\sigma(\mathcal{R}') = \mathcal{B}(E')$ . From the lattice version of the monotone class theorem (see e.g. [2], Theorem 22.3) and the fact that  $x \rightarrow \nu^\alpha(x, f)$  is  $\mathcal{B}(E')$  measurable (even continuous) for every  $f \in \mathcal{R}'$ , it follows that  $x \rightarrow \nu^\alpha(x, f)$  is  $\mathcal{B}(E')$  measurable for every bounded  $\mathcal{B}(E')$  measurable  $f$ . Another monotone class argument shows that  $\nu^\alpha$  is a resolvent.

In the next section, we start with a state space  $E$ , a family of real-valued functions  $\mathcal{R}_+$  on  $E$ , and a resolvent  $u^\alpha$  which all together satisfy the conditions which we have found that  $E'$ ,  $\mathcal{R}'_+$  and  $\nu^\alpha$  satisfy. From these we show that such a resolvent gives rise to a unique Markov transition function and that from this transition function we can construct a strong Markov process in the space of right continuous trajectories that have left hand limits.

### 3. FROM RESOLVENT TO STRONG MARKOV PROCESS

In this section, we take as our starting point properties of the intrinsic state space that were developed in Section 2. Hence we make the following assumptions. Let  $(E, \mathcal{B})$  be a closed subset of  $\mathbb{R}^d$  with the Borel  $\sigma$ -algebra. Suppose that there is given a convex subcone  $\mathcal{R}_+$  of  $C^+(E)$  that contains the nonnegative constant functions and that separates points in  $E$ . Finally, suppose that we have a Markov resolvent  $U^\alpha$  on  $(E, \mathcal{B})$  with the following properties:

3.1  $U^\alpha \mathcal{R}_+ \subset \mathcal{R}_+$ ;

3.2 every  $f \in \mathcal{R}_+$  is  $\alpha$ -supermedian for some  $\alpha > 0$ .

Put  $\mathcal{R} = \mathcal{R}_+ - \mathcal{R}_+$ . We need one additional assumption that is not a carry-over from Section 2 (or maybe it is—this is an open problem):

3.3  $C_0(E) \subset \text{cl}(\mathcal{R})$ .

The notation  $\text{cl}(\mathcal{R})$  means the closure of  $\mathcal{R}$  in the sup norm. If  $E$  is a bounded subset of  $\mathbb{R}^d$  then  $C_0(E) = C(E)$  and, by the lattice version of the Stone-Weierstrass theorem,  $C(E) = \text{cl}(\mathcal{R})$ . Hence, in this case assumption (3.3) is automatic. The examples in Section 4 show that, even when  $E$  is unbounded, assumption (3.3) is easily verified. A resolvent that satisfies (3.1) and (3.2) will be called an *Engelbert resolvent*.

The assumptions in force in this section are very similar to those of a Ray resolvent. Indeed for a *Markov Ray resolvent*  $U^\alpha$  one assumes that  $E$  is a compact metric space,  $\mathcal{B}$  is the Borel  $\sigma$ -algebra, and

3.4  $U^\alpha C(E) \subset C(E)$ ;

3.5 the continuous  $\alpha$ -supermedian functions separate points in  $E$ .

We have the following connection between Ray and Engelbert resolvents.

3.6 THEOREM. If  $E$  is a compact subset of  $\mathbb{R}^d$  and  $U^\alpha$  is a Markov Ray resolvent, then  $U^\alpha$  is an Engelbert resolvent.

*Proof.* We must find a family  $\mathcal{R}_+$  with the desired properties. We may take  $\mathcal{R}_+$  to be the family of all continuous functions that are  $\alpha$ -supermedian for some  $\alpha > 0$ . For any  $f \in \mathcal{B}^+$ ,  $U^\alpha f$  is  $\alpha$ -supermedian and, by (3.4),  $U^\alpha$  maps continuous functions into continuous functions. Hence  $U^\alpha \mathcal{R}_+ \subset \mathcal{R}_+$ . Property (3.5) says that  $\mathcal{R}_+$  separates points. It is clear that  $\mathcal{R}_+$  is a convex subcone of  $C^+(E)$  and so the proof is complete.  $\square$

With the aim of constructing a strong Markov process, our first goal is to invert the Laplace transform to recover  $p_t$  from  $U^\alpha$ .

3.7 THEOREM. Let  $U^\alpha$  be an Engelbert resolvent. Then there exists a unique Markov transition function  $p_t$  such that the following hold:

- (i) For all  $\alpha > 0$ ,  $f \in \mathcal{B}$ :  $U^\alpha f = \int_0^\infty e^{-\alpha t} P_t f dt$ .
- (ii) For all  $f \in \mathcal{R}_+$ , there exists  $\alpha > 0$  such that  $e^{-\alpha t} P_t f \leq f$ .
- (iii) For all  $x \in E$ ,  $f \in \mathcal{R}$ , the mapping  $t \rightarrow P_t f(x)$  is right continuous on  $[0, \infty)$ .

*Proof.* We will need several general properties of resolvents and  $\alpha$ -supermedian functions. For convenience, we have summarized these properties in the appendix. Fix  $f \in \mathcal{R}_+$ . Then there is an  $\alpha > 0$  such that  $f$  is  $\alpha$ -supermedian. According to 5.3 the  $\lim_{\beta \rightarrow \infty} (\beta - \alpha)U^\beta f(x)$  exists. We denote the limit by  $\hat{f}(x)$ . Fix  $x \in E$  and put  $g(\beta) = \hat{f}(x) - (\beta - \alpha)U^\beta f(x)$ . By 5.3,  $g$  is nonnegative and by (5.2)

$$(-1)^n \left( \frac{d}{d\beta} \right)^n g(\beta) = n!(U^\beta)^n [1 - (\beta - \alpha)U^\beta]f.$$

