# Lectures on
# Stochastic Control and Nonlinear Filtering

By

**M. H. A. Davis**

**Tata Institute of Fundamental Research**
**Bombay**
**1984**

# Lectures on
# Stochastic Control and Nonlinear Filtering

**By**

**M. H. A. Davis**

Lectures delivered at the
**Indian Institute of Science, Bangalore**
under the
**T.I.F.R.–I.I.Sc. Programme in Applications of Mathematics**

**Notes by**

**K. M. Ramachandran**

Published for the
**Tata Institute of Fundamental Research**

**Springer-Verlag**
Berlin Heidelberg New York Tokyo
**1984**

**Author**

**M. H. A. Davis**
Department of Electrical Engineering
Imperial College of Science and Technology
London SW 7
United Kingdom



ISBN 3-540-13343-7 Springer-Verlag, Berlin. Heidelberg.
New York. Tokyo
ISBN 0-387-13343-7 Springer-Verlag, New York. Heidelberg.
Berlin. Tokyo



Printed by M. N. Joshi at The Book Centre Limited,
Sion East, Bombay 400 022 and published by H. Goetze,
Springer-Verlag, Heidelberg, West Germany

**Printed in India**

# Preface

These notes comprise the contents of lectures I gave at the T.I.F.R. Centre in Bangalore in April/May 1983. There are actually two separate series of lectures, on controlled stochastic jump processes and nonlinear filtering respectively, and the corresponding two parts of these notes are almost disjoint. They are united however, by the common philosophy (if that is not too grand a work for it) of treating Markov processes by methods of stochastic calculus, and I hope the reader will, at least, be convinced of the usefulness of this and of the 'extended generator' concept in doing calculations with Markov precesses.

The first part is aimed at developing optimal control theory for a class of Markov processes called piecewise-deterministic (PD)processes. These were only isolated rather recently but seen general enough to include as special cases practically all the non-diffusion continuous time processes of applied probability. Optimal control for PD processes occupies a curious position just half way between deterministic and Stochastic optimal control theory in such a way that no standard theory from either side is adequate to deal with it. The only applicable theory that exists at all is very recent work of D. Vermes based on the generalized dynamic programming ideas of R.B. Vinter and R.M. Lewis, and this is what I have attempted to describe here. Undoubtedly, further development of control theory for PD processes will be a fruitful field of enquiry.

Part II concentrates on the "pathwise" theory of filtering for diffusion processes and on more sophisticated extensions of it due primarily to H. Kunita. The intriguing point here is to see how stochastic partial differential equations can be dealt with by stochastic flow theory through what amounts to a "doubly stochastic" version of the FeynmanKac formula. Using this, Kunita has given an elegant argument to show the existence of smooth conditional densities under Hörmander-type conditions. This is included. Ultimately, it rests on results obtained by Bismut and others using Malliavin calculus, since one needs a version of the Hörmander theorem which is valid for continuous (rather than $C^\infty$) $t$-dependence of the coefficients. It was unfortunately impossible to go into such questions in the time available.

I would like to thank Professor K.G. Ramanathan for his kind invitation to visit Bangalore and K.M. Ramachandran for his heroic efforts at keeping up-to-date notes on a rapidly accumulating number of lectures, and for preparing the final version of the present text. I would also like to thank the students and staff of the T.I.F.R. Centre and of the I.I.Sc. Guest House for their friendly hospitality which made my visit such a pleasant one.

# Contents

# Part I

# Stochastic Jump Processes and Applications

# Chapter 1

# Stochastic Jump Processes

## 0 Introduction

Stochastic jump processes are processes with piecewise constant paths. **1**
The Poisson process, the processes arising in Inventory problems (stocks of items in a store with random ordering and replacement) and queuing systems (arrivals at a queue with each customer having random demand for service) are examples of stochastic jump processes. Our aim here is to develop a theory suitable for studying optimal control of such processes.

In Section 1, martingale theory and stochastic calculus for jump processes are developed. Gnedenko-Kovalenko [16] introduced piecewise-linear process. As an example of such a process, consider virtual waiting time process ($VWT$) for queueing systems, where $VWT(t)$ is the time customer arriving at time $t$ would have to wait for service, see Fig. (0.1).

Later Davis [7] and Vermes [25] introduced the concept of piecewise deterministic processes which follow smooth curves (not necessarily straight lines) between jumps. In Section 2, we will study some applications to piecewise-deterministic processes. The idea there is to derive Markov properties, Dynkin's formula, infinitesimal generators etc., using the calculus developed in Section 1.

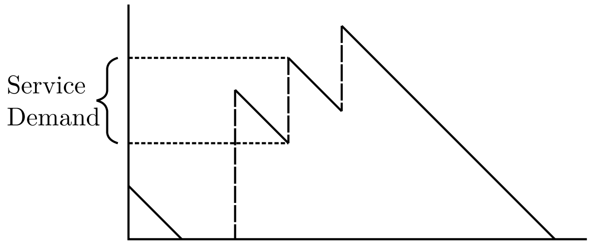


Figure 0.1: Arrival time of customers

2

## 1 Martingale Theory for Jump Processes

Let $(X, S)$ be a Borel space.

**Definition 1.** *A jump process is defined by sequences* $T_1, T_2, T_3, \ldots$, $Z_1, Z_2, Z_3, \ldots$ *of random variables,* $T_i \in \mathbb{R}_+$ *and* $T_{i+1} > T_i$ *a.s. and* $Z_i \in (X, S)$*. Set*

$$T_\infty = \lim_{k\to\infty} T_k.$$

Let $z_0, z_\infty$ be fixed elements of $X$. Define the path $(x_t)_{t\geq 0}$ by

$$x_t = \begin{cases} z_0 & \text{if} \quad t < T_1 \\ Z_i & \text{if} \quad t \in [T_i, T_{i+1}[ \\ z_\infty & \text{if} \quad t \geq T_\infty. \end{cases}$$

Then the probability structure on the process is determined by either joint distribution for $(T_i, Z_i, i = 1, 2, \ldots)$ or specifying

3 (i) distribution of $(Z_1, T_1)$

(ii) for each $k = 1, 2, \ldots$, conditional distribution of $(S_k, Z_k \mid T_{k-i}, i = 1, 2, \ldots)$, where $S_k = T_k - T_{k-i}$ is then $k^{\text{th}}$ inter-arrival time.

We will start studying the process $(x_t)$ having a single jump, i.e.,

$$x_t \begin{cases} z_0 & \text{if} \quad t < T(\omega) \\ Z_{(\omega)} & \text{if} \quad t \geq T(\omega). \end{cases}$$

If $T = \infty$, let $Z = z_\infty$, a fixed point of $X$.

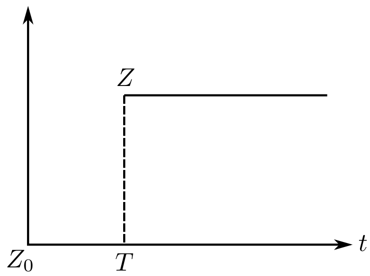


Figure 1.2:

Define the probability space $(\Omega, F, P)$ as the canonical space for $T$, $Z$,

i.e., $$((\mathbb{R}_+ \times X)U\{(\infty, z_\infty)\}, B(\mathbb{R}_+) \ * S, \{(\infty, z_\infty)\}, \mu)$$

where $\mu$ is a probability measure on

$$((\mathbb{R}_+ \times X)U\{(\infty, z_\infty)\}, B(\mathbb{R}_+) \ * S, \{(\infty, z_\infty)\}).$$

The random function $(x_t)$ generates the increasing family of $\sigma-$ fields **4**
$(F_t^0)$, i.e.,

$$F_t^0 = \sigma\{x_S, s \leq t\}.$$

We suppose

$$\mu(([0, \infty] \times \{z_0\})U\{0\} \times X) = 0.$$

This assumption guarantees that the process $x_t$ does jump at its jump time $T$, i.e.,

$$P(T > 0 \text{ and } Z \neq z_0) = 1.$$

Recall that an $\mathbb{R}_+$ - valued random variable $\tau$ is a stopping time of a filtration $F_t$, if $(\tau \leq t \in)F_t, \forall t$. Let

$$F_t = \text{ Completion of } F_t^0 \text{ with all } F_\infty^0 - \text{ null sets }.$$

**Proposition 1.1.** *$T$ is not an $F_t^0$ stopping time, but $T$ is an $F_t$ stopping time.*

*Proof.* Let $A = \{Z = z_0\}$ and $K$ be any set in $X$. Then

$$x_S^{-1}(K) = \begin{cases} ([s,\infty]\times X)U([o,s]\times\{z_0\}) & \text{if } z_0 \in K \text{ and } Z(E-A)\cap K = \phi \\ ((s,\infty]\times X)U([o,s]\times K) & \text{if } z_0 \in K \text{ and } Z(E-A)\cap K \neq \phi. \\ [o,s]\times K & \text{if } z_0 \notin K. \end{cases}$$

where $E = \mathbb{R}_+ \times X - A$.

Clearly $[0,t]\times X$ cannot be in the $\sigma$- algebra generated by sets of the above form. So $T$ is not an $F_t^0$ stopping time. Let $B = X - \{z_0\}$. By assumption, $P(A) = 0$; so $A \in F_t$.

$$x_t^{-1}(B) = [0,t]\times X - A \in F_t^0.$$

So

$$[0,t]\times X \in F_t.$$

But $\{T \leq t\} = [0,t]\times X$. Hence $T$ is an $F_t$ stopping time. □

It can be seen that

$$F_t = B[o,t] \ * S\, U(]t,\infty]\times X)U \text{ null sets of } F_\infty^\circ.$$

The stopped $\sigma$- field $F_T$ is given by

$$F_T = \{G \in F_\infty : G \cap (T \leq t) \in F_t, \forall t\}.$$

Clearly

$$F_T = F_\infty.$$

**Definition 1.2.** *A process $(M_t)$ is an $F_t$ -martingale if $E \mid M_t \mid < \infty$ and for $s \leq t$*

$$E[M_t \mid F_s] = M_s \ a.s.$$

*$(M_t)$ is a local $F_t$- martingale if there exists a sequence of stopping times $S_n \uparrow \infty$ a.s. such that $M_t^n := M_{t\Lambda S_n}$ is a uniformly integrable martingale for each n; here $t\Lambda S_n := Min(t, S_n)$.*

**Proposition 1.2.** *If $M_t$ is a local martingale and $S$ is a stopping time such that $S \geq T$ a.s., then $M_S = M_T$ a.s.*

*Proof.* Let $S_n$ be stopping times such that $S_n \uparrow \infty$ a.s. Then $M_{t \wedge S_n}$ is *u.i.* martingale. Let $M_t^n = M_{t \wedge S_n}, \forall n$. Then by optional sampling theorem

$$E[M_S^n F_{T \wedge S_n}] = M_T^n;$$

but

$$F_T = F_\infty.$$

So 6

$$E[M_S^n \mid F_T] = M_S^n.$$

Also

$$\lim_{n\to\infty} M_T^n = M_T$$

and

$$\lim_{n\to\infty} M_S^n = M_S.$$

So

$$M_S = M_T \ a.s.$$

□

**Proposition 1.3.** *Suppose $\tau$ is an $F_t$-stopping time. Then there exist $t_0 \in \mathbb{R}_+$ such that $\tau \wedge T = t_0 \wedge T$ a.s.*

*Proof.* If $\tau$ is a stopping time, then $(\tau \wedge T \leq t) \in F_t, \forall t$. But if $T \wedge \tau$ is not constant a.s. on $(\tau \leq T)$, then

$$(\tau \wedge T \leq t) \cap (]t, \infty] \times X) \underset{\neq}{\subset} ]t, \infty[ \times X \text{ for some } t \in \mathbb{R}_+.$$

But $[t, \infty] \times X$ is an atom of $F_t$. This contradicts the fact that $\tau$ is a stopping time. So

$$\tau \wedge T = t_0 \wedge T a.s.$$

The general definition of a stopped $\sigma$-field is that if $U$ is a stopping time. Then

$$F_U = \{A \in F \mid A \cap (U \leq t) \in F_t, \forall t\}.$$

But this is an implicit definition of the $\sigma$-field. □

**Exercise 1.1.** *Suppose* $\tau = t_0 \Lambda T$. *Show that*

(i) $F_\tau = F_{t_o}$

(ii) $F_\tau = \sigma\{x_{\tau\Lambda s}, S \geq 0\}$.

7

**Definition 1.3.** *For* $A \in S$, *define*

$$F^A(t) = \mu([t, \infty] \times A)$$

*and*

$$F(t) = F^X(t) = P[T > t].$$

*Note that* $F(0) = 1$ *and* $F(.)$ *is monotone decreasing and right continuous. Define*

$$c = \begin{cases} \inf\ \{t : F(t) = 0\} \\ +\infty \ \ \textit{if}\ \ \{t : F(t) = 0\} = \phi. \end{cases}$$

**Proposition 1.4.** *Suppose* $(M_t)_{t \geq 0}$ *is an* $F_t$ *local martingale. Then*

(a) *if* $c = \infty$ *or* $c < \infty$ *and* $F(c-) = 0$, *then* $M_t$ *is a martingale on* $[0, c[$.

(b) *if* $c < \infty, F(c-) > 0$, *then* $(M_t)$ *is a uniformly integrable martingale. Here* $F(c-) = \lim_{t \uparrow c} F(t)$.

*Proof.* (a) If $\tau_k \geq T a.s.$ for some $k$, then

$$M_{t\Lambda\tau_k} = M_{t\Lambda\tau_k\Lambda T} = M_{t\Lambda T} = M_t.$$

So $M_t$ is a *u.i.* martingale. Hence suppose $P[\tau_k < T] > 0$ for all $k(*)$; then by Proposition 1.3,

$$\tau_k \Lambda T = t_k \Lambda T \ \text{ for some fixed } t_k$$

8 and $t_K < c$ because of $(*)$. Also $t_k \uparrow c$ since $\tau_k \uparrow \infty$.

So

$$M_{t\Lambda\tau_k} = M_{t\Lambda\tau_k\Lambda T} = M_{t\Lambda T\Lambda t_k} = M_{t\Lambda t_k}.$$

Hence $M_{t\Lambda t_k}$ is a u.i. martingale. So $(M_t)_{t<c}$ is a martingale.

(b) $c < \infty, F(c_-) > 0$.

$$F(c_-) = P(T = c)$$

$P(T = c) > 0$; so it must be the case that $t_k = c$ for some $k$. Otherwise $P(\tau_k < c) \geq F(c-) > 0$; so "$\tau_k \uparrow \infty$ *a.s.*" fails. For this $k$,

$$M_{t\wedge t_k} = M_t$$

So $(M_t)$ is a *u.i.* martingale.

Our main objective is to show that all local martingales can be represented in the form of "stochastic integrals". So we introduce some "elementary martingales" associated with the process $(x_t)$. For $A \in S$ and $t \in \mathbb{R}_+$, define

$$\begin{aligned} p(t, A) &= \tilde{I}_{(t\geq T)} I_{(Z\in A)} \\ \tilde{p}(t, A) &= -\int_{]o,T\wedge t[} \frac{1}{F(s_-)} dF^A(s). \end{aligned}$$

□

**Proposition 1.5.** *Let $q(t, A) = p(t, A) - \tilde{p}(t, A)$. Then $(q(t, A))_{t\geq 0}$ is an $F_t$ - martingale, i.e., $p(\tilde{t}, A)$ the "compensator" of the point process $p(t, A)$.*

*Proof.* (Direct calculation). Take $t > s$, then **9**

$$E[p(t, A) - p(s, A) \mid F_s] = I_{(s<T)} \frac{F^A(s) - F^A(t)}{F(s)}.$$

$$E[\tilde{p}(t, A) - \tilde{p}(s, A) \mid F_s] = I_{s<T}$$

$$\left\{ \frac{F(t)}{F(s)} \int_{[s.t]} \frac{dF^A(u)}{F(u-)} - \frac{1}{F(s)} \int_{[s.t]} \int_{[s.r]} \frac{dF^A(u)}{F(u-)} dF(r) \right\}$$

and

$$\int_{[s,t]} \int_{[s,r]} \frac{dF^A(u)}{F(u-)} dF(r) = \int_{[s,t]} \frac{1}{F(u-)} \int_{[u,t]} dF(r) dF^A(u)$$

$$= \int_{[s.t]} \frac{1}{F(u-)}(F(t) - F(u-))dF^A(u)$$
$$= F(t) \int_{[s.t]} \frac{dF^A(u)}{F(u-)} + F^A(t) - F^A(s).$$

So

$$E[q(t,A) - q(s,A) \mid F_s] = 0$$

□

**Another expression for $\tilde{p}(t, A)$:** We have $F^A(.) << F(.)$ (i.e. $F^A(.)$ is absolutely continuous w.r.t. $F(.)$). So there exists a function $\lambda(s,A)$ such that

$$F^A(0) - F^A(t) = -\int_{]o.t]} \lambda(s,A)dF(s).$$

In fact

$$\lambda(s,A) = P(Z \in A \mid T = s).$$

Suppose $X$ is such that a regular version of this conditional probability exists (which is the case, since $X$ is Borel space). Then $\dfrac{-dF^A(s)}{F(s-)} = \lambda(s,A)d\Lambda(s)$ where $d\Lambda(s)\dfrac{-dF(s)}{F(s-)}$.

Then

$$\tilde{p}(t,A) = \int_{]o,T \wedge T]} \lambda(s,A)d\Lambda(s).$$

## 10 Stochastic Integrals

Let $I$ denote the set of measurable functions $g : \Omega \to \mathbb{R}$ such that $g(\infty, z_\infty) = 0$.

**(a) Integrals w.r.t. $p(t, A)$:** Suppose $N_t$ is a counting process. Since its sample functions are monotone increasing and there is a one-to-one correspondence between monotone increasing functions and measures, and since in this case, mass is concentrated at the jump points and they are only countable; the function $N_t$ defines a random measure on $(\mathbb{R}, B(\mathbb{R}))$ say, $\pi = \sum\limits_i \delta_{T_i}$ where $\delta_X$ is the Dirac measure at $x$. Similarly, the one jump process can be identified with the random measure $\delta_{(T, x_T)}$ on $R_+ \times X$. So we can define Stieltjes integrals of the form $\int g(t,x)p(dt,dx)$ for suitable integrands $g \in I$ as

$$\int\limits_{\Omega} g(t,x)\, p(dt,dx) = g(T, x_T).$$

We say $g \in L_1(p)$ if

$$E \int\limits_{\Omega} |g(t,x) \mid p(dt,dx) < \infty$$

and denote

$$\| g \|_{L_1(p)} = E \int\limits_{\Omega} \mid g(t,x) \mid p(dt,dx)$$

Clearly $g \in L_1(p)$ if and only if

$$\int\limits_{\mathbb{R}_+ \times X} \mid g(t,x) d\mu < \infty$$

**(b) Integrals w.r.t. $\tilde{p}(t, A)$:** **11**

Recall $\tilde{p}(t,A) = \int\limits_{]o,T\ \wedge\ t]} \lambda(s,A) d\Lambda(s)$.

So we define

$$\int\limits_{\Omega} g(t,x)\tilde{p}(dt,dx) = \int\limits_{[o,T]} \int\limits_{X} g(t,x)\lambda(t,dx)d\lambda(t)$$

and say $\quad g \in L_1(\tilde{p})$ if $\int\limits_{\Omega} |g(t,x) \mid \tilde{p}(dt,dx) < \infty$

and $\quad \| g \|_{L_1(\tilde{p})} = \int\limits_{\Omega} \mid g(t,x) \mid \tilde{p}(dt,dx).$

**Proposition 1.6.**

$$\| g \|_{L_1(p)} = \| g \|_{L_1(\tilde{p})}$$

*and so*

$$L_1(p) = L_1(\tilde{p}).$$

*Proof.*

$$\begin{aligned}
\| g \|_{L_1(\tilde{p})} &= -\int_{\mathbb{R}_+}\int_{[o,T]} \frac{1}{F(s_-)} \mid g(s,x) \mid d\mu(s,x) dF(T) \\
&= \int_{\Omega} \frac{1}{F(s_-)} \mid g(s,x) \mid (-\int_{[s,\infty]} dF(t)) d\mu(s,x) \\
&= \int_{\Omega} \mid g(s,x) \mid d\mu(s,x) \\
&= \| g \|_{L_1(p)} \ .
\end{aligned}$$

Define

$L_1^{\text{loc}}(p) = \{g \in I \mid g(s,x) I_{s \le t} \in L_1(p), \forall t < c\}$
$L_1^{\text{loc}}(\tilde{p}) = \{g \in I \mid g(s,x) I_{s \le t} \in L_1(p), \forall t < c\}$. Clearly
$L_1^{\text{loc}}(p) = L_1^{\text{loc}}(\tilde{p})$.

12 Following is the main result of this section, which gives an integral representation for $F_t$ local martingales. □

**Proposition 1.7.** *All $F_t$-local martingales are of the form*

$$\begin{aligned}
M_t &= \int_{\Omega} (g(s,x) I_{(s \le t)} dq(s,x) \\
&= \int_{\Omega} (g(s,x) I_{s \le t} dp(s,x) - \int_{\Omega} (g(s,x) I_{(s \le t)} d\tilde{p}(s,x).
\end{aligned}$$

*for some $g \in L_1^{\text{loc}}(p)$.*

We need the following result.

**Lemma 1.1.** *Suppose $(M_t)_{t>0}$ is u.i.$F_t$ martingale with $M_0 = 0$. Then there exists a function $h : \Omega \to \mathbb{R}$ such that*

$$E \mid h(T,Z) \mid < \infty \tag{1}$$

*and*

$$M_t = I_{(t\geq T)}h(T,Z) - I_{(t<T)}\frac{l}{F(t)}\int\limits_{]0,t]\times X} h(s,z)\mu(ds,dz) \tag{2}$$

*Proof.* If $(M_t)$ is a *u.i.* martingale, then $M_t = E[\xi \mid F_t]$ for some $F_\infty$-measurable r.v. $\xi$ and from the definition of $F_\infty$, we have

$$\xi = h(T,Z)\ a.s.$$

for some measurable $h : \Omega \to \mathbb{R}$. Expression (1) is satisfied since $M_t$ is **13**
u.i., and $M_o = 0$ implies

$$\int\limits_{\Omega} h.d\mu = 0. \tag{3}$$

Now

$$\begin{aligned} M_t &= E[h(T,Z)|F_t] \\ &= I_{(t\geq T)}h(T,Z) + I_{(t<T)}\frac{1}{F(t)}\int\limits_{]t,\infty]\times X} h(s,x)\mu(ds,dx). \end{aligned} \tag{4}$$

From (3) and (4), we have (2). ◻

For $g \in L_I^{\text{loc}}(p)$, the stochastic integral

$$M_t^g = \int\limits_{]0,t]\times X} g(s,x)q(ds,dx)$$

is defined by

$$M_t^g = \int\limits_{\mathbb{R}^+\times X} I_{(s\leq t)}g(s,x)p(ds,dx) - \int\limits_{\mathbb{R}^+\times X} I_{(s\leq t)g(s,x)}\tilde{p}(ds,dx).$$

Then the question is whether $M_t$ given by (2) is equal to $M_t^g$ for some $g$. As a motivation to the answer consider the following example.

**Example 1.1.** Let $(X, S) = (\mathbb{R}, B(\mathbb{R}))$ and

$$\mu(ds, dx) = \psi(s, x)\ ds\ dx.$$

Then

$$M_t^g = I_{(t\geq T)}\left\{g(T,Z) - \int_0^T\int_{\mathbb{R}} \frac{1}{F(s)} g(s,x)\psi(s,x)dxds\right\} - I_{(t\geq T)}\left\{\int_0^t\int_{\mathbb{R}} \frac{1}{F(s)} g(s,x)\psi(s,x)dxds\right\} \quad (5)$$

14 If $M_y^g$ given by (5) is equal to $M_t$ given by (2), then the coefficients of $I_{(t\geq T)}$ and $I_{(t<T)}$ must agree. Comparing the coefficients of $I_{(t>T)}$, we require

$$h(t,z) = g(t,z) - \int_0^t\int_{\mathbb{R}} \frac{1}{F(s)} g(s,x)\psi(s,x)dxds.$$

Let

$$\eta(t) = h(t,z) - g(t,z).$$

Define

$$\gamma(t) = \int_{\mathbb{R}} \psi h dx$$

and

$$f(t) = \int_{\mathbb{R}} \psi dx.$$

Then

$$\begin{aligned}\eta(t) &= \int_0^T \frac{1}{F(s)}\left(\int_{\mathbb{R}} h(s,z) + \eta(s)\right)\psi(s,x)dx)ds \\ &= \int_0^t \frac{1}{F(s)}\gamma(s)ds + \int_0^t \frac{1}{F(s)}\eta(s)f(s)ds;\end{aligned}$$

that is

$$\frac{d}{dt}\eta(t) = \frac{f(t)}{F(t)}n(t) + \frac{1}{F(t)}\gamma(t)$$
$$\eta(o) = 0$$

which has a unique solution

$$\eta(t) = \int_0^t \phi(t,s)\frac{1}{F(s)}\gamma(s)ds,$$

where

$$\begin{aligned}\phi(t,s) &= \exp\int_s^t \frac{f(u)}{F(u)}du\\ &= \frac{F(s)}{F(t)},\ \text{since } f(t) = -\frac{dF(t)}{dt}.\end{aligned}$$

So

**15**

$$\eta(t) = \frac{1}{F(t)}\int_0^t \gamma(s)ds.$$

Hence

$$g(t,z) = h(t,z) + \frac{1}{F(t)}\int_0^t\int_{\mathbb{R}} h(s,x)\psi(s,x)dxds \tag{6}$$

Now it can be checked that with this choice of $g$ the coefficients of $I_{(t<T)}$ in (2) coincides with that of (5). So $M_t = M_t^g$.

Now we can prove the general case given in Proposition 1.7.

**Proof of Proposition 1.7.**

**Case 1.** $c < \infty, F(c_-) > 0$. *Take a local martingale $M_t$ with $M_o = 0$. Then $(M_t)$ is u.i. So $M_t = E[h(T,Z)|F_t]$ for some measurable $h$ such that $E|h| < \infty, Eh = 0$. Then we claim that $M_t = M_t^g$ where*

$$g(t,z) = h(t,z) + I_{(t<c)}\frac{1}{F(s)}\int_{]\circ,t]\times x} h(s,z)\,\mu\,(ds,dz).$$

But this can be seen algebraically following similar calculations as of the example 1.1. Now to show that $g \in L_1^{\text{loc}}(p)$.

$$\begin{aligned}
\int |g| d\mu &\leq \int |h| d\mu - \int_{]o,c[} \frac{1}{F(t)} \int_{]o,t]\times X} |h| d\mu dF(t) \\
&\leq \int |h| d\mu - \frac{1}{F(c_-)} \int_{]o,c[} \int_{]o,t]\times X} |h| d\mu dF(t) \\
&\leq \int |h| d\mu + \frac{1}{F(c_-)} \int_{]o,c[\times X} (F(t) - F(c_-)) |h| d\mu \\
&\leq \left(1 + \frac{1}{F(c_-)}\right) \int |h| d\mu < \infty.
\end{aligned}$$

16 **Case 2.** *$c = \infty$, or $c < \infty$ and $F(c_-) = 0$. Then from proposition 1.4, $M_t$ is a martingale on $[0, c]$, and so it is u.i. on $[o, t]$ for $t < c$. Therefore*

$$M_s = E[h(T, Z)|F_\infty]$$

*for some function h satisfying*

$$\int_{]o,t]xX} |g(s, x)| d\mu(s, x) < \infty \textit{ for all } t < c.$$

*Define $g(s, x)$ as in (6). Then calculations as in case 1. Show that $M_s = M_s^g$ for $s \leq t < c$. Now*

$$\begin{aligned}
\int_{[o,t]\times X} |g| d\mu &\leq \int_{]o,t\times X} |h| d\mu - \int_{]o,t]} \frac{1}{F(s)} \leq \int_{]o,s]\times X} |h| d\mu \, dF(s) \\
&\leq \int_{]o,t]\times X} |h| d\mu \Big(1 - \int_{]o,t]} \frac{1}{F(s)} dF(s)\Big) \\
&< \infty \textit{ for } t < c.
\end{aligned}$$

*Hence $g \in L_1^{log}(p)$.*

Conversely, suppose $g \in L_1^{log}(p)$. Then it can be checked that $M_t^g$ is a local martingale.