Since  $f$  is  $\alpha$ -supermedian and  $U^\beta$  is a positive operator, we see that  $g$  is completely monotone. Also  $g(0) = g(0+) = \hat{f}(x) + \alpha U^0 f(x)$ , which might be infinite ( $U^0 f(x) = \lim_{\beta \downarrow 0} U^\beta f(x)$  exists for every  $f \in \mathcal{B}^+$ ). By the Hausdorff-Bernstein-Widder Theorem (see e.g. [10], p.439) there exists a positive measure  $\lambda_{x,\alpha}(f, \cdot)$  on  $[0, \infty)$  of total mass  $\hat{f}(x) + \alpha U^0 f(x)$  such that, for  $\beta \geq 0$ ,

$$3.8 \quad \hat{f}(x) - (\beta - \alpha)U^\beta f(x) = \int_{[0, \infty)} e^{-\beta t} \lambda_{x,\alpha}(f, dt).$$

Note that  $\lambda_{x,\alpha}(f, \cdot)$  does not charge  $\{0\}$ :

$$\begin{aligned} \lambda_{x,\alpha}(f, \{0\}) &= \lim_{\beta \rightarrow \infty} \int_{[0, \infty)} e^{-\beta t} \lambda_{x,\alpha}(f, dt) \\ &= \lim_{\beta \rightarrow \infty} [\hat{f}(x) - (\beta - \alpha)U^\beta f(x)] \\ &= 0. \end{aligned}$$

Put, for  $f \in \mathcal{R}_+$ ,

$$3.9 \quad P_t^{(\alpha)} f(x) = e^{\alpha t} \int_{(t, \infty)} e^{-\alpha s} \lambda_{x,\alpha}(f, ds).$$

Then for each  $x \in E$ , the mapping  $t \rightarrow e^{-\alpha t} P_t^{(\alpha)} f(x)$  is decreasing, right continuous and

$$e^{-\alpha t} P_t^{(\alpha)} f(x) \uparrow \int_{(0, \infty)} e^{-\alpha s} \lambda_{x,\alpha}(f, ds) = \hat{f}(x)$$

as  $t \downarrow 0$ , since  $\lambda_{x,\alpha}(f, \cdot)$  does not charge  $\{0\}$ . Moreover, for  $\beta > 0$ ,

$$\begin{aligned}
\int_0^\infty e^{-\beta t} P_t^{(\alpha)} f(x) dt &= \int_0^\infty e^{-\beta t} e^{\alpha t} \int_t^\infty e^{-\alpha s} \lambda_{x,\alpha}(f, ds) dt \\
&= \int_0^\infty \int_0^s e^{-(\beta-\alpha)t} dt e^{-\alpha s} \lambda_{x,\alpha}(f, ds) \\
&= \frac{1}{\beta-\alpha} \int_0^\infty [1 - e^{-(\beta-\alpha)s}] e^{-\alpha s} \lambda_{x,\alpha}(f, ds) \\
&= \frac{1}{\beta-\alpha} [\hat{f}(x) - \hat{f}(x) + (\beta-\alpha)U^\beta f(x)] \\
&= U^\beta f(x).
\end{aligned}$$

Hence  $P_t^{(\alpha)}$  does not depend on  $\alpha$  (by the uniqueness of the Laplace transform) and so we will now denote it simply by  $P_t$ .

Consider  $f_1, f_2, \dots, f_n \in \mathcal{R}_+$  and nonnegative real numbers  $r_1, r_2, \dots, r_n$ . Put  $f = r_1 f_1 + \dots + r_n f_n$ . Choose  $\alpha$  sufficiently large so that  $f_1, f_2, \dots, f_n$  are all  $\alpha$ -supermedian. Then  $f$  is also  $\alpha$ -supermedian. Since

$$\hat{f}(x) - (\beta - \alpha)U^\beta f(x) = \sum_{i=1}^n r_i \left[ \hat{f}_i - (\beta - \alpha)U^\beta f_i(x) \right], \quad \beta > 0,$$

it follows from the uniqueness of Laplace transforms that  $\lambda_{x,\alpha}(f, \cdot) = \sum_{i=1}^n r_i \lambda_{x,\alpha}(f_i, \cdot)$ . It then follows from (3.9) that  $P_t f(x) = \sum_{i=1}^n r_i P_t f_i(x)$ . Hence  $P_t$  is a cone map on  $\mathcal{R}_+$ . Extend by linearity the definition of  $P_t$  to  $\mathcal{R} = \mathcal{R}_+ - \mathcal{R}_+$ . Denote the extension again by  $P_t$ . It is easy to see that this extension is well defined and that  $f \rightarrow P_t f(x)$  is a linear functional on  $\mathcal{R}$ . Moreover, by linearity,  $t \rightarrow P_t f(x)$  is right continuous and

$$U^\beta f(x) = \int_0^\infty e^{-\beta t} P_t f(x) dt, \quad f \in \mathcal{R}.$$

Using the first property in (5.2), we see that

$$\begin{aligned}
\left( \frac{d}{d\beta} \right)^n \left[ \frac{f}{\beta} - U^{\beta+\alpha} f \right] &= \frac{(-1)^n n!}{\beta^{n+1}} f - n! (-1)^n (U^{\beta+\alpha})^{n+1} f \\
&= \frac{(-1)^n n!}{\beta^{n+1}} n! [f - (\beta U^{\beta+\alpha})^n f].
\end{aligned}$$

For  $f \in \mathcal{R}_+$ , choose  $\alpha$  so that  $f$  is  $\alpha$ -supermedian. Then we can iterate the inequality  $\beta U^{\beta+\alpha} f \leq f$  to conclude that  $\frac{f}{\beta} - U^{\beta+\alpha} f$  is completely monotone. Since it is clearly the Laplace transform of  $f - e^{-\alpha t} P_t f$ , we see that  $e^{-\alpha t} P_t f \leq f$  for all  $t \geq 0$ .

Since  $\mathcal{R}$  contains the constant functions and  $U^\alpha 1 = 1/\alpha$ , we see that  $P_t 1 = 1$  for all  $t \geq 0$ .