**Remark 1.1.** If $g \in L_1^{log}(p)$ then $M_t^g$ is a martingale. But the result does not say $M_t$ is a martingale if and only if $M = M^g$ for $g \in L_1(p)$, it only characterizes *local* martingales.

**Remark 1.2.** All preceding results hold if $z_o$ is a random variable; then $\mu$ should be taken as conditional distribution of $(T, Z)$ given $z_o$

**The multi-jump case:** The process $x_t$ has jump times $T_1, T_2, \ldots$ with **17**
corresponding states $Z_1, Z_2, \ldots$ Let $(Y, y)$ denote the measurable space

$$(Y, y) = ((\mathbb{R}_+ \times X)U\{(\infty, z_\infty)\}, \sigma\{B(\mathbb{R}_+) * S, \{(\infty, z_\infty)\}\}).$$

Define

$$\Omega = \prod_{i=1}^{\infty} Y_i, \Omega_K = \prod_{i=1}^{K} Y_i$$
$$F^o = \sigma\left\{\prod_{i=1}^{\infty} y_i\right\}$$

where $(Y_i, y_i)$ denote a copy of $(Y, y)$. Let

$$S_k(\omega) = T_k(\omega) - T_{k-1}(\omega)$$

and $$w_k(\omega) = (S(\omega), Z_1(\omega), \ldots S_k(\omega), Z_k(\omega)).$$

Then

$$T_k(\omega) = \sum_{i=1}^{k} S_i(\omega)$$
$$T_\infty(\omega) = \lim_{k \to \infty} T_k(\omega).$$

As before, $(x_t(w))_{t \geq o}$ is defined by

$$x_t(\omega) = \begin{cases} z_o & \text{if } t < T(\omega) \\ Z_k & \text{if } t \in [T_k(\omega), T_{k+1}(\omega)] \\ z_\infty & \text{if } t \geq T_\infty(\omega) \end{cases}$$

A probability measure $\mu$ on $(\Omega, F^o)$ is specified by a family $\mu^i$ : $\Omega_{i-1} \times y \to [0,1]$ (with $\Omega_o = \phi$) satisfying

(i) $\mu^i(.;\Gamma)$ is measurable for each fixed $\Gamma$

(ii) $\mu^i(w_{i-1}(\omega);.)$ is a probability measure on $(Y, y)$ for each fixed $\omega \in \Omega$,

18 (iii) $\mu^i(w_{i-1}(\omega);.(\{0\} \times X) \cup (\mathbb{R}_+ \times \{z_{i-1}(\omega)\})) = 0$ for all $\omega$,

(iv) $\mu^i(w_{i-1}(\omega);.\{(\infty, z_\infty)\}) = 1$ if $S_{i-1}(\omega) = \infty$.

Then for $\Gamma \in y$ and $\eta \in \Omega_{i-1}, \mu$ is defined by

$$\mu[(T_1, Z_1) \in \Gamma] = \mu^1(\Gamma)$$
$$\mu[(S_i, Z_i) \in \Gamma | w_{i-1} = \eta] = \mu^i(\eta : \Gamma) i = 2, 3 \cdots$$

Notice that as in the single jump case, here (iii) ensure that two "jump times" $T_{i-1}, T$ do not occur at one and that the process $x_t$ does effectively jump at its jump times (iv) ensures that $\mu[Z_k = z_\infty | T_k = \infty] = 1$.

As before, $F_t^o = \sigma\{x_s, s \leq t\}$ and
$F_t$ = completion of $F_t^o$ with all $\mu$-null sets of $F^o$.

**Proposition 1.8.** *(i) $F_\infty = F$, the completion of $F^o$*

*(ii)* $F_{T_n} = \prod_{i-1}^{n} y_i \times \prod_{i=n+1}^{\infty} y_i$.

The idea here is to reduce everything to one jump case. That is, the process "restarts" at each $T_k$. We need the following result.

**Proposition 1.9.**

$$F_{(T_{k-l}+t)} \Lambda T_k = F_{T_{k-l}} V \sigma \left\{ x_{(T_{k-l}+s)} \Lambda T_k, o \leq s \leq t) \right\}.$$

Proof of this is an application of the "Galmarino test" (Dellacherie and Mayer [12], theorem IV, pp. 100).

19 Recall that in one jump case $F_{t_o \Lambda T} = F_{t_o}$. Now we conjecture that,

if $U = (T_{k-1} + t_o)\Lambda T_k$, then

$$F_U = \left(\prod_{i-1}^{k-1} y_i\right) * y_{t_o}^k * \left(\prod_{i=k+1}^{\infty} Y_i\right),$$

where $$y_{t_o}^k = S * B[o, t_o] \cup (x \times [t_o, \infty]).$$

As an example, see the following exercise.

**Exercise 1.2.** *Consider a point process with* $k = 2$*, and take the probability space as* $\mathbb{R}_+^2$*. Then*

$$x_t = \sum_{i=1}^{2} I_{(t \geq T_i)}.$$

Then

(a) Show that

$$\begin{aligned} F_t = &\text{ Borel sets in } \{S_1 + S_2 \leq t\} \\ &+ (A \times \mathbb{R}_+) \bigcap B + [t, \infty] \times \mathbb{R}_+, A \in B(\mathbb{R}) \end{aligned}$$

where

$$B = \{S_1 + S_2 \geq t\} \bigcap \{S_2 \leq t\}.$$

(b) With $U = (T_1 + t_o)\Lambda T_2$, show that

$$F_U = B(\mathbb{R}_+ \times [o, t]) + \{(A \times \mathbb{R}_+) \bigcap (\mathbb{R}_+ \times ]t_o, \infty]) : A \in B(\mathbb{R})\}.$$

**Elementary Martingales, Compensator's** Define

$$p(t, A) = \sum_i I_{(t \geq T_i)} I_{(z_i \in A)}$$

which counts the jump of $(x_t)$ ending in the set $A$. Define

$$\phi_1^A(s) = -\int_{[o,s]} \frac{1}{F^1(u)} dF^{1A}(u)$$

where 20

$$F^1 = \mu^1([t,\infty]\times X)$$

and

$$\phi_k^A(w_{k-1};s) = -\int\limits_{[o,s]} \frac{1}{F^k(u_-)} dF^{kA}(u)$$

where

$$F^{kA}(u) = \mu^k(w_{k-1};[u,\infty]\times A).$$

Now define

$$\tilde{p}(t,A) = \phi_1^A(T_1) + \phi_2^A(w_1;S_2) + \cdots \phi_j^A(w_{j-1};t-T_{j-1}(\omega))$$
$$\text{for } t \in ]T_{j-1},T_j].$$

**Exercise 1.3.** *Consider a renewal process*

$$x_t = \sum_i I_{(t\geq T_i)}$$

*and $S_i$'s are independent, $P(S_i > t) = F(t)$ is continuous. Then show that the compensator for $x_t$ is*

$$\tilde{p}(t) = -\ell n(F(S_1)F(S_2)\cdots F(S_{k-1})F(t-T_{k-1})) \text{ for } t \in [T_{k-1},T_k],$$

*and* $\quad x_t - \tilde{p}(t)$ *is a martingale.*

**Example 1.** If $F(t) = e^{-\alpha t}$, then $\tilde{p}(t) = \alpha t$.

21 **Proposition 1.10.** *For fixed k, and $A \in S$,*

$$q(t\Lambda T_k,\ A) = p(t\Lambda T_k,\ A) - \tilde{p}(t\Lambda T_k,\ A), t \geq 0$$

*is an $F_t$ - martingale.*

*Proof.* Calculation as in proposition 1.5. ◻

The class of integrands $I$ consists of measurable function $g(t,x,\omega)$ such that

$$g(t,x,\omega) = \begin{cases} g^1(t,x), & t \leq T_1(\omega) \\ g^k(w_{k-1},t-T_{k-l},x), & t \in ]T_{k-l}(\omega),T_k(\omega)], \\ 0, & t \geq T_\infty(\omega) \end{cases}$$

for some function $g^k$ such that $g^1(\infty, x) = g^k(w_k; \infty, x) \equiv 0$. Such $g's$ are $F_1-$predictable processes. Now we define $L_1(p), L_1(\tilde{p})$, etc. exactly as in one jump case:

$$\int g\,dp = \sum_i g(T_i, Z_i)$$

$$L_i(p) = \left\{ g \in I : E \sum_i |g(T_i, Z_i)| < \infty \right\}$$

$$\int g\,d\tilde{p} = -\sum_k \int\limits_{]o,T_k - T_{k-l}] \times X} g(\omega_{k-l}, s, x)\lambda(\omega_{k-l}, s, dx)\frac{dF^k(s)}{F^k(s_-)}$$

where

$$\lambda(\omega_{k-l}, s, A) = \frac{dF^{kA}}{dF^k}(s)$$

$g \in \Gamma_1^{\mathrm{loc}}(p)$ if there exists a sequence of stopping times $\sigma_k \uparrow T_\infty$ a.s. and $gI_{(t \le \sigma_n} \in L_1(p), \forall n$. For $g \in L_1^{1oc}(p)$ we define

$$M_t^g = \int\limits_{[o,t]\times X} g(s,x)\,q\,(ds,dx)$$

$$= \int\limits_{[o,t]\times X} g(s,x)\,p\,(ds,dx) - \int\limits_{[o,t]\times X} g(s,x)\,\tilde{p}\,(ds,dx).$$

**Proposition 1.11.** *If $g \in L_1^{1oc}$ then there exists a sequence of stopping* 22
*times $T_n < T_\infty$ such that $\tau_n \uparrow T_\infty$ and $M^g_{t\wedge T_n}$ is a u.i. martingale for each n.*

*Proof.* Take $\tau_n = n \wedge T_n \wedge \sigma_n$. Then the result follows by direct calculations using the optional sampling theorem. □

Now let $(M_t)_{t\ge 0}$ be a u.i. $F_t-$martingale. Then

$$M_t = M_{t\wedge T_1} + \sum_{k=2}^{\infty}(M_{t\wedge T_k} - M_{T_{k-1}}), I_{t \ge T_{k-l}} \tag{7}$$

because this is an identity if $t < T_\infty$ and the right-hand side is equal to $\lim M_{T_K}$ is $t \ge T_\infty$. Here we have $M_{T_\infty -} = M_{T_\infty}$. Now we state the main result.

**Theorem 1.1.** *Let $(M_t)$ be a local martingale of $F_t$. Then there exists $g \in L_1^{\text{loc}}(p)$ such that*

$$M_t - M_o = \int\limits_{[o,t]\times x} g(s, x, \omega)\, q(ds, dx).$$

*Proof.* Suppose first that $M_t$ in a u.i. martingale. define

$$x_t^1 = M_{t\Lambda T_1}$$
$$x_t^k - M_{(t+T_{k-l})\Lambda T_k} - M_{T_{k-1}}, k = 2, 3, \ldots$$

Then from (7)

$$M_t = \sum_{k=l}^{\infty} x^k_{(t-T_{k-l})Vo}.$$

We can now use proposition 1.7 to represent each $x^k$. Fix $k$ and define for $t \geq 0$.

$$H_t = F_{(t+T_{k-1})\Lambda T_k}.$$

23 Then $x_t^k$ is an $H_t$ martingale. Then there exists a measurable function $h^k$ such that

$$x_t^k = E(h^k(\omega_{k-1}; S_k, Z_k)|H_t).$$

Then using proposition 1.7, there exists $g^k(\omega_{k-1}; s, z)$ such that

$$x_t^k = \int\limits_{]o,t]\times X} g^k(\omega_{k-1}; s, z) q^k(ds, dz)$$

where $q^k(t, A) = q((t + T_{k-1})\Lambda T_k, A)$ and $g^k \in L_1^{\text{loc}}(p^k)$ for all $\omega^{k-1}$ a.s. Piecing these results together for $k = 1, 2, 3, \ldots$ gives the desired representation with $g = (g^k)$. It remains to prove that $g \in L_1^{\text{loc}}(p)$ as defined, for which we refer to Davis [6].

If $(M_t)$ is only a local martingale with associated stopping time sequence $\tau_n \uparrow \infty$ such that $M_t\Lambda\tau_n$ is a u.i. martingale, apply the above arguments to $M_t\Lambda\tau_n$ to complete the proof. $\square$

**Corollary 1.1.** *If $T_\infty = \infty$ a.s. then the result says $(M_t)$ is a local martingale of $F_t$ if and only if $M_t = M_t^g$ for some $g \in L_1^{\text{loc}}(p)$.*

**Remark 1.3.** It would be useful to determine the exact class of integrands $g$ required to represent u.i. martingales (as opposed to local martingales) when the jump times $T_i$ are totally inaccessible, Boel, Varaiya and Wong [4] show that $\{M^g, g \in L_1(p)\}$ coincides with the set of $u.i.$ 24
martingales of integrable variation. It seems likely that this coincides with the set u.i. martingales if $Ep(t, E) < \infty$ for all $t$ (a somewhat stronger condition than $T_i \to \infty$ a.s.) but no proof of this is available as yet.

## 2 Some Discontinuous Markov Processes

### Extended Generator of a Markov Process

Let the process $x_t \in (E, E)$, some measurable space. Then $(x_t, F_t)$ is a Markov process if for $s \leq t$

$$E[f(x_t)|F_s] = E[f(x_t)|x_s] a.s.$$

A *transition function* $p(s, x, t, \Gamma)$ is a function such that

$$\begin{aligned} p(s, x_s, t, \Gamma) &= P(x_t \in \Gamma | x_s) \\ &= E[I_\Gamma(x_t)|x_s]\ a.s.\ for\ t \geq s. \end{aligned}$$

$p$ satisfies the Chapman-Kolmogorov equation

$$p(s, x, t, \Gamma) = \int_E p(s, x, u, dy) p(u, y, t, \Gamma) \text{ for } s \leq u \leq t.$$

Not every Markov process has a transition function, but usually one wants to start with transition function and construct the corresponding process. This is possible if $(E, E)$ is a Borel space (required to apply Kolmogorov extension theorem; refer Wentzel [27]). One constructs a Markov family,

$$\{P_{x,s}, (x, s) \in E \times \mathbb{R}_+\} p_{x,s}$$

being the measure for the process starting at $x_s = x$. All measures $P_{s,x}$ have the same transition function $p$. Denote by $E_{x,s}$ integration w.r.t

$P_{x,s}$. Let $B(E)$ be the set of bounded measurable functions

$$f : E \to \mathbb{R} \text{ with } \|f\| = \sup_{x \in E} |f(x)|$$

25 Define

$$T_{s,t} f(x) = E_{x,s} |f(x_t)|], s \leq t.$$

$T_{s,t}$ is an operator on $B(E)$ such that

(i) it is contraction, $\|T_{s,t} f\| \leq \|f\|, T_{s,t} 1 = 1$.