If  $f \in \mathcal{R}$  and  $f$  is nonnegative, then  $\left(\frac{d}{d\beta}\right)^n U^\beta f = (-1)^n n! (U^\beta)^{n+1} f$  and since  $(U^\beta)^{n+1} f \geq 0$  we see that  $\beta \rightarrow U^\beta f$  is completely monotone. Since  $U^\beta f$  is the Laplace transform of  $P_t f$ , we see that  $P_t f \geq 0$  for all  $t \geq 0$ . Hence  $P_t$  is a positive linear operator on  $\mathcal{R}$  and, since  $P_t 1 = 1$ , it is bounded and therefore continuous. We can now extend the definition of  $P_t$  by continuity to  $\text{cl}(\mathcal{R})$ . Fix  $t$  and  $x$  and consider the linear functional  $f \rightarrow P_t f(x)$  defined on  $C_0(E)$ . Let  $f_n$  be a sequence of nonnegative functions in  $C_0(E)$  that decreases to zero pointwise. Then by Dini's lemma  $f_n$  converges uniformly to zero. Hence  $P_t f_n(x)$  converges to zero and so we can apply Daniel's theorem to conclude that there exists a unique measure  $p_t(x, \cdot)$  on  $\mathcal{B}$  such that  $P_t f(x) = \int p_t(x, dy) f(y)$  for all  $f \in C_0(E)$ .

By Proposition A.4, the map  $(t, x) \rightarrow P_t f(x)$  is  $\mathcal{B}([0, \infty)) \times \mathcal{B}$  measurable for each  $f \in \mathcal{R}$ . Applying the lattice version of the monotone class theorem, it follows that  $(t, x) \rightarrow P_t f(x)$  is  $\mathcal{B}([0, \infty)) \times \mathcal{B}$  measurable for every  $f \in \mathcal{B}$  and also that

$$U^\alpha f(x) = \int_0^\infty e^{-\alpha t} P_t f(x) dt$$

for every  $f \in \mathcal{B}$ .

Finally, we must show that  $P_t$  forms a semigroup. Again by a monotone class argument, it suffices to show that  $P_t P_s f = P_{t+s} f$  for all  $f \in \mathcal{R}_+$ , and  $s, t > 0$ . Fix  $f \in \mathcal{R}_+$ . Then  $P_t P_s f$  and  $P_{t+s} f$  are both right continuous in  $s$  and so it suffices to show that, for each  $t \geq 0$ , they have the same Laplace transform on  $s$ . That is,

$$P_t U^\alpha f(x) = \int_0^\infty e^{-\alpha s} P_{t+s} f(x) ds.$$

Since  $U^\alpha f \in \mathcal{R}$ , the left-hand side is right continuous in  $t$ . So is the right-hand side. Hence, it suffices to check that they have the same Laplace transform on  $t$ . Thus

$$\begin{aligned} U^\beta U^\alpha f(t) &= \int_0^\infty e^{-\beta t} \int_0^\infty e^{-\alpha t} P_{t+s} f(x) ds dt \\ &= \int_0^\infty \int_0^\infty e^{-(\beta-\alpha)t} e^{-\alpha s} P_s f(x) ds dt \\ &= \int_0^\infty e^{-\alpha s} \left[ \int_0^\infty e^{-(\beta-\alpha)t} dt \right] P_s f(x) ds. \end{aligned}$$

Since

$$\int_0^s e^{-(\beta-\alpha)t} dt = \begin{cases} \frac{1}{\beta-\alpha} [1 - e^{-(\beta-\alpha)s}] & \beta \neq \alpha \\ s & \beta = \alpha \end{cases}$$

we must check that

$$U^\beta U^\alpha f = \begin{cases} \frac{1}{\beta-\alpha} [U^\alpha f - U^\beta f] & \beta \neq \alpha \\ (U^\alpha)^2 f & \beta = \alpha \end{cases}$$

which is true since  $U^\alpha$  is a resolvent. This completes the proof.  $\square$

The transition function constructed in Theorem 3.7 will be called an *Engelbert transition function*. Let  $\mathcal{W}$  denote the space of all functions from  $\mathbb{R}^+ = [0, \infty)$  into  $E$  that are right continuous and have left-hand limits. Put  $Y_t(w) = w(t)$  for  $w \in \mathcal{W}$  and let  $\mathcal{G}^0 = \sigma\{Y_t | t \geq 0\}$  and  $\mathcal{G}_t^0 = \sigma\{Y_s | 0 \leq s \leq t\}$ .

**3.10 THEOREM.** Let  $p_t$  be an Engelbert transition function. Then, for each probability measure  $\mu$  on  $(E, \mathcal{B})$ , there exists a probability measure  $\mathbb{P}_\mu$  on  $(\mathcal{W}, \mathcal{G}^0)$  such that

- (i)  $\mathbb{E}_\mu\{f(Y_{t+s}) | \mathcal{G}_s^0\} = P_t f(Y_s)$ , a.s.  $\mathbb{P}_\mu$  for  $f \in \mathcal{B}$  and all  $s, t \geq 0$ ;
- (ii)  $\mathbb{E}_\mu f(Y_t) = \int \mu(dx) P_t f(x)$ , for  $f \in \mathcal{B}$  and all  $t \geq 0$

( $\mathbb{E}_\mu Z$  denotes the expectation of  $Z(w)$  with respect to the probability measure  $\mathbb{P}_\mu$ ).

*Proof.* The proof is quite standard and we will present essentially the same outline as in [11], Theorem 5.1. Let  $\Omega = E^{[0, \infty)}$  and  $\mathcal{F} = \mathcal{B}^{[0, \infty)}$ . Put  $X_t(\omega) = \omega(t)$  and  $\mathcal{F}_t = \sigma\{X_s | 0 \leq s \leq t\}$ . Given  $\mu$ , we can apply the Kolmogorov extension theorem to obtain a measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$  such that  $(X_t, \mathcal{F}_t, \mathbb{P})$  satisfies properties (i) and (ii).

We must show that this process has a modification that is right continuous and has left-hand limits. By Theorem 3.7-(ii), for each  $f \in \mathcal{R}_+$ , there exists an  $\alpha > 0$  such that  $e^{-\alpha t} P_t f \leq f$ . This fact together with (i) implies that  $e^{-\alpha t} f(X_t)$  is an  $\mathcal{F}_t$ -supermartingale. Hence, with  $\mathbb{P}$ -probability one, the map  $t \rightarrow f(X_t)$  restricted to the positive rationals  $\mathbb{Q}^+$  has left- and right-hand limits at each point in  $\mathbb{R}^+$ . Since  $\text{cl}(\mathcal{R}) \supset C_0(E)$ , it follows that for each  $f \in C_0(E)$ , with  $\mathbb{P}$ -probability one, the map  $t \rightarrow f(x_t)$  restricted to  $\mathbb{Q}^+$  has left and right-hand limits at each point in  $\mathbb{R}^+$ . The space  $C_0(E)$  contains a countable subset  $(f_n)_{n \geq 0}$  that separates the points in  $E$ . Let  $N_n$  be a  $\mathbb{P}$ -null set off of which  $f_n(X_t)$  restricted to  $\mathbb{Q}^+$  has left- and right-hand limits at each point of  $\mathbb{R}^+$ . Put  $N = \bigcup_n N_n$ . Since  $(f_n)_{n \geq 0}$  separates points, it follows that, on  $\Omega \setminus N$ , the map  $t \rightarrow X_t$ , restricted to  $\mathbb{Q}^+$  has right- and left-hand limits in  $E$  at each point in  $\mathbb{R}^+$ . Take  $\Omega \setminus N$  as our new sample space  $\Omega$ . For each  $t \in \mathbb{R}^+$  and  $\omega \in \Omega$ , put

$$X_{t+}(\omega) = \lim_{\substack{r \downarrow t \\ r \in \mathbb{Q}^+}} X_r(\omega).$$

Fix  $f, g \in \mathcal{R}_+$ ,  $t \in \mathbb{R}^+$  and let  $r_n$  be a sequence in  $\mathbb{Q}^+$  such that  $r_n \downarrow t$ . Then, applying (ii) we get

$$\begin{aligned}
 \mathbb{E}f(X_t)g(X_{t+}) &= \lim_{n \rightarrow \infty} \mathbb{E}f(X_t)g(X_{r_n}) \\
 &= \lim_{n \rightarrow \infty} \mathbb{E}f(X_t)P_{r_n-t}g(X_t) \\
 3.11 \qquad &= \mathbb{E}f(X_t)\hat{g}(X_t)
 \end{aligned}$$

where  $\hat{g}(x) = \lim_{t \downarrow 0} P_t g(x) = P_0 g(x)$ . Since  $e^{-\alpha t} P_t g \leq g$  (for some  $\alpha > 0$ ) and  $g \geq 0$ , it follows that  $0 \leq \hat{g} \leq g$ . Since  $P_t g = P_t P_0 g = P_t \hat{g}$ , we see that

$$\mathbb{E}g(X_t) = \int \mu(dx) P_t g(x) = \int \mu(dx) P_t \hat{g}(x) = \mathbb{E}\hat{g}(X_t).$$

Hence  $g(X_t) = \hat{g}(X_t)$   $\mathbb{P}$ -a.s. and so (3.11) implies that

$$3.12 \qquad \mathbb{E}f(X_t)g(X_{t+}) = \mathbb{E}f(X_t)g(X_t).$$

Since  $\text{cl}(\mathcal{R}) \supset C_0(E)$ , (3.12) actually holds for all  $f, g \in C_0(E)$ . A monotone class argument now shows that  $\mathbb{E}F(X_t, X_{t+}) = \mathbb{E}F(X_t, X_t)$  for all  $F \in \mathcal{B} \times \mathcal{B}$ . Taking  $F(x, y) = 1_{x=y}$ , we see that

$$\mathbb{P}(X_t = X_{t+}) = 1.$$

Now define  $Y_t(\omega) = X_{t+}(\omega)$  for each  $t \in \mathbb{R}^+$ ,  $\omega \in \Omega$ . Then the map  $t \rightarrow Y_t(\omega)$  is right continuous on  $\mathbb{R}^+$  and has left limits on  $(0, \infty)$ . Put  $\mathcal{H}_t = \sigma(Y_s | s \leq t)$ . Since  $Y_t = X_t$   $\mathbb{P}$ -a.s., it follows that  $(Y_t, \mathcal{H}_t, \mathbb{P})$  satisfies (i) and (ii). Finally, we define a map  $\pi$  from  $\Omega$  into  $\mathcal{W}$  by the formula  $\pi(\omega)(t) = Y_t(\omega)$ . Then the measure  $\mathbb{P}_\mu = \mathbb{P} \circ \pi^{-1}$  has all the properties stated in the theorem.  $\square$

Formula (3.10)-(i) is the basic form of the Markov property. We now extend this property in the usual way. Indeed, for each  $t \geq 0$  let  $\theta_t$  denote the *shift operator* in  $\mathcal{W}$ :

$$(\theta_t w)(s) = w(t + s).$$

For each probability  $\mu$  on  $(E, \mathcal{B})$  let  $\mathcal{G}^\mu$  denote the completion of  $\mathcal{G}^0$  with respect to  $\mathbb{P}_\mu$  and let  $N^\mu = \{B \in \mathcal{G}^\mu \mid \mathbb{P}_\mu(B) = 0\}$ . Put  $\mathcal{G}_t^\mu = \sigma\{\mathcal{G}_t^0 \cup N^\mu\}$ ,  $\mathcal{G}_t = \bigcap_{\mu} \mathcal{G}_t^\mu$  and  $\mathcal{G} = \bigcup_{\mu} \mathcal{G}^\mu$ . For any bounded  $\mathcal{G}$ -measurable function  $Z$ , put

$$\theta_t Z(w) = Z(\theta_t w).$$

Using a standard monotone class argument, formula ( 3.10)-(i) can be extended as follows:

$$3.13 \quad \mathbb{E}_\mu\{\theta_t Z | \mathcal{G}_t\} = \mathbb{E}_{Y_t} Z, \quad Z \in \mathcal{G}, t \geq 0.$$

Formula ( 3.13) is called the *Markov property*.