(ii) Semi group property: $r \leq s \leq t$,

$$T_{r,t} = T_{r,s} T_{s,t}$$

for

$$\begin{aligned} T_{r,t}(T_{s,t} f)(x) &= E_{x,r}[E_{x_s,s}(f(x_t))] \\ &= E_{x,r}[E(f(x_t)|F_s)] \\ &= E_{x,r} f(x_t)) \\ &= T_{r,t} f(x). \end{aligned}$$

$T_{s,t}$ is time invariant if $T_{s+r,t+r} = T_{s,t}$ for all $r \geq -s$. Then $T_{s,t} = T_{0,t-s} \equiv T_{t-s}$. So $T$ is a one parameter family; this happens when the transition function is time invariant i.e., $p(s, x, t, \Gamma) = p(s + r, x, t + r, \Gamma)$. Then get a one parameter family of measures $(P_x, x \in E)$ and the connection is

$$T_t f(x) = E_x f(x_t); T_0 f = f.$$

Let

$$B_0(E) = \{f \in B(E) : \|T_t f - f\| \to 0, t \downarrow 0\}.$$

An operator $\overset{\circ}{A}$ with domain $D(\overset{\circ}{A}) \subset B_\circ(E)$ is the *strong infinitesimal generator* of $T_t$ if

$$\lim_{t \downarrow 0} \|(T_t f - f) - \overset{\circ}{A} f\| = 0.$$

26 So

$$\overset{\circ}{A} f = \frac{d}{dt} T_t f(x)|_{t=0}$$

Take $f \in \mathcal{D}(\overset{\circ}{A})$. Then

$$\begin{aligned}
\lim_{t \downarrow 0} \frac{1}{t}(T_t T_s f - T_s f) &= \lim_{t \downarrow 0} \frac{1}{t}(T_{t+s} f - T_s f) \\
&= \lim_{t \downarrow 0} T_s \frac{1}{t}(T_t f - f) \\
&= T_s \overset{\circ}{A} f.
\end{aligned}$$

So $f \in \mathcal{D}(\overset{\circ}{A})$ implies $T_s f \in \mathcal{D}(\overset{\circ}{A})$ and $\overset{\circ}{A} T_s f = T_s \overset{\circ}{A} f$. So we get backward Kolmogorov equation

$$\frac{d}{ds} T_s f = \overset{\circ}{A}(T_s f). \tag{1}$$

The main results of the "analytic theory" of Markov semigroups are the following;

(i) Hille-Yosida theorem: Necessary and sufficient conditions for an operator $\overset{\circ}{A}$ to be the generator of some semigroup.

(ii) If $\overset{\circ}{A}$ satisfies these conditions, then $\mathcal{D}(\overset{\circ}{A})$ is dense in $B_0(E)$ and $(\overset{\circ}{A}, \mathcal{D}(\overset{\circ}{A}))$ determines $T_t$ (via the so called resolvent operator).

**NB:** The domain $\mathcal{D}\overset{\circ}{A}$ provides essential information.

Integrating (1), we get *Dynkin's formula*

$$T_t f(x) - f(x) = \int_0^t T_s \overset{\circ}{A} f(x) ds$$

i.e.,
$$E_x f(x_t) - f(x_0) = E_x \int_0^t \overset{\circ}{A} f(x_s) ds, \; f \in \mathcal{D}(\overset{\circ}{A}).$$

**Proposition 2.1.** *If $f \in \mathcal{D}(\mathring{A})$ then the process*

$$C_t^f = f(x_t) - f(x_0) - \int_0^t \mathring{A} f(x_s) ds$$

*is a martingale.*

27 *Proof.* For $t \geq s$

$$E|C_t^f - C_s^f|F_s] = E\left[f(x_t) - f(x_s) - \int_s^t \mathring{A} f(x_u) du|F_s\right]$$

$$= E_{x_s} f(x_t) - f(x_s) - E_{x_s} \int_s^t \mathring{A} f(x_s) ds. \qquad 0. \qquad \square$$

$\square$

**Definition 2.1.** *Let $M(E)$ be the set of measurable functions $f : E \to \mathbb{R}$. Then $A, \mathcal{D}(A)$ with $\mathcal{D}(A) \subset M(E)$ is the* extended generator *of $(x_t)$ if $C_t^f$ is a local martingale, where*

$$C_t^f = f(x_t) - f(x_0) - \int_0^t A f(x_s) ds.$$

*This is an extension of $(\mathring{A}, \mathcal{D}\mathring{A}))$ in that $\mathcal{D}(\mathring{A}) \subset \mathcal{D}(A)$ and $\mathring{A} f = A f$ for $f \in \mathcal{D}(\mathring{A})$. We have uniqueness of $A$ in the following sense. Write*

$$f(x_t) = f(x_0) + \int_0^t A f(x_s) ds + C_t^f.$$

*This shows that $(f(x_t))$ is a "special semi-martingale" (=local martingale + predictable bounded variation process). The decomposition is unique. So, if $B$ is another generator then*

$$\int_0^t (A f(x_s) - B f(x_s)) ds = 0\, P_x\, a.s. \forall t.$$

*Thus $Af(x) = Bf(x)$ except on a set of potential zero. where a set of $\Gamma$ has potential zero, where a set $\Gamma$ has potential zero if*

$$E_x \int_0^\infty I_\Gamma(x_s)ds = 0 \ \forall x.$$

**Example 2.1.** Suppose $x_t \in \mathbb{R}^d$ satisfies

$$dx_t = b(x_t)dt + \sigma(x_t)\, dw_t$$

with standard Ito conditions. If $f \in C^2$, then **28**

$$df(x_t) = \left(\sum_i b_i(x_t)\frac{\partial f}{\partial x_i}(x_t) + \frac{1}{2}\sum_{i,j}(\sigma\sigma')_{ij}\frac{\partial^2 f}{\partial x_i \partial x_i \partial x_j}\right)dt + \int_0^t \nabla f'\sigma dw.$$

So $C^2 \subset \mathcal{D}(A)$ and

$$Af(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{ij}(\sigma\sigma')_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j} C_t^f = \int_0^t \nabla f'\sigma dw.$$

**NB:** This is not a characterisation of $\mathcal{D}(A)$.

**Remark 2.1.** If we had required $C_t^f$ to be a martingale rather than a local martingale in definition 2.1, then not every $f \in C^2$ would be in $\mathcal{D}(A)$ because of the properties of I to integrals.

**Exercise 2.1.** *For $i = 1, 2, \ldots$,let $N_t^i$ be a Poisson process with rate $\lambda_i$ where $\sum_i \lambda_i < \infty$. Define*

$$X_t = \sum_{i=1}^\infty \ell_i N_t^i$$

*where $\ell_i \geq 0$ and*

$$\sum_{i=1}^\infty \ell_i \lambda_i = r < \infty.$$

*Find the extended generator of $x_t$*

*(This is also an example where jump times are not isolated).*

**Piecewise-linear Markov Process** 29

Gnedenko-Kovalenko introduced the concept of piecewise linear Markov process. Later, Vermes [24] simplified the definition as follows. A *piecewise linear process* is a two component Markov process $(x_t) = (\nu_t, \xi_t)$ where $\nu_t$ is integer-valued and $\xi_t$ takes values in an interval $[a_n, b_n]$ of the real line if $\nu_t = n$ ($b_n$ may be $+\infty$). Let $E$ be the state space, i.e., $E = \{(n, \xi) \in Z \times R : \xi \in Z \times R : \xi \in [a_n, b_n]\}$ Then the probabilistic description is that if the motion starts at $(n, z) \in E$ and $x_t$ os given by $\nu_t = n, \xi_t = z + t$ for $t < T_1$, the first jump time. " Spontaneous jumps" happen at rate $\lambda(x_t)$, i.e., probability "jump occurs" in $(t, t + dt)$, is $\lambda(x_t)dt$, and process must jump if $\xi_t - = b_n$. Let the transition measure be given by $Q(A; x)$ for $A \in B(E)$. Then $x_{T_1}$ is selected from the probability distribution $Q(A; x_{T_1})$. After a jump, motion restarts as before. Thus the law of the process is determined by specifying the intervals $[a_n, b_n]$, the jump intensity $\lambda(x)$ and the transition measure $Q(A; x)$.

**Example 2.2.** *Non-stationary countable state Markov Process* $(\xi_\tau)$

$(\xi_\tau)$ takes integer values with the-dependent transition rates $a_{ij}(t)$ such that

$$P\Gamma\xi_{t+h} = i|\xi_t = j] = a_{ij}(t)\delta + \circ(\delta), i \neq j.$$

Then $x_t = (\xi_t, t)$ is a $PL$ process with no barriers, i.e., $a_n = 0, b_n = +\infty$.

30 **Example 2.3.** *Countable state process with non-exponential sojourn times*

Here, jump times of the process $(x_t)$ form a renewal process with inter arrival density $b(.)$ and transition matrix $q_{ij} = P[x_{T_k}, = i, x_{T_k} = j]$. This is a $PL$ process with $\nu_t = x_t$, and $\xi_t$ the time since last arrival. The jump rate is

$$\lambda(\nu, \xi) = \frac{b(\xi)}{\int\limits_{\xi}^{\infty} b(t)dt}$$

Here again $a_n = 0, b_n = +\infty$.

**Example 2.4.** *Virtual waiting times.* (The M/ G/ 1 queue)

Customers arrive at a single-server queue according to Poisson process with rate $\mu$, and have *i.i.d.* service time requirements with distributions $F$. The virtual waiting time $\xi_t$ is the time a customer arriving at time $t$ would have to wait for service. Piecewise linear process structure is :

$$\nu_t = \begin{cases} 1 & \text{if queue is not empty} \\ 0 & \text{if queue is empty} \end{cases}$$

and $a_1 = 0, b_1 = \infty, a_\circ = b_\circ = 0$. Here $\xi_t$ moves to left with uniform speed and transition to $(0,0)$ is certain if $x_{t-} = (1,0)$.

A more general definition of *PL* process of Gnedenko and Kovalenko allows $\xi_t$ to move in an open subset $0_{\nu_t}$ of $\mathbb{R}^d(\nu_t)$ with uniform speed in a fixed direction $V(\nu_t)$. Again transition must take place if $\xi_{t-}, \in \partial 0_{\nu_t}$, the boundary of $0_{\nu_t}$.

**Example 2.5.** *VWT with renewal process arrivals.* (The GI/G/I queue) **31**

Suppose the inter arrivals times in Example 2.4 are not exponential, but form a renewal process with inter arrival density $b(.)$. Now the appropriate structure is $\nu$ is 0 or 1 as before, $d(1) = 2$, $d(0) = 1$. (When $\nu = 1$ we have to remember both the value of VMT and the time since the last arrival.)

We cannot accommodate this in previous framework, because there $[a_n, b_n]$ is fixed, whereas here the length of the interval is random.

Davis [7] introduced the piecewise deterministic (*PD*) process which is a further generalization. It is similar to the piecewise linear process, except that $\xi_t$ satisfies some ordinary differential equation, rather than moving in straight line.

**Example 2.6.** *Shot noise*. This has sample functions similar to the *VWT* process except that decay between arrivals is exponential rather than linear (fig. 2.1).

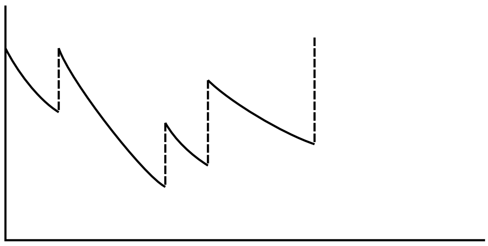

Figure 2.1:

32 **Example 2.7.** *A model for capacity expansion.* Suppose that the demand for some utility is monotone increasing and follows a Poisson process with rate $\mu$. Each 'unit' of supply provides $q$ units of capacity. These are built one at a time at a cost of Rs.p. Investment takes place at a rate of Rs.$u(t)$/week and $u(t) \leq$ constant. When

$$\int_0^t u(s)ds = p,$$

then the project is finished, capacity is increased by $q$ and investments are channelled into next project.

Denote $d_t$ = demand; $c_t$ = capacity at time $t$; $\xi_t$ = cumulative investment in current project

$$= \int_\tau^t u(s)ds,$$

where $\tau$ is the last time project was completed. Investment is determined by some "policy" $\psi$, i.e.,

$$u(t) = \frac{d}{dt}\xi_t = \psi(c_t, d_t, \xi_t)$$

where $(c_t, d_t, \xi_t)$ is the current "situation". Define $v_t = (c_t, d)$. Then the process $x_t = (v_t, \xi_t)$ evolves in the state space $E = Z_+^2 \times [o, p]$ ($Z_+^2$ is the 2-dimensional positive integer lattice). Then for $v = (c, d)$ if $g_v(\xi) = \psi(c, d, \xi), \xi_t$ satisfies $\dfrac{d}{dt}\xi_t = g_{v_t}(\xi_t)$.

**The piecewise-deterministic process;**

Let $K$ be a countable set and $d : K \to \mathbb{N}$ (= natural numbers) be a given function. For each $\nu \in K, M_\nu$ is an open subset of $\mathbb{R}^{d(\nu)}(M_\nu$ can be a $d(\nu)$-dimensional manifold). Then the state space of the $PD$ process is **33**

$$E = \underset{\nu\in K}{U} M_\nu = \{(\nu,\xi); \nu \in K, \xi \in M_\nu\}.$$

Let

$$E = \left\{ \underset{\nu\in K}{\cup'} A_\nu; A_\nu \in B(M_\nu) \right\}$$

Then $(E, E)$ is a Borel space. Then the process is $x_t = (\nu_t, \xi_t)$. The probability law of $(x_t)$ is specified by The probability law of $(x_t)$ is specified by

(i) Vector fields $(X_\nu, \nu \in K)$

(ii) A 'rate' function $\lambda : E \to \mathbb{R}_+$

(iii) A transition measure $Q : E \times E \to [0, 1]$

Assume that corresponding to each $X_\nu$ there is a unique integral curve $\phi_\nu(t,z)$, i.e., $\phi_\nu(t,z)$ satisfies

$$\frac{d}{dt} f(\phi_\nu(t,z)) = X_\nu f(\phi_\nu(t,z))$$
$$\phi_\nu(o,z) = z$$

for every smooth function $f$, and $\phi(t,z)$ exists for all $t \geq o$. Let $\partial M_\nu$ be the boundary of $M_\nu$. $\partial^* M_\nu$ is those points in $M_\nu$ at which integral curves exit from $M_\nu$, i.e., $\partial^* M_\nu = \{z \in \partial M_\nu : \phi_\nu(t,\xi) = z$ for some $(t,\xi) \in \mathbb{R} \times M_\nu\}$.

Let

$$\Gamma^* = \{\nu, z : \nu :\in K, z \in \partial^* M_\nu\}.$$

So $\Gamma^*$ is the set of points on the boundary at which jumps may take place. For $x = (\nu, z) \in E$, denote

$$t * (x) = \inf\{t > 0 : \phi_\nu(t,z) \in \partial^* M_\nu\}.$$

Write $Xh(x)$ for the function whose value at $x = (\nu, z)$ is $X_\nu h(\nu, .)(z)$. For $\lambda$, we suppose that the function $t \to \lambda(\gamma, \phi_\nu(t, z))$ is Lebesgue integrable on $[o, \in]$ for some $\in > 0, 0(., x)$ is a probability measure on $(E, E)$ **34**
for each $x \in EU\Gamma*$.

The motion of the process $(x_t)$ starting from $x = (n, z) \in E$ is described as follows. Define

$$F(t) = \begin{cases} \exp\left(-\int\limits_0^t \lambda(n, \phi_n(s, z))ds\right), & t < t^*(x) \\ \qquad 0, & t \geq t^*(x). \end{cases}$$

This is the distributions of $T_1$, the first jump time. More precisely, $F(t)$ is the *survivor function*

$$F(t) = P_x[T_1 > t].$$

Now let $Z_1$ be an $E$-valued random variable with distribution $Q(, ; \phi_n)(T_1, z))$. Then define

$$x_t = \begin{cases} (n, \phi_n(t, z)) & t < T_1 \\ \quad Z_1 & t = T_1 \end{cases}$$

and restart with $(n, z)$ replaced by $Z_1$. Assume $T_k \uparrow \infty$ a.s. Then $x_t$ defines a measurable mapping from $(\Omega, a, P)$ (countable product of unit interval probability spaces) into space of right continuous $E$- valued functions. This defines a measure $P_x$ on the Canonical space.

**NB:** The condition on $\lambda$ ensures that $T_1 > 0$ a.s and hence that $T_k - T_{k-1} > 0$ a.s.

**Proposition 2.2.** $(x_t, P_x)$ *is a Markov process.*

*Proof.* Suppose that $T_x \leq t < T_{k+1}$. The distributions of $T_{k+1} - T_k$ is given by

$$P[T_{k+1} - T_k > s] = \begin{cases} \exp\left(-\int\limits_0^s \lambda\left(\nu_{T_k}, \phi_{T_k}\right) du\right), & s < t^*(x_{T_k}) \\ \qquad 0, & s \geq t*(x_{T_k}). \end{cases}$$

35 Denote $\nu = \nu_{T_k}, \xi = \xi_{T_k}$, Then for $s > t$ and $s < t^*(x_{T_k})$

$$
\begin{aligned}
P[T_{k+l} > s|T_k, T_{k+l} > t] &= P[T_{k+l} - T_k > s - T_k|T_k, T_{k+l} - T_k > t - T_k] \\
&= \exp\left[-\int\limits_{t-T_k}^{s-T_k} \lambda(\nu, \phi(u,\xi))du\right] \\
&= \exp\left[-\int\limits_{o}^{s-t} \lambda(\nu_t, \phi_{\nu_t}(u,\xi_t))du\right]
\end{aligned}
$$

where we used the semigroup property of $\phi$. Since the process "restarts" at $T_{k+1}$, the law of the process for $s > t$ given part upto $t$ coincides with the law given $x_t$. Hence the Markov Property. ◻

Let $\Gamma \subset \Gamma^*$ be the subset for which, if $y = (\nu, \xi) \in \Gamma^*$

$$P[T_1 = T^*(x)] \to asx = (n,z) \to y.$$

Then $\Gamma$ is called the "essential" boundary

**Exercise 2.2.** *Prove that $y \in \Gamma$ if and only if $P_x[T_1 = T^*(x)] > 0$ for some $x = (\nu, z)$.*

So $P_x[x_{s-} \in \Gamma - \Gamma^*]$ for some $s > 0] = 0$, and $Q(A, x)$ need not be specified for $x \in \Gamma - \Gamma^*$.

**Example 2.8.** Here $\nu$ has only a single value; so delete it. $\xi$ takes values in $M = [0.1] \times \mathbb{R}_+, \lambda = 0$ and $X = \dfrac{\partial}{\partial \xi_1}$. Then $\Gamma^* = \{(l,y); y \in \mathbb{R}_+\}$. Let $Q(.,(l,y)) = \delta(1 - \frac{1}{2}y, \frac{1}{2}y)$.

Then starting at $x_0 = (0,1)$ we have 36

$$T_n = \sum_{k=1}^{n} \frac{1}{k-1} \text{ so that } \lim_{n\to\infty} T_n < \infty.$$

The same effect could be achieved by the combined effect of $\lambda$ and $Q$, suitably chosen. So, we prefer to assume $T_\infty = \infty$ a.s. rather than stating sufficient conditions on $\chi, \lambda, Q$ to ensure this. To illustrate the difference between $\Gamma$ and $\Gamma^*$, suppose that $\lambda(\xi_1, 1) = \dfrac{1}{1-\xi_1}$ (This is equivalent to saying $T_1$ is uniformly distributed on $[0, 1]$ if the process starts at $(0, 1)$. Then (1,1) os never hit, whatever be the starting point. So $(1,1) \in \Gamma - \Gamma^*$.

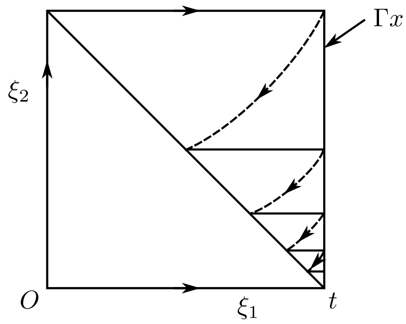


Figure 2.2:

**The Associated Jump Process.**

Let $(x_t)$ be an PD process. Definite the associated jump process $(z_t)$ by

$$z_t = x_{T_k}, t \in [T_k, T_{k+1}[ \tag{2}$$

37 This is an $E$-valued jump process such that $Z_{T_K} = x_{T_K}$. Let $F_t = \sigma\{x_s, s \leq t\}$ and $F_t^z = \sigma\{z_s, s \leq t\}$.

**Proposition 2.3.** *$F_t = F_t^z$ for each $t$*

*Proof.* This follows form the fact that there is a one-to-one mapping from $x_{[o,t]}$ to $z_{[o,t]}.x \rightarrow z$ is given by (2). Conversely, if $z_{[o,t]}$ is given then $x[o, t]$ can be constructed since the motion in the interval $[T_k, T_{k+1}[$ is deterministic. ☐

**NB:**

(1) $x_t$ and $z_t$ are not in one-to-one correspondence at each fixed time $t$.

(2) $(z_t)$ is not a Markov process.

Since $F_t = F_t^z$, we can apply jump process theory. Define

$$p(t, A) = \sum_{T_i \le t} I_{(x_{T_i} - \in A)}$$
$$p_t^* = \sum_{T_i \le t} I_{(x_{T_i} - \in \Gamma)}$$
$$\tilde{p}(t, A) = \int_0^t Q(A, x_s)\lambda(x_s)ds + \int_0^t Q(A, x_{s-})dp_s^* \tag{3}$$

**Proposition 2.4.** *Suppose $E(p(t, E)) < \infty$. Then for each*

$$A \in E, q(t, A) = p(t, A) - \tilde{p}(t, A) \tag{4}$$

*is an $F_t$-martingale.*

*Proof.* From previous results, the compensator of $p(t\Lambda T_1, A)$ is

$$\tilde{p}(t\Lambda T_1, A) = -\int_{]o,t\Lambda T_1]} Q(A, x_{s-})\frac{dF_s}{F_{s-}}.$$

But 38

$$F_t = \begin{cases} \exp\left(-\int_0^t \lambda(x_s)ds\right) & t < t_1^*(x) \\ 0 & t \ge t_1^*(x). \end{cases}$$

Thus -$\dfrac{dF_t}{F_t} = \lambda(x_t)dt$ for $t < t_1^*(x)$ and

$$\frac{\triangle F_{t_t^*}}{F_{t_1^*}-} = 1.$$

□

This verifies the result for $t \le T_1$. As before, we show by considering intervals $[T_{k-1}, T_k]$ that the compensator of $p(t\Lambda Tn, A)$ is $\tilde{p}(t\Lambda T_n, A)$

given by (3). Since $p(t,A)$ and $\tilde{p}(t,A)$ are monotonic increasing functions and $T_n \uparrow \infty$ a.s. $E(p(t,E)) < \infty$, taking the limits, we have

$$q(t,A) = p(t,A) - \tilde{p}(t,A)$$

is a martingale.

**Exercise 2.3.** *Show that $p_t^*$ is an $F_t-$ predictable process.*

Then (4) is the Doob-Meyer decomposition of the submartingale $p$.

The next step is to use stochastic integrals to calculate the extended generator of $x_t$. Choose the following integrands. For Measurable $f : \underline{E} \to \mathbb{R}$, define

$$Bf(x,s,\omega) = f(x) - f(x_{s-}(\omega))$$

Then Bf $\in L_1(p)$ if

$$E \sum_{T_i \le t} |f(x_{T_i}) - f(x_{T_i-})| < \infty$$

39 for each $t \geq 0$. This certainly holds if $f$ is bounded and $E\ p\ (t,E) < \infty$.

$$\int_o^t \int_E Bf(y,s,w)\tilde{p}(ds,dy) = \int_{[0,t]} \int_E (f(y) - f(x_{s-}))Q(dy;x_{s-})\lambda(x_s)ds + \int_{[0,t]} \int_E (f(y) - f(x_{s-}))Q(dy;x_{s-})ds_s. \quad (5)$$

Suppose that $f$ satisfies the boundary condition

$$f(x) = \int_E f(y)Q(dy;x),\ x\ \in\ \Gamma. \quad (6)$$

Then the second integral in (5) is zero. The following result characterizes the extended generator $A$ of $(x_t)$.

**Theorem 2.1.** *The domain $D(A)$ of the extended generator $A$ of $(x_t)$ consists of those functions $f$ satisfying*

(i) *For each* $(n,z) \in E$ *the function* $t \to f(n, \phi_n(n,z))$ *is absolutely continuous for* $t \in [0, t^*(n,z)[$.

(ii) *The boundary condition (6) is satisfied.*

(iii) $Bf \in L_1^{\text{loc}}(p)$.

*Then for* $f \in D(A)$

$$Af(x) = Xf(x) + \lambda(x) \int\limits_E [f(y) - f(x)]Q(dy;x). \tag{7}$$

*Proof.* Suppose that $f$ satisfies (i)-(iii). Then $\int$ Bf $dq$ is a local martingale, and

$$\int\limits_0^t Bf dq = \sum_{T_i \le t} f(x_{T_i}) - f(x_{T^-_{i^-}}) - \int\limits_0^t \int\limits_E [f(y) - f(x_s)]Q(dy;x_s)\lambda(x_s)ds.$$

Now, 40

$$\sum_{T_i \le t} f(x_{T_i}) - f(x_{T_i^-}) = \left[\sum_{T_i \le t}(f(x_{T_i}) - f(X_{T_{i-1}})) + f(x_t) - f(x_{T_n})\right]$$
$$- \left[\sum_{T_i \le t}(f(x_{T^-_i}) - f(x_{T_{i-1}})) + f(x_t) - f(x_{T_n})\right]$$

where $T_n$ is the last jump time before $t$. The first bracket is $(f(x_t) - f(x_o))$. Note that

$$f(x_{T_i^-}) - f(x_{T_{i-1}}) = \int\limits_{T_{i-1}}^{T_i} X_{\nu_{T_{i-1}}} f(\nu'_{T_{i-1}} \phi_{\nu_{T_{i-1}}}(\xi T_{i-1}, s) ds \text{ a.s}).$$

□

So the second bracket is equal to $\int\limits_0^t xf(x_s)ds$ and

$$\int Bf dq = f(x_t) - f(x_0) -$$
$$\int\limits_o^t \int\limits_E (f(y) - f(x_s))Q(dy, x_s)\lambda(x_s)ds - \int\limits_o^t Xf(x_s)ds.$$

So $Af$ is given by (7) and $C_t^f = \int\limits_0^t$ Bf $dq$. Conversely, suppose $f \in D(A)$. Then there exists a function $h$ such that $s \to h(x_s)$ is Lebesgue integrable and $M_t = f(x_t) - f(x_0) - \int\limits_0^t h(x_s)ds$ is a local martingale. By the martingale representation theorem, $M_t = M_t^g$ for some $g \in L_1^{1oc}(p)$. Now the jumps of $M_t$ and $M_t^g$ must agree, these only occur when $t = T_i$ for some $i$ and are then given by

$$\begin{aligned}\triangle M_t &= M_t - M_{t-} = f(x_{T_i}) - f(x_{T_i^-}).\\ \triangle M_t^g &= M_t^g - M_{t-}^g\\ &= g(x_t, t, w) - \int\limits_E g(y, t, w)Q(dy, x_{t-})I_{(x_{t-}\in\Gamma)}\end{aligned}$$

at $t = T_i$. It follows that

$$g(x, t, w)I_{(x_t \notin/\Gamma)} = (f(x) - f(x_{t-}))I_{(x_{t-}\notin\Gamma)}$$

41 except possibly on a set $G \in E * p$ such that

$$E_y \int\limits_{\mathbb{R}_+\times E} I_G p(dt, dx) = 0 \text{ for all } y \in E.$$

Now suppose $X_{T_i} = z \in \Gamma$; then

$$f(x) - f(z) = g(x, t, \omega) - \int\limits_E g(y, t, \omega)Q(dy; z)$$

for all $x$ except a set $A \in E$ such that $Q(A, z) = 0$. Since only the first terms on the left and right involve $x$ it must be the case that

$$f(x) = g(x, t, \omega) + \tilde{f}(t, \omega)$$

and $$f(z) = \int\limits_E g(y, t, \omega) Q(dy; z) + \tilde{f}(t, w)$$

for some predictable process $\tilde{f}$. Since $g = f - \tilde{f}$,

$$f(z) = \int\limits_E f(y) Q(dy; z)$$

for $z \in \Gamma$, i.e., $f$ satisfies condition (ii). Hence

$$g(x, t, \omega) = f(x) - f(x_{t-}).$$

Hence we get

$$\|(Bf - g) I_{(t<\sigma_n)}\| L_1(p) = 0.$$

So condition (iii) is satisfied. Fix $\omega$ and consider $(M_t)_{o \le t < t_1(\omega)}$ starting at $(\nu_o, \xi_o)^{\cdot}$, then

$$M_t = f(\nu_o, \phi_{\nu_o}(t, \xi_o)) - f(\nu_o, \xi_o) - \int\limits_o^t h(x_s) ds$$

$$M_t^g = \int\limits_o^t \int\limits_E (f(y) - f(x_s)) Q(dy; x_s) \lambda(x_s) ds.$$

Hence $f(\nu_o, \phi_{\nu_o}(t, \xi_o))$ is absolutely continuous for $t < T_1(\omega)$. Since 42
$(\nu_o, \xi_o)$ is arbitrary and $T_1(w) > 0$ a.s. this shows that (i) is satisfied.