A map  $\tau$  from  $\mathcal{W}$  to  $[0, \infty]$  is called a *stopping time* if  $\{w \mid \tau(w) \leq t\} \in \mathcal{G}_t$  for all  $t \geq 0$ . If  $\tau$  is a stopping time, we define the *pre- $\tau$   $\sigma$ -algebra*  $\mathcal{G}_\tau$  to be the collection of all sets  $A \in \mathcal{G}$  such that  $A \cap \{\tau \leq t\} \in \mathcal{G}_t$  for all  $t \geq 0$ . The following theorem extends ( 3.13) to stopping times.

3.14 THEOREM. Let  $\mathbb{P}_\mu$  be the extension to  $(\mathcal{W}, \mathcal{G})$  of the measure constructed in Theorem 3.10. Then, for each stopping time  $\tau$  and for each  $Z \in \mathcal{G}$ ,

$$3.15 \quad \mathbb{E}_\mu\{\theta_\tau Z | \mathcal{G}_\tau\} = \mathbb{E}_{Y_\tau} Z \quad \text{a.s. } \mathbb{P}_\mu \text{ on } \{\tau < \infty\}.$$

Also  $\mathcal{G}_t = \bigcap_{s>t} \mathcal{G}_s$ .

*Proof.* The proof follows immediately from the results proved in [1] Sections 1.6–1.8. We only remark that in Theorem 1.8.11 of [1] we must replace his  $\mathbb{L}$  by our class  $\mathcal{R}$ . Then in his proof, the monotone class theorem can still be used since  $\sigma(\mathcal{R}) = \mathcal{B}$ .  $\square$

Formula ( 3.15) is the *strong Markov property*.

#### 4. EXAMPLES

4.1. **Uniform Motion.** Let  $(E, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  and let  $p_t$  be the transition function for uniform motion:

$$p_t(x, B) = 1_B(x - t).$$

Then

$$P_t f(x) = f(x - t)$$

and

$$U^\alpha f(x) = e^{-\alpha x} \int_{-\infty}^x e^{\alpha y} f(y) dy.$$

It is clear that, for every  $f \in \mathcal{B}$ ,  $U^\alpha f \in C(\mathbb{R})$ . Since  $C(\mathbb{R})$  is a linear lattice, it follows that  $\mathcal{R} \subset C(\mathbb{R})$ .

Now suppose that  $g \in C^1(\mathbb{R})$  and put  $f = g' + \alpha g$ . Then, a simple calculation shows that  $U^\alpha f = g$ . Hence,  $C^1(\mathbb{R}) \subset \mathcal{C} \subset \mathcal{R} \subset C(\mathbb{R})$ . The classes  $C^1(\mathbb{R})$  and  $C(\mathbb{R})$  generate



the same uniformity and this uniformity is also generated by  $\mathcal{E} = \{e(x) = x\}$ . Hence the intrinsic state space coincides with the real line.

**4.2. Brownian Motion.** Let  $(E, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  and let  $p_t$  be the transition function for  $d$ -dimensional Brownian motion:

$$p_t(x, B) = \int p_t(y - x)1_B(y)dy$$

where

$$p_t(y) = (2\pi t)^{-d/2} \exp\{-|y|^2/2t\}.$$

Then

$$P_t f(x) = \int p_t(y - x)f(y)dy.$$

It is easy to check that for every  $\epsilon > 0$  there exists a  $K < \infty$  such that:

$$|p_t(y - x) - p_t(y - x_0)| \leq Kp_{2t}(y - x_0),$$

whenever  $|x - x_0| < \epsilon$ . Hence, the dominated convergence theorem shows that for every  $f \in \mathcal{B}$ ,  $P_t f$  belongs to  $C(\mathbb{R}^d)$ . Put  $g = U^\alpha f$ . By Proposition 2.10-(i),  $\lim_{t \downarrow 0} \|P_t g - g\| = 0$ . Since  $x \rightarrow P_t g(x)$  is continuous, it follows that  $g$  belongs to  $C(\mathbb{R}^d)$ . Thus,  $\mathcal{C} \subset C(\mathbb{R}^d)$  and since  $C(\mathbb{R}^d)$  is a linear lattice we have  $\mathcal{C} \subset \mathcal{R} \subset C(\mathbb{R}^d)$ .

Let  $g$  be a function of class  $C^2(\mathbb{R}^d)$  and let  $f = -\frac{1}{2}\Delta g + \alpha g$  where  $\Delta = \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_d^2}$  is the Laplacian in  $\mathbb{R}^d$ . If we integrate by parts (in  $x$ ), then use the fact that  $\frac{1}{2}\Delta p_t = \frac{d}{dt}p_t$ , and integrate by parts again (in  $t$ ), we get

$$\int_s^\infty e^{-\alpha t} P_t f(x) dt = e^{-\alpha s} \int p_s(y - x)g(y)dy.$$

Letting  $s \downarrow 0$  and using the continuity of  $g$  we get  $U^\alpha f = g$ . Hence,  $C^2(\mathbb{R}^d) \subset \mathcal{C} \subset \mathcal{R} \subset C(\mathbb{R}^d)$ . Since the classes  $C^2(\mathbb{R}^d)$  and  $C(\mathbb{R}^d)$  generate the same uniformity on  $\mathbb{R}^d$  we see that we may take

$$\mathcal{E} = \{e_1(x) = x_1, e_2(x) = x_2, \dots, e_d(x) = x_d\}.$$

Hence the intrinsic state space coincides with  $\mathbb{R}^d$ .

**4.3. Pure Jump Process.** Let  $(E, \mathcal{B})$  be a general state space, let  $q$  be a stochastic kernel on  $(E, \mathcal{B})$  and put

$$p_t(x, B) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} q^{(n)}(x, B).$$

Then

$$P_t f(x) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} Q^n f(x)$$

and

$$U^\alpha f(x) = \frac{1}{\alpha + \lambda} \sum_{n=0}^{\infty} \left( \frac{\lambda}{\alpha + \lambda} \right)^n Q^n f(x)$$

where  $Q$  is the operator on  $\mathcal{B}$  corresponding to  $q$ .

Suppose that  $g \in \mathcal{B}$  and put  $f = \lambda(1 - Q)g + \alpha g$ . Then a simple calculation shows that  $U^\alpha f = g$ . Hence  $\mathcal{B} \subset \mathcal{C}$  and so it follows that  $\mathcal{B} = \mathcal{C} = \mathcal{R}$ . The class  $\mathcal{B}$  generates a finite dimensional uniformity if and only if  $E$  is a countable set. In this case, if points are measurable, then we can take  $E = \{e(x_n) = n\}$  which maps  $E$  into the set of nonnegative integers.

**4.4. Uniform Motion with Sticky Origin.** Let  $(E, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  and let  $p_t$  be the transition function for uniform motion with “stickum” at zero (discussed in the Introduction):

$$p_t(x, B) = \begin{cases} e^{x-t} 1_B(0) + \int_0^{t-x} e^{-u} 1_B(x-t+u) du & 0 \leq x < t \\ 1_B(x-t) & \text{else.} \end{cases}$$

Then

$$U^\alpha f(x) = \begin{cases} \frac{e^{-\alpha x}}{1 + \alpha} \left[ f(0) + \int_{-\infty}^0 e^{\alpha t} f(t) dt + (1 + \alpha) \int_0^x e^{\alpha t} f(t) dt \right] & x \geq 0 \\ e^{-\alpha x} \int_{-\infty}^x e^{\alpha t} f(t) dt & x < 0. \end{cases}$$

It is easy to see from this formula that if  $f \in \mathcal{B}$  then  $U^\alpha f$  belongs to  $C((-\infty, 0), [0, \infty))$ . Since this last class is a linear lattice, it follows that  $\mathcal{R} \subset C((-\infty, 0), [0, \infty))$ .

Now suppose that  $g \in C^1((-\infty, 0), [0, \infty))$  and  $g'(0) = 0$  (this is the right-hand derivative at zero). Put  $f = g' + \alpha g$ . Then a simple integration by parts shows that  $U^\alpha f = g$ . Hence,

$$C^1((-\infty, 0), [0, \infty)) \cap \{h \mid h'(0) = 0\} \subset \mathcal{C} \subset \mathcal{R} \subset C((-\infty, 0), [0, \infty))$$

where by  $h'(0)$  we mean the right-hand derivative at zero. It is easy to see that we can take  $\mathcal{E}$  consisting of one function

$$e(x) = \begin{cases} x & x \geq 0, \\ x - 1 & x < 0. \end{cases}$$

This function makes small negative numbers “far away” from small positive numbers. Also the space  $\psi(E) = (-\infty, -1) \cup [0, \infty)$  is not closed and so the intrinsic state space  $E' = (-\infty, -1] \cup [0, \infty)$  turns out to have one point more than  $E$ .

**4.5. Uniform Motion with a Jump.** Let  $E = (-\infty, -1] \cup [0, \infty) = E^- \cup E^+$  and let  $\mathcal{B}$  be the Borel  $\sigma$ -algebra on  $E$ . Let  $p_t$ , be “severed” uniform motion:

$$p_t(x, B) = \begin{cases} 1_B(x - t - 1) & 0 \leq x < t, \\ 1_B(x - t) & \text{else.} \end{cases}$$

Then

$$P_t f(x) = \begin{cases} f(x - t - 1) & 0 \leq x < t, \\ f(x - t) & \text{else,} \end{cases}$$

and

$$U^\alpha f(x) = \begin{cases} e^{-\alpha x} \left[ e^\alpha \int_{-\infty}^{-1} e^{\alpha y} f(y) dy + \int_0^x e^{\alpha y} f(y) dy \right] & x \in E^+, \\ e^{-\alpha x} \int_{-\infty}^x e^{\alpha y} f(y) dy & x \in E^-. \end{cases}$$

Note that  $U^\alpha f(0) = U^\alpha f(-1)$ . Put  $I_{-1,0} = \{h \mid h(-1) = h(0)\}$ . We see that if  $f \in \mathcal{B}$  then  $U^\alpha$  belongs to  $C(E) \cap I_{-1,0}$ .

Now suppose that  $g \in C^1(E) \cap I_{-1,0}$  and put  $f = g' + \alpha g$ . Again, integrating by parts shows immediately that  $g = U^\alpha f$ . Hence

$$C^1(E) \cap I_{-1,0} \subset \mathcal{C} \subset \mathcal{R} \subset C(E) \cap I_{-1,0}.$$

In this example,  $\mathcal{R}$  does not separate  $-1$  and  $0$ . We may choose for  $\mathcal{E}$  the collection consisting of the single function

$$e(x) = \begin{cases} x & x \in E^+ \\ x + 1 & x \in E^- \end{cases}$$

On the intrinsic state space the process is just uniform motion.

**4.6. Uniform Motion with Fork in the Road.** In  $\mathbb{R}^2$ , let  $R$  denote the closed right half-line on the  $x$ -axis, let  $U$  denote the open upper half-line on the  $y$ -axis, and let  $L$  denote the open lower half-line on the  $y$ -axis. Let  $E = R \cup U \cup L$  and let  $\mathcal{B}$  be the Borel  $\sigma$ -algebra on  $E$ . Let  $p_t$  be the transition function for the process that proceeds

deterministically left, up, or down on  $R$ ,  $U$ , or  $L$ , respectively. When it arrives at the origin, it proceeds upward/downward each with probability  $1/2$ . Then

$$P_t f(x, y) = \begin{cases} f(0, y + t) & (x, y) \in U, \\ f(0, y - t) & (x, y) \in L, \\ f(x - t, 0) & (x, y) \in R, 0 < t \leq x, \\ \frac{1}{2}f(0, t - x) + \frac{1}{2}f(0, x - t) & (x, y) \in R, 0 \leq x < t, \end{cases}$$

and

$$U^\alpha f(x, y) = \begin{cases} e^{\alpha y} \int_y^\infty e^{-\alpha t} f(0, t) dt & (x, y) \in U, \\ e^{-\alpha y} \int_{-\infty}^y e^{\alpha t} f(0, t) dt & (x, y) \in L, \\ e^{-\alpha x} \int_0^x e^{\alpha t} f(t, 0) dt \\ \quad + \frac{1}{2} e^{-\alpha x} \int_0^\infty e^{\alpha t} [f(0, t) + f(0, -t)] dt & (x, y) \in R. \end{cases}$$

From this formula for  $U^\alpha$ , we see that  $\mathcal{C} \subset \mathcal{R} \subset C(R, U, L)$ .