**A "Feynman-Kac" formula.**

This is used to calculate expected values of functionals such as

$$E_x \left[ \int\limits_o^t e^{-\alpha s} c(s, x_s) ds + e^{-\alpha t} \phi(x_t) \right].$$

There is no extra generality in allowing a *P.D.* Process to be time-varying, because time can always be included as one component of $\xi_t$.

However, it is sometimes convenient to consider the joint process $(t, x_t)$ with generator $\tilde{A} = \dfrac{\partial}{\partial t} + A$. Then for $f \in D(\tilde{A})$

$$f(t, x_t) - f(o, x_o) = \int_o^t \left(\frac{\partial}{\partial s} + A\right) f(s, x_s) ds + \int_o^t Bf \, dq.$$

If $\left(\dfrac{\partial}{\partial s} + A\right) f(s, x_s) = o$ and $Bf \in L_1(p)$, then $f(t, x_t)$ is a martingale, so it has constant expectation

$$E_{x_o} f(t, x_t) = f(o, x_o).$$

Then

$$f(o, x_o) = E_{x_o} \phi(x_t)$$

where $$f(t, x) = \phi(x) \quad (\phi \text{ prescribed}).$$

**Proposition 2.5.** *Let $t > o$ be fixed and $\alpha : [o, t] \times E \to \mathbb{R}_+, c : [o, t] \times E \to \mathbb{R}$ and $\phi : E \to \mathbb{R}$ be measurable functions. Suppose $f : [o, t] \times E \to \mathbb{R}$ satisfies:*

$$\left.\begin{array}{ll} \text{(i)} & f(s, . \in D(\tilde{A})) \\ \text{(ii)} & f(t, x) = \phi(x), \ x \in E \\ \text{(iii)} & Bf \in L_l(p) \end{array}\right\} \tag{8}$$

$$\frac{\partial f(s, x)}{\partial s} + Af(s, x) - \alpha(s, x) f(s, x) + c(s, x) = 0 \tag{9}$$

$$(s, x) \in [o, t[ \times E.$$

43 *Then*

$$f(o, x) = E_{o,x} \left[ \int_0^t \exp\left( - \int_o^s \alpha(u, x_u) du \right) c(s, x_s) ds + \exp\left( - \int_o^t \alpha(u, x_u) du \right) \phi(x_t) \right] \tag{10}$$

*Proof.* Suppose $f$ satisfies (8). Define

$$e_s = \exp\left(-\int_o^s \alpha(u, x_u)du\right).$$

Then

$$\begin{aligned} d(e_s f(s, x_s)) &= e_s df(s, x_s) + f(s, x_s)de_s \\ &= e_s\left(\frac{\partial f}{\partial s} + Af\right)ds + e_s\, Bf\, dq - \alpha(s, x_s)e_s f\, ds \\ &= -e_s\, c(s, x_s)ds + e_s\, Bf\, dq \quad \text{(by (9))}. \end{aligned}$$

Now by (iii), $e_s\, Bf \in L_1(p)$ since $e_s \leq 1$. Thus the last term is a martingale and

$$E_x[e_t f(t, x_t) - f(o, x)] = -E_x\left[\int_o^t e_s c(s, x_s)ds\right].$$

This with (ii) gives (10). ◻

**Example 2.9.** *The Renewal Equation:*

Let $(N_t)$ be a renewal process with inter arrival density $f(.)$. Let $m(t) = EN_t$. Since the process "restarts" at renewal times,

$$E[N_t|T_1 = s] = \begin{cases} 0, & s > t \\ m(t-s) + 1, & s < t. \end{cases}$$

So

$$m(t) = \int_o^\infty E[N_t|T_1 = s]\, f(s)ds$$

which gives the renewal equation **44**

$$m(t) = \int_o^t (1 + m(t-s))f(s)\, ds. \tag{11}$$

This can be solved by Laplace transforms. Defining

$$\hat{f}(p) = \int_{o}^{\infty} e^{-pt} f(t)dt$$

etc., we get

$$\hat{m}(p) = \hat{m}(p)\ \hat{f}(p) + \frac{1}{p}\ \hat{f}(p).$$

So

$$\hat{m}(p) = \frac{\frac{1}{p}\hat{f}(p)}{1 - \hat{f}(p)}.$$

In particular, for the Poisson process $f(t) = \lambda e^{-\lambda t}$,

$$\hat{f} = \frac{\lambda}{\lambda + p}$$

will give

$$\hat{m}(p) = \frac{\lambda}{p^2}$$

to get

$$m(t) = \lambda t.$$

**Exercise 2.4.** *Compute $M_\tau(t) = E_\tau N_t$, where the component in service at time 0 has age $\tau$ (and is replaced by a new component when it fails).*

$(N_t)$ is a *PD* process if we take $x_t = (\nu_t, \xi_t)$ where $\nu_t = N_t$ and $\xi_t$ is the time since last renewal. Then

$$X_\nu = \frac{\partial}{\partial \xi}, \quad \lambda(\xi) = \frac{f(\xi)}{\int_{\xi}^{\infty} f(u)du}$$

45 and $Q(.; \nu, \xi) = \delta_{(\nu+1,0)}$, so that

$$Af(\nu, \xi) = \frac{\partial}{\partial \xi} f(\nu, \xi) + \lambda(\xi)[f(\nu + 1, 0) - f(\nu, 0)].$$

Use proposition 2.5 with $\alpha = c = 0$ and $\phi(x) = \nu$ to get

$$f(0, \nu, \xi) = E_{(\nu,\xi)}\nu_t.$$

Clearly

$$f(s, \nu + 1, \xi) = f(s, \nu, \xi) + 1.$$

Define

$$f(s, 0, \xi) = h(s, \xi).$$

Then the equation for $f$ (or $h$) becomes

$$\frac{\partial}{\partial s} h(s, \xi) + \frac{\partial}{\partial \xi} h(s, \xi) + \lambda(\xi)[1 + h(s, 0) - h(s, \xi)] = 0 \tag{12}$$
$$h(t, \xi) = 0.$$

Define

$$z(u) = h(u, u).$$

Then

$$\frac{d}{du} z(u) = -\lambda(u)[1 + h(u, 0) - z(u)]$$
$$z(t) = 0.$$

Thus $z(u)$ satisfies

$$\dot{z}(u) = \lambda(u)\, z(u) - \lambda(u)[1 + h(u, 0)] \tag{13}$$

where

$$\lambda(u) = -\frac{\dot{F}}{F} = \frac{f}{F}, F(u) = \int_u^\infty f(s)\, ds.$$

Equation (13) is a linear $ODE$ satisfied by $z(u)$. The transition function corresponding to $\lambda(u)$ is

$$\phi(u, v) = \frac{F(u)}{F(v)}$$

Hence (13) has the following solution at time 0: **46**

$$z(o) = h(o, o) = \int_{o}^{t} f(u)[1 + h(u, o)]du \tag{14}$$

Define

$$m(s) = h(t - s, o).$$

Then (14) coincides with the renewal equation (11). Having determined $h(u, o)\, o \leq u \leq t, h(s, \xi)$ for $s \neq \xi \neq o$ can be calculated from (12). The result will be equivalent to that of Exercise 2.4.

# Chapter 2

# Optimal Control of pd Processes

General formulations of stochastic control problems have been studied 47
using martingale theory, where the conditions for optimality, existence of optimality are derived (E1 Karoui [15]). But this does not give ways of computing optimal control. Control of Markov jump processes has been studied using dynamic programming (Pliska [21]). In this Chapter, we will be dealing with control theory for *PD* processes, following Vermes [25].

Let $Y$ be a compact metric space. Control arises when the system functions $X, \lambda, Q$ contain a parameter $y \in Y$ i.e., for $x = (\nu, \xi)$

$$\begin{aligned} X^y f(x) &= \sum_i b(\nu, \xi, y) \frac{\partial f(\nu, \xi)}{\partial \xi_i} \\ Q &= Q(A, x, y). \end{aligned}$$

A feedback *policy* (or *strategy*) is a function $u : \mathbb{R}_+ \times E \to Y$. Let $\mathscr{U}$ denote the set of all strategies. $u$ is *stationary* if there is no $t$-dependence, i.e., $u : E \to Y$. Corresponding to policy $u$ we get a *PD* process with characteristics $X^u, \lambda^u, Q^u$ given by

$$X^u f = \sum_i b(\nu, \xi, u(x)) \frac{\partial f}{\partial \xi_i}(x)$$

$$\lambda^u(x) = \lambda(x, u(x))$$
$$Q^u(A; x) = Q(A; x, u(x)).$$

More conditions on $u$ will be added when required. Then we get a *PD* process $x_t$ with probability measure $P^u$ determined by $X^u, \lambda^u, Q^u$.

48 Given a cost function, say, for example,

$$J_x(u) = E_x^u \left[ \int_o^t e^{-\alpha s} c(x_s, u_s) ds + e^{-\alpha t} \phi(x_t) \right]$$

where $E_x^u$ is the expectation w.r.t. $P^u$ starting at $x$ and $\alpha > 0$. The control problem is to choose $u(.)$ to minimise $J_x(u)$. The "usual" approach to such problems is via "dynamic programming". Let $V(s, x)$ be a function of $(s, x)$. Introduce the Bellman-Hamilton-Jacobi equation

$$\frac{\partial V(s.x)}{\partial s} + \min_{y \in Y} [A^y V(s, x) + c(x, y)] - \alpha\ V(s, x) = 0 \tag{B}$$

where $A^y$ is the generator corresponding to $X^y$, $\lambda(., y)$, $Q(., y)$.

If $Y$ has one point, then this coincides with the equation for $J_x$ as before.

**Proposition 1.** *Suppose* $(B)$ *has a "nice" solution (i.e., satisfies boundary condition etc.). Then*

$$V(o, x) = \min_{u \in \mathscr{U}} J_x(u)$$

*and the optimal strategy* $u^o(s, x)$ *satisfies*

$$A^{u^o(s,x)} V(s, x) + c(x, u^o(s, x)) = \min_{y \in Y} (A^y\ v + c).$$

*Proof.* Same calculations arise as before. Let $x_t$ correspond to an arbitrary control policy $u$. Then

$$\begin{aligned} d(e^{-\alpha s} V(s, x_s)) = -\alpha e^{-\alpha s} V(s, x_s) ds + e^{-\alpha s} \left( \frac{\partial V}{\partial s} + A^u V \right) ds \\ + e^{-\alpha s} BV\ dq \geq -e^{-\alpha s} c(x_s, u(s, x_s)) ds + e^{-\alpha s} Bu\ dq \end{aligned} \tag{1}$$

49 So

$$V(o,x) \leq E_x \left[ \int\limits_o^t e^{-\alpha s}\, c(x_s,u_s)ds + e^{-\alpha t}\, \phi(x_t) \right]$$
$$= J_x(u).$$

Now suppose $u = u^0$, then "equality" holds in place of "inequality" in (1). So

$$V(0,x) = J_x(u_0).$$

So $u^o$ is optimal. ◻

**Objections:**

(1) There is no general theory under which $(B)$ has solution.

(2) $u^o(x)$ constructed as above may fail to be an admissible control: to make sense of it, we must be able to solve the ODE

$$\frac{d}{ds}\xi(s) = b_\nu^{u_o}(\xi_s) = b(\nu,\xi,u^o(s,\xi)).$$

There is no guarantee that $u^o$ leads to a "solvable" ODE.
So we must redefine "admissible controls" so that this is avoided.

**Remark 1.** In control of diffusion processes, the equation is

$$dx_t = b(x_t,u(x_t))dt + \sigma(x_t)dW_t.$$

Here we "handle" nonsmooth $u$ by using weak solutions.

**Remark 2.** In deterministic control, one uses open-loop controls depending only on time. The equation here is of the form

$$\dot{x} = b(x_t,u(t)).$$

Then solution is well defined for the measurable $u(.)$.

**Special cases:** 50

(1) Control only appears in $Q$. Then the problem reduces to a sequential decision problem where a "decision" is taken each time a jump occurs. (Rosberg, Varaiya and Walrand [22]).

(2) $X = 0$. Here Markov jump process with piecewise constant paths are considered. Control appears in $\lambda$ and $Q$.