Now suppose that  $g \in C^1(R, U, L)$  and  $g(0, 0) = \frac{1}{2} [g(0, 0+) + g(0, 0-)]$ . Put

$$f = \begin{cases} -\frac{\partial g}{\partial x_2} + \alpha g & (x, y) \in U \\ \frac{\partial g}{\partial x_2} + \alpha g & (x, y) \in L \\ \frac{\partial g}{\partial x_1} + \alpha g & (x, y) \in R. \end{cases}$$

Then, as in the previous examples, simple calculations show that  $g = U^\alpha f$ . Hence  $g \in \mathcal{C}$ . Since the collection of all such functions  $g$  generates the same uniformity as  $C(R, U, L)$  does, we see that we can take  $E = \{e_1(x, y), e_2(x, y)\}$  where

$$e_1(x, y) = 1_R(x, y)$$

$$e_2(x, y) = \begin{cases} y + 1 & (x, y) \in U \\ y - 1 & (x, y) \in L \\ 0 & (x, y) \in R. \end{cases}$$

Hence  $\psi(E) = \{(x, y) \mid x \geq 0, y = 0 \text{ or } x = 0, y > 1 \text{ or } x = 0, y < -1\}$  and so the closure  $\mathcal{E}^1 = \overline{\psi(E)}$  contains two new points:  $(0, 1)$  and  $(0, -1)$ . Since the intrinsic strong Markov process is right continuous we see that, with probability one, it visits one of these two new points and it never visits the branch point  $(0, 0)$ .

**4.7. Brownian Motion with an Absorbing State.** Let  $(E, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  and let  $p_t$  be the transition function for  $d$ -dimensional Brownian motion except it is altered at  $x = 0$  as follows:

$$p_t(x, B) = \begin{cases} \int p_t(y - x) 1_B(y) dy & x \neq 0 \\ 1_B(0) & x = 0 \end{cases}$$

where  $p_t(y)$  is as defined in Subsection 4.2. Since the origin is a set of measure zero, it is easy to check that this  $p_t$  is a Markov transition function. It is also easy to see that  $\mathcal{R}$  coincides with the  $\mathcal{R}$  for Brownian motion except that the value of each function at zero can be changed arbitrarily. Hence we may take

$$\mathcal{E} = \{e_1(x) = x_1, e_2(x) = x_2, \dots, e_d(x) = x_d, e_{d+1}(x) = 1_0(x)\}.$$

Hence the intrinsic state space is a closed subset of  $\mathbb{R}^{d+1}$  that consists of two parts: the  $d$ -dimensional hyperplane corresponding to  $x_{d+1} = 0$  and the single point corresponding to  $x_{d+1} = 1$  with the rest of the coordinates vanishing. On the hyperplane the transition function is that for ordinary Brownian motion. The single point is an absorbing point.

**4.8. Instantaneous Jump to Limiting Distribution.** Let  $(E, \mathcal{B})$  be a general measurable space and let  $p_t(x, B) = \pi(B)$  where  $\pi$  is a probability measure on  $(E, \mathcal{B})$ . Then

$$P_t f(x) = \int f(y) \pi(dy)$$

and

$$U^\alpha f(x) = \frac{1}{\alpha} \int f(y) \pi(dy).$$

Hence, for  $f \in \mathcal{B}$ ,  $U^\alpha f$  is a constant function. Hence, both  $\mathcal{C}$  and  $\mathcal{R}$  coincide with the collection of all constant functions. In this example, every point gets identified as the same. We may take  $\mathcal{E}$  consisting of one function

$$e(x) = 1.$$

In this case, the entire state space collapses to a single point.

## 5. APPENDIX

In this section, we collect certain results that are either well-known or secondary to the main theme of the paper.

**5.1 PROPOSITION.** Let  $E$  be a closed subset of  $\mathbb{R}^d$  and let  $\mathcal{H}$  be a class of continuous real-valued functions defined on  $E$ . Then the  $\sigma$ -algebra generated by  $\mathcal{H}$  coincides with the Borel  $\sigma$ -algebra on  $E$  if and only if  $\mathcal{H}$  separates points in  $E$ .

*Proof.* The “only if” direction is trivial. Suppose that  $\mathcal{H}$  separates points in  $E$ . For  $a \in \mathbb{R}^d$  and  $r > 0$ , let  $B_r(a)$  denote the intersection of  $E$  and the closed ball of radius  $r$  centered at  $a$ . Fix  $a \in E$  and  $0 < r_0 < r_1$ . Put  $A_n = \{x \in E \mid nr_1 \leq |x - a| \leq (n + 1)r_1\}$ . Fix  $x \in B_{r_0}(a)$  and  $y \in B_{r_0}(a)^c$ . Let  $f$  be a function in  $\mathcal{H}$  for which  $f(x) \neq f(y)$ . Suppose that  $f(x) < f(y)$ . Pick  $s$  such that  $f(x) < s < f(y)$  and let  $V_{x,y} = f^{-1}((-\infty, s))$  and  $W_{x,y} = f^{-1}((-\infty, s])$ . Then, in the relative topology,  $V_{x,y}$  is open,  $W_{x,y}$  is closed,  $x \in V_{x,y} \subset W_{x,y}$  and  $y \in W_{x,y}^c$ . The collection  $\{V_{x,y} \mid x \in B_{r_0}(a)\}$  forms an open cover of the compact set  $B_{r_0}$  and hence there is a finite subcover  $V_{x_1,y}, \dots, V_{x_n,y}$ . Put  $F_y = \bigcup_{i=1}^n W_{x_i,y}$ . Then  $F_y$  is a closed set,  $y \in F_y^c$  and  $B_{r_0}(a) \subset F_y$ . Also,  $F_y \in \sigma(\mathcal{H})$ .