Then

$$A^u_f = \lambda(x, u(x)) \left( \int_E (f(z) - f(x)) Q(dz; x, u(x)) \right)$$

is a bounded operator on $B(E)$. Regard $(B)$ as an ODE in Banach space $B(E)$. Let $V(s) := V(s, .)$, then

$$\frac{dV}{ds} = g(V(s)) = \min_{y \epsilon Y} (A^y V + c).$$

So $g$ is a nonlinear function, but it is Lipschitz continuous in $V$ [Pliska [21]].

(3) Piecewise linear processes (Vermes [25]).

Here $\xi_t$ is on dimensional and $X = \frac{\partial}{\partial \xi}$. Control appears in $\lambda$ and $Q$.

Consider a 'stationary' control problem, where the Bellman equation takes the form

$$\min_{y \epsilon Y} (A^y V + c(x, y)) = 0$$
$$V(x) = \Phi(x), \quad x \epsilon E_T.$$

This corresponds to minimising

$$E_x \left( \int_0^\tau c(x_s, u_s) ds + \Phi(x_\tau) \right),$$

where $\tau$ is the first hitting time of some target set $E_T$. Then

$$A^y V(x) = \frac{\partial}{\partial \xi} V(\nu, \xi) + \lambda(x, y) \int_E (V(z) - V(x)) Q(dz, x, y).$$

51 Suppose $\nu \epsilon \{1, 2, \ldots, n\}$ and

$$V(\xi) = \begin{pmatrix} V(l, \xi) \\ V(N, \xi) \end{pmatrix}$$

Then Bellman equation takes the form

$$\frac{d}{d\xi} V(\xi) = g(V(.)).$$

This is an "ordinary" functional differential equation with non-standard boundary condition. Vermes showed existence of an optimal feedback strategy in special cases.

**'Generalised' Dynamic Programming Conditions:**

Let us consider next optimal control of the deterministic differential system:

$$\dot{x}_t = f(x_t, t, u_t), \ t \epsilon [t_o, t_l]. \tag{2}$$

Then the control problem is to

$$\text{minimize} \int_{t_o}^{t_l} \ell(x_t, t, u_t) dt$$

over "admissible" control/trajectory pairs $u_t, x_t$ i.e., pairs of functions for which

(i) (2) is satisfied,

(ii) $x(t_l) = x_l$, $x(t_o) = x_o$ with $x_o, x_l$ given,

(iii) $x_t \epsilon \bar{A}$, $u_t \epsilon \Omega$, where $\bar{A}, \Omega$ are compact and $A = \bar{A} \times [t_o, t_l]$.

This will be called the *strong problem* $(S)$.

We assume (a) an admissible pair $(x_t, u_t)$ exists, and also we make a temporary assumption

(b) $\begin{pmatrix} f(x, t, \Omega) \\ \ell(x, t, \Omega) \end{pmatrix}$ is convex. 52

This enables "relaxed controls" to be avoided. Define

$$\eta(S) = \text{“value” of } S$$

i.e., $$\inf_{(x_t,u_t)ad} \int \ell dt.$$

**Theorem 1.** *There exists an optimal admissible pair* $(x_t, u_t)$ *for the strong problem.*

This is a "standard" result in optimal control theory (Vinter and Lewis [26]). It depends critically on the convexity assumption (b).

**A Sufficient Condition for Optimality:** (Standard Dynamic Programming). Suppose $(x_t, u_t)$ is admissible and $\Phi$ is in $C^1(A)$ such that

$$\Phi_t(t,x) + \max_{u\epsilon\Omega}(\Phi_x(x,t) f(x,t,u) - \ell(x,t,u)) = 0$$
$$\Phi(x_l, t_l) \in A \times \Omega$$

and $$\Phi_t(t,x_t) + \Phi(x_t,t) f(x_t,t,u_t) - \ell(x_t,t,u_t) = 0 \quad \text{a.a.t.}$$

then $(x_t, u_t)$ is optimal and $\eta(S) = -\Phi(x_o, t_o)$. The main result of Vinter and Lewis is as follows.

**Theorem 2.** *The strong problem has a solution (i.e., there exists an optimal pair* $(x_t, u_t)$*). There exists a sequence* $\{\Phi^i\}$ *in* $C^1(A)$ *such that*

$$\Phi^i_t + \max_{u\epsilon\Omega}(\Phi_x f - \ell) \leq o, (x,t)\epsilon A$$
$$\Phi^i(x_1, t_1) = 0$$

53 *and* $(x_t, u_t)$ *is optimal if and only if*

$$\lim_{t\to\infty} H^i(t) = 0 \text{ in } L_l[t_o, t_1]$$

*where* $$H^i(t) = \Phi^i_t(x_t,t) + \Phi^i_x(x_t,t) f(x_t,t,u_t) - \ell(x_t,t,u_t).$$

**The Weak Problem**

For $(x_t, u_t)$ admissible, define $\mu_{x,u}\epsilon C^*(A \times \Omega)$, the dual of $C(A \times \Omega)$, by

$$< g, \mu_{x,u} >= \int_{t_o}^{t_l} g(x_t,t,u_t)\, dt$$

for arbitrary $g\epsilon C(A \times \Omega)$. $\mu_{x,u}$ satisfies

(i) $\mu_{x,u} \epsilon P^+$ (i.e., if $g \geq o$ then $< g, \mu_{x,u} > \geq o$).

(ii) Take $\phi \epsilon C^1(A)$ and $g(x,t,u) = \phi(x,t) + \phi_x(x,t) f(x,t,u)$

then

$$< \phi_t + \phi_x f, \mu_{x,u} >= \phi(x_l, t_l) - \phi(x_o, t_o).$$

Define

$$\mu = \{\mu \epsilon C^*(A \times \Omega) : \text{ (i) and (ii) are satisfied}\}.$$

**Proposition 2.** *$\mu$ is weak* compact and convex.*

**Note 1.** The cost function for $(x_t, u_t)$ is $< \ell, \mu_{x,u} >$.

**Weak Problem (W):** Minimise $< \ell, u >$ over $\mu \epsilon \mu$. So 54

$$\eta(W) \leq \eta(S).$$

**Theorem 3.** *$\eta(S) = \eta(W)$. There exists an optimal $x, u$ for $S$, so $\mu_{x,u}$ is optimal for $W$.*

Now we incorporate the constraints on $\mu$ into the cost function in the following way. Define extended real valued functions $p, q$ on $C^*(A \times \Omega)$ as follows:

$$p(\mu) = \begin{cases} < \ell, \mu > & \text{if } |\mu| \leq t_l - t_o, \mu \epsilon P^+ \\ +\infty & \text{otherwise}. \end{cases}$$

Let $\mathscr{M}_2 = \{\mu \epsilon C^* : \text{ Condition (ii) is satisfied}\}$. Then

$$q(\mu) = \begin{cases} o & \text{if } \mu \epsilon M_2 \\ -\infty & \text{otherwise}. \end{cases}$$

**Proposition 3.** *$p$ is $\ell$. s.c. and convex,*
*$q$ is u.s.c. and concave,*

*and*

$$\eta(W) = \inf_{\mu \epsilon C^*} \{p(\mu) - q(\mu)\}.$$

The Fenchel dual problem is as follows:

$$\max_{\xi\epsilon C(A\times\Omega)}(q^*(\xi)-p^*(\xi)) \tag{D}$$

where $p^*, q^*$ are "dual" functionals defined by

$$p^*(\xi)=\sup_{\mu\epsilon C^*}<\xi,\mu>-p(\mu).$$
$$q^*(\xi)=\inf_{\mu\epsilon C^*}<\xi,\mu>-q(\mu).$$

**Proposition 4.**

$$p^*(\xi)=\max_{(x,t,u)\epsilon A\times\Omega}(\xi(t,x,u)-\ell(t,x,u))^+\times(t_l-t_o).$$

55 *where* $a^+ := \max(a,0)$.

**Sketch of proof:**

$$\begin{aligned} p^*(\xi) &= \sup_{\substack{|\mu|\le t_l-t_o\\ \mu\epsilon P^+}}[<\xi,\mu>-<\ell,\mu>]\\ &= \sup_{\substack{|\mu|\le t_l-t_o\\ \mu\epsilon P^+}}[<\xi-\ell,\mu>]. \end{aligned}$$

If $<\xi,\mu>-<\ell,\mu>$ is negative, then the optimum is zero. If $<\xi,\mu>-<\ell,\mu>\ge o$, then put Dirac measure $x(t_l-t_o)$ on maximum point to get the result.

**Proposition 5.**

$$W=\{\xi\epsilon C:\xi=\phi_t+\phi_x f \textit{ for some } \phi\epsilon C^l(A)\}.$$

*Then*

$$q^*(\xi)=\begin{cases}-\infty & if\,\xi\notin\overline{W}\\ \lim\limits_i(\phi^i(x_1,t_1)-\phi^i(x_o,t_o)) & if\,\xi\epsilon\overline{W}\end{cases}$$

*where* $$\xi=\lim_i\xi^i \textit{ and } \xi^i=\phi^i_t+\phi^i_x f.$$

*Proof.* For $\xi \epsilon W$, by definition of $q$ and $q^*$, we get $q^*(\xi) = \phi(x_1, t_1) - \phi(x_0, t_o)$.

A similar argument gives the result for $u \epsilon \overline{W}$. For $\xi \notin \overline{W}$, there exists a separating hyperplane, i.e., $\overline{\mu} \epsilon C^*$ such that $< \overline{\xi}, \overline{\mu} > \neq 0, \overline{\xi} \epsilon \overline{W}$ and $< \xi, \overline{\mu} >= 0$. If $\mu \epsilon \mathscr{M}_2$, then $\mu + c\overline{\mu} \epsilon M_2$. So

$$q^*(\xi) = \inf_{\mu \epsilon M_2} < \xi, \mu >= -\infty.$$

□

Characterizing the solution of (D): **56**

$$\begin{aligned}\eta(D) = \max_{\xi \epsilon \overline{W}} \Bigg[ & \lim_i (\phi^i(x_1, t_1) - \phi^i(x_o, t_o)) \\ & - \max_{(x,t) \epsilon A \times \Omega} (\xi(t, x, u) - \ell(t, x, u)^+ (t_l, t_o)) \Bigg] \\ = \sup_{\phi \epsilon C^*} & (\phi(x_1, t_1) - \phi(x_o, t_o) - \max_{x,t,u} (\phi_t + \phi_x f - \ell)^+ (t_l - t_o)).\end{aligned}$$

It is no restriction to assume $\phi(x_1, t_1) = 0$. Then Vinter and Lewis show by an ingenious argument that

$$\eta(D) = \sup(-\phi(x_o, t_o))$$

where the supremum over $\phi \epsilon C^l$ such that $\phi(x_1, t_1) = 0$ and $(\phi_t + \phi_x f - \ell) \leq 0 \ \forall \ (x, t, u)$.

**Theorem 4.**

$$\eta(D) = \eta(W) = \eta(S).$$

*Proof.* This follows from a "standard" result in duality theory: $q^*$ is finite at some point in its domain where $p^*$ is continuous. □

Proof of the main results now follow easily. The strong problem has a solution since $\eta(S) = \eta(D)$.

$$\eta(S) = \lim(-\phi^i(x_o, t_o))$$

for some sequence of $\phi^i$'s satisfying the "Bellman inequality". The characterization of optimal pairs $(x_t, u_t)$ follows.

**Remark 3.** If the set 57
$$\begin{pmatrix} f(x,t,\Omega) \\ \ell(x,t,\Omega) \end{pmatrix}$$
is not convex, them the results are still valid but *relaxed controls* must be used.

A relaxed control $\mu_t$ is a $C^*(\Omega)$-valued function on $[t_o, t_1]$ such that $\mu_t$ is a probability measure for every $t$ and $t \to \int g(t,u)\mu_t(du)$ is measurable for every continuous function $g$.

**Interpretation:** $x_t$, $\mu_t$ is an admissible pair, whenever
$$\frac{dx_t}{dt} = \int\limits_{\Omega} f(x_t,t,u)\mu_t(du);$$
the cost is
$$\int\limits_{t_o}^{t_l}\int\limits_{\Omega} f(x_t,t,u)\mu_t(du)dt.$$

**Optimal Control PD Processes (Vermes [25])**

In this section, we adopt a slightly modified definition of the PD process $(x_t)$. It will take values in $E$, a closed subset of $\mathbb{R}^d$, and we suppose that
$$E = E_o U E_\partial U E_T \qquad \text{(disjoint)}$$
where $E_T$ is a closed set, $E_o$ is an open set and
$$E_\partial = (\bar{E}_o - E_o) - E_T.$$

Let $E'_o$, $E'_\partial$, $E'_T$ be compactification of $E_o$, $E_\partial$, $E_T$ respectively and $E'$ be the disjoint union of $E'_o$, $E'_\partial$, and $E'_T$. Then a controlled PD process is determined by functions
$$f : E'_o \times Y \to \mathbb{R}^d;$$
$$\lambda : E'_o \times Y \to \mathbb{R}_+,$$
and
$$Q : (E'_o U E') \times Y \to m_1(E_o)$$
58 where $m_l(E_o)$ is the set of probability measures on $E_o$ and $Y$ is a compact

(control) space. Function $f$ gives the deterministic motion by

$$\dot{x}_t = f(x_t, y_t).$$

We assume $f$ satisfies a Lipschitz condition in $x$.

**Admissible Controls:** Feed back controls $u(t) = u(x_t)$ are not the "right" class of controls because the equation $x = f(x, u(x))$ only has a unique solution under strict conditions on $u(.)$. Let

$$\begin{aligned} \alpha_t &= \text{ last jump time before } t. \\ n(t) &= x_{\alpha(t)} \\ z(t) &= t - \alpha(t). \end{aligned}$$

Then $n(t)$, $z(t)$ determine $x_t$; in fact, for fixed $y \epsilon Y$,

$$x_t = X_{n(t),z(t)}$$

where
$$X_{n,z} = n + \int\limits_o^z f(X_{n,s}, y) ds.$$

Then admissible controls are $Y$-valued measurable functions $u(n(t), z(t))$. By Caratheodory's theorem, the equation

$$X_{n,z} = n + \int\limits_o^z f(X_{n,s}, u(n, s)) ds$$

has a unique solution, and PD process is well defined for such $u$. We **59**
will consider the three component process $(x_t, z_t, n_t)$ for notational convenience.