Now fix  $n \geq 1$ . The collection  $\{F_y^c \mid y \in A_n\}$  forms an open cover of the compact set  $A_n$ . Hence there is a finite subcover  $\{F_{y_{j,n}}^c \mid j = 1, 2, \dots, k_n\}$ . The collection  $\{F_{y_{j,n}}^c \mid j = 1, 2, \dots, k_n; n = 1, 2, \dots\}$  is a countable collection that covers  $\{x \mid |x - a| \geq r_1\}$ . Hence,

$$B_{r_0}(a) \subset \bigcap_{j,n} F_{y_{j,n}} \subset B_{r_1}(a)$$

and  $\bigcap_{j,n} F_{y_{j,n}}$  is evidently  $\sigma(\mathcal{H})$ -measurable and closed. Since  $r_0$  and  $r_1$  were arbitrary, we see that for any  $0 < r_0 < r_1$  we can find a closed  $\sigma(\mathcal{H})$ -measurable set  $G$  such that  $B_{r_0}(a) \subset G \subset B_{r_1}(a)$ .

Now fix  $r > 0$  and put  $B = B_r(a)$  and  $B_n = B_{r+1/n}(a)$ . There exist closed  $\sigma(\mathcal{H})$  measurable sets  $G_n$  such that  $B \subset G_n \subset B_n$ . Hence  $B \subset \bigcap_n G_n \subset \bigcap_n B_n = B$  and so  $B = \bigcap_n G_n$  is a  $\sigma(\mathcal{H})$ -measurable set. Since  $a$  and  $r$  were arbitrary, we see that  $\{B_r(a) \mid a \in E, r > 0\} \subset \sigma(\mathcal{H})$  and since these sets generate the Borel  $\sigma$ -algebra it follows that  $B(E) \subset \sigma(\mathcal{H})$ . Since every function in  $\mathcal{H}$  is continuous, we see that  $\sigma(\mathcal{H}) \subset \mathcal{B}(E)$ .  $\square$

The following properties follow directly from the resolvent identity ( 2.5).

**5.2 PROPOSITION.** Let  $U^\alpha$  be a resolvent on  $(E, \mathcal{B})$ . Then for each  $f \in \mathcal{B}$  and  $x \in E$ , the map  $\alpha \rightarrow U^\alpha f(x)$  is infinitely differentiable on  $(0, \infty)$  and

$$\left(\frac{d}{d\alpha}\right)^n U^\alpha f = (-1)^n n! (U^\alpha)^{n+1} f$$

and

$$\left(\frac{d}{d\alpha}\right)^n (\alpha U^\alpha f) = (-1)^{n+1} n! (U^\alpha)^n [1 - \alpha U^\alpha] f.$$

The proof can be found in Gettoor [11] page 7. The next proposition is also proved in [11] on pages 5 and 12, respectively.

5.3 PROPOSITION. If  $f$  is  $\beta$ -supermedian, then the map  $\alpha \rightarrow (\alpha - \beta)U^\alpha f(x)$  is increasing for  $\alpha > \beta$  and the limit  $\widehat{f} = \lim_{\alpha \rightarrow \infty} \alpha U^\alpha f$  is  $\beta$ -supermedian. Also  $\widehat{f} \leq f$  and  $U^\alpha \widehat{f} = U^\alpha f$  for all  $\alpha > 0$ .

## REFERENCES

- [1] R. M. Blumenthal and R. K. Gettoor. *Markov Processes and Potential Theory*. Academic Press, New York, 1968. 16
- [2] C. Dellacherie and P. A. Meyer. *Probabilities and Potential*. Hermann, Paris, 1978. 9
- [3] H. J. Engelbert. Markov processes in general state spaces. part I. *Math. Nachr.*, 80:19–36, 1977. 2
- [4] H. J. Engelbert. Markov processes in general state spaces. part II. *Math. Nachr.*, 82:191–203, 1977. 2
- [5] H. J. Engelbert. Markov processes in general state spaces. part III. *Math. Nachr.*, 83:47–71, 1977. 2
- [6] H. J. Engelbert. Markov processes in general state spaces. part IV. *Math. Nachr.*, 84:277–300, 1977. 2
- [7] H. J. Engelbert. Markov processes in general state spaces. part V. *Math. Nachr.*, 85:111–130, 1977. 2
- [8] H. J. Engelbert. Markov processes in general state spaces. part VI. *Math. Nachr.*, 85:235–266, 1977. 2
- [9] H. J. Engelbert. Markov processes in general state spaces. part VII. *Math. Nachr.*, 86:67–83, 1977. 2
- [10] W. Feller. *Introduction to Probability Theory and its Applications. Vol. 2*. Wiley, New York, second edition, 1971. 11
- [11] R. K. Gettoor. *Markov Processes: Ray Processes and Right Processes*, volume 440 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-Heidelberg-New York, 1975. 2, 14, 22
- [12] J. L. Kelley. *General Topology*. Van Nostrand; Springer, 1955. 2
- [13] F. Knight. Note on regularization of Markov-processes. *Illinois Journ. Math.*, 9:548–552, 1965. 2
- [14] D. Ray. Resolvents, transition function, and strongly Markovian processes. *Ann. Math.*, 70:43–75, 1959. 2
- [15] A. A. Yushkevich. On strong Markov processes. *Thy. Prob. Appl.*, 2:181–205, 1957. 2

DEPT. OF MATHEMATICS, UNIV. OF ILLINOIS, CHAMPAIGN-URBANA, IL.