**Relaxed controls** are functions $\mu : E \times \mathbb{R}_+ \to m_1(Y)$ such that $(n, z) \to \int \phi(n, z, y)\mu(dy; n, z)$ is a measurable function for $(n, z)$ for all continuous $\phi$. Corresponding to $\mu$, define

$$\begin{aligned} f^\mu(x, n, z) &= \int f(x, y)\mu(dy, n, z) \\ \lambda^\mu(x, n, z) &= \int \lambda(x, y)\mu(dy; n, z) \end{aligned}$$

$$Q^{\mu}(A,x,n,z) = \int Q(A,x,y)\mu(dy,n,z).$$

Then we construct a PD process $(x_t, n_t, z_t)$ corresponding to $f^{\mu}$, $\lambda^{\mu}$, $Q^{\mu}$ in the usual way.

The *strong problem* is to minimise $J_{\hat{x}_o}(\mu)$ over admissible relaxed controls $\mu$, where $\hat{x}_o = (x,x,o)$, and

$$J_{\hat{x}_o}(\mu) = E^{\mu}_{x_o}\left(\int\limits_{o}^{\tau}\int\limits_{Y} \ell_o(x_t,y)\mu(dy;n_t,z_t)dt \right.$$
$$\left. + \sum_{\{t:x_{t-}\epsilon E_\partial\}} \int \ell_o(x_{t-},y)\mu(dy;n_{t-},z_{t-}) + \ell_T(x_\tau)\right).$$

Here $\tau$ is the first hitting time of set $E_T$.

**Main Results:**

**Theorem 5.** *There exists an optimal (relaxed) control.*

**Theorem 6.** *The value function $\psi(x) = \sup\phi(x,x,o)$ where the supremum is over all functions $\phi\epsilon C^1(E)$ such that*

$$\phi_z(x,n,z) + \min_{y\epsilon Y}\left(\nabla_x\phi(x,n,z)f(x,y) + \lambda(x,y)\right.$$
$$\left.\left(\int\phi(\xi,\xi,o)Q(d\xi,x,y) - \phi(x,n,z)\right) + \ell_o(x,y)\right)\quad (x,n,z)\epsilon\tilde{E}_o \tag{3}$$
$$\geq 0$$

$$\phi(x,n,z) \leq \min_{y\epsilon Y}\left\{\int\phi(\xi,\xi,o)Q(d\xi,y,x) + \ell_o(x,y)\right\} (x,n,z)\epsilon\tilde{E}_\partial \tag{4}$$

$$\phi(x,n,z) \leq \ell_T(x),\quad x\epsilon E_T \tag{5}$$

60 *and $\tilde{E}$ is the space of triplets $(x,n,z)$.*

**Theorem 7.** *There exists a sequence $\phi^k$ satisfying (3), (4), (5) above such that $\mu^o$ is optimal if and only if*

$$\phi^k(x,n,z) + \int\limits_{Y}\left\{\nabla_x\phi^k(x,n,z)f(x,y) + \lambda(x,y)\right\}$$

$$\left[\int \phi^k(\xi,\xi,0)\times Q(d\xi,x,y) - \phi^k(x,n,z)\right]$$
$$+\ \ell_o(x,y)\mu^o(dy,n,z) \to o \ in\ L_l(Q^o_o). \qquad (6)$$

$$\int\limits_Y \left\{\int\limits_{E_o} \phi^k(\xi,\xi,o)Q(d\xi,x,y) + \ell_o(x,y)\right\}\mu^o(dy,n,z)$$
$$-\ \phi^k(x,n,z) \to 0\ in\ L_l(Q^o_\partial) \qquad (7)$$
$$\phi^k(x,n,z) - \ell_T(x) \to 0\ in\ L_1(Q^o_T). \qquad (8)$$

*The measures $Q^o_o$, $Q^o_\partial$ and $Q^o_T$ are defined as follows.*
*Denote $\tilde{x}_t = (x_t, n_t, z_t)$. For $A\epsilon\tilde{E}_o$,*

$$Q^o_o(A) = \tilde{E}^{\mu_o}_{\hat{x}_o}\int\limits_o^\tau \psi_A(\tilde{x}_t)dt$$

*which is a measure on $\tilde{E}_o$ and is called potential measure of $\tilde{x}_t$.* **61**

$$Q^o_\partial(A) = \tilde{E}^{\mu_o}_{\hat{x}_o}\sum_{t\le\tau}\psi_A(\tilde{x}_{t-})$$

*where $A\in\tilde{E}_\partial$.*

$$Q^o_T(A) = \tilde{P}^{\mu_o}_{\hat{x}_o}[\tilde{x}_T\in A]$$

*for $A\in\tilde{E}_T$.*

Comparing with deterministic case, the necessary and sufficient condition there was that $(x_t,\mu_t)$ is optimal if and only if

$$\phi^i_t(x_t,t) + \int\left\{\phi^i_x(x_t,t)f(x_t,t,u) - \ell(x_t,t,u)\right\}\mu_t(du) \to o \text{ in } L_l(t_0,t_1).$$

The "probability measure" corresponding to $\mu_t$ is Dirac measure on $x(.)$ and $Q^o_o(A)$ is the time spent by $x(.)$ in $A$. Thus the conditions stated are a direct generalization of the deterministic ones.

**Remark 4.** Note that if we define

$$Q^o_o(A) = E_{\tilde{x}}\int_o^\tau \psi_A(\tilde{x}_s)ds$$

then for any positive measurable function $g$,

$$E_{\tilde{x}} \int_o^\tau g(\tilde{x}_s)ds = \int_{\tilde{E}_o} g(\xi) Q_o^o(d\xi);$$

for, if

$$g(\tilde{x}) = \sum_i c_i \psi_{A_i}(\tilde{x})$$

62 then

$$\begin{aligned} E_{\tilde{x}} \int_o^\tau g(\tilde{x}_s)ds &= \sum_i c_i E_x \int_0^\tau \psi_{A_i}(\tilde{x}_s)ds \\ &= \sum_i c_i Q_\circ^\circ(A_i) \\ &= \int_{E_o} g(\xi)\ Q(d\xi). \end{aligned}$$

The general case follows by monotone convergence.

**Remark 5.** The $Q_i^\circ$ are "potentials of additive functionals" $\ell_t$ is an *additive functional* if $\ell \geq o$ and[1]

$$\ell_{t+s} = \ell_t + \ell_s o \theta_t$$

$t, p_t^*, I_{(t \geq \tau)}$ are some example of additive functionals.

The potential of an additive functional is an operator

$$U_\ell g(\tilde{x}) = E_{\tilde{x}} \int_o^\tau g(\tilde{x}_s) d\ell_s.$$

Here $Q_O^o, Q_\partial^o, Q_T^o$ correspond precisely to this with $\ell_t = t, p_t^*, I_{(t>T)}$ respectively.

**The Weak Problem:**

[1] $\theta_t$ is the shift operator on the space of right continuous functions: $(\theta_t w)_s = \omega_{t+s}$

The deterministic weak problem involved the fact that

$$\phi(x_1, t_1) - \phi(x_o, t_o) = \int_{t_o}^{t_1} (\phi_t + \phi_x f) ds$$

for any $\phi \in C^1$. The stochastic equivalent of this is Dynkin's formula. To get this in the appropriate form, *define operators* $A^y$, $B^y$ as follows.

$$\begin{aligned} A^y \phi(x, n, z) &= \phi_z(x, n, z) + \nabla_x \phi(x, n, z) f(x, y) \\ &\quad + \lambda(x, y) \int_{\tilde{E}_O} (\phi(\xi, \xi, o) - \phi(x, n, z)) Q(d\xi, x, y) \end{aligned}$$

and $$B^y \phi(x, n, z) = \int_{E_\partial} \phi(\xi, \xi, o) Q(d\xi, x, y) - \phi(x, n, z)$$

for $(x, n, z) \in \tilde{E}_\partial$. Then the Dynkin formula on the interval $(o, \tau)$ is **63**

$$\begin{aligned} &\tilde{E}^\mu_{\hat{x}} \phi(x_\tau, n_\tau, z_\tau) - \phi(\hat{x}) \\ &\qquad = E^\mu_{\hat{x}} \left[ \int_o^\tau \int_Y A^y \phi(x_t, n_t, z_t) \mu(dy; n_t, z_t) dt \right. \\ &\qquad\quad \left. + \int_o^\tau \int_Y B^y \phi(x_t, n_t, z_t) \mu(dy, z_t, n_t) dp^*_t \right] \\ &\qquad = \int_{\tilde{E}_o} \int_Y A^y \phi(x, n, z) \mu(dy; n, z) Q^\mu_o(dx, dn, dz) \\ &\qquad\quad + \int_{\tilde{E}_d} \int_Y B^y \phi(x, n, z) \mu(dy; x, z) Q^\mu_o(dx, dn, dz) \end{aligned}$$

Now

$$\tilde{E}^\mu_{\hat{x}} \phi(x_\tau, n_\tau, z_\tau) = \int_{\tilde{E}_T} \phi(x, n, z) Q^\mu_T(dx, dn, dz).$$

So we can express the Dynkin formula as follows:

$$\phi(\hat{x}_o) = \int_{\tilde{E}\times Y} L\phi(\tilde{x}, y) M^{\mu}(d\tilde{x}, dy)$$

where $L\phi(x, n, z, y) = \psi_{\tilde{E}_T}\phi(x, n, z) + \psi_{\hat{E}_o} A^y \phi(x, n, z) + \psi_{\tilde{E}_\partial} B^Y(x, n, z)$.

$$M^{\mu}(S_1 \times S_2) = Q^{\mu}_T\left(S_1 \bigcap \tilde{E}_T\right) + \int_{S_1 \bigcap \tilde{E}_o} \int_{S_2} \mu(dy; n, z) Q^{\mu}_o(dx, dn, dz)$$
$$+ \int_{S_1 \bigcap \tilde{E}_\partial} \int_{S_2} \mu(dy; n, z) Q^{\mu}_\partial(dx, dn, dz).$$

64 The cost for the relaxed control $\mu$ is

$$J_{\hat{x}_o}(\mu) = \int_{E\times Y} \ell(\tilde{x}, y) M^{\mu}(d\tilde{x}, dy).$$

The following supplementary assumption is required.

$$\inf_{u\in u} J_{\hat{x}_o}(\mu) = \inf_{\mu\in u_o} J_{\hat{x}_o}(\mu)$$

for some $c > o$ and $u$ is the set of relaxed controls,

$$u_c = \{\mu \in u : \mu \in [\tau + p^*_\tau] \leq c\}$$

with this assumption the weak problem is to minimize $\int_{\bar{E}\times Y} \ell dM$ over measures $M \in m_{1+c}(\tilde{E} \times Y)$ (where $m_a$ is the set positive measures of total mass less than or equal to a) such that

1. $M = M_o + M_\partial + M_T$

   where
   $$M_T \in m_1(\tilde{E}_T).$$
   $$M_\partial \in m(\tilde{E}_\partial \times Y)$$
   $$M_O \in m(\tilde{E}_o \times Y).$$

2. $\phi(\hat{x}_o) = \int L\phi dM, \phi \in C^1(\tilde{E})$.

65 From this point on, the development follows the Vinter-Lewis arguments closely. We reformulate the weak problem as a convex optimization problem by incorporating the constrains in the cost function and obtain the characterization of optimality by studying the dual problem. The reader is referred to Vermes [25] for the details.

**Remark 6.** The optimality condition involves the measures $Q^o_O, Q^o_\partial, Q^o_T$ corresponding to $\mu^o$. These can be computed from the following system of equations.

$$\begin{aligned}
&A^{\mu^o} h(\tilde{x}) + \psi_{\Gamma \bigcap E_o} = 0, \tilde{x} \in \tilde{E}_O\\
&B^{\mu^o} h(\tilde{x}) + \psi_{\Gamma \bigcap \tilde{E}_\partial} = 0, \tilde{x} \in \tilde{E}_\partial\\
&h(\tilde{x}) + \psi_{\Gamma \bigcap \tilde{E}_T}(\tilde{x}), \tilde{x} \in \tilde{E}_T.
\end{aligned}$$

Then

$$Q^o(\Gamma) = h(\hat{x}_o).$$

**Example 1.** If $\Gamma \subset \tilde{E}_o$, then Dynkin's formula says

$$\begin{aligned}
h(\hat{x}_o) &= E^{\mu^o}_{\hat{x}_O} \int_O^\tau \psi_\Gamma(\tilde{x}_s) ds\\
&= Q^o_O(\Gamma).
\end{aligned}$$

The results outlined above are the first general results on optimal control of *PD* processes. Obviously much work remains to be done; natural next steps would be to determine necessary conditions for optimality of Pontrjagin type; to develop computational methods; and to study optimal stopping and "impulse control" for *PD* processes. For some related work, see van der Duyn Schouton [29], Yushkevich [30] and Rishel [31].

# Part II

# Filtering Theory

# 0 Introduction

Suppose $\{x_t\}$ is a signal process which represents the state of a system, **66**
but cannot be observed directly. We observe a related process $\{y_t\}$. Our aim is to get an expression for the "best estimate" of $x_t$, given the history of $\{y_t\}$ upto time $t$.

In Section 1, we give quick derivations of the "Kalman filter" for the linear systems, and nonlinear filtering equations, that of Fujisaki, Kallianpur and Kunita and Zakai's equation for unnormalized conditional density (Kallianpur [19], Davis and Marcus [8]). In section 2, we will study pathwise solutions of differential equations. In section 3, we will study the "Robust" theory of filtering as developed by Clark [5], Davis [10] and Pardoux [20]. Here the above filtering equations are reduced to quasi-deterministic form and solved separately for each observation sample path. Also, we will look here into some more general cases of filtering developed by Kunita [17], where the existence of conditional density functions is proved using methods related the theory of "Stochastic flows".

# 1 Linear and Nonlinear Filtering Equations Kalman Filter (Davis [11])

Suppose that the "signal process" $x_t$ satisfies the liner stochastic differ- **67**
ential equation

$$dx_t = Ax_t dt + cdV_t \tag{1}$$

where $V_t$ is some Wiener process. "Observation" $y_t$ is given by

$$dy_t = Hx_t dt + dW_t \tag{2}$$

where $W_t$ is a Wiener process independent of $V_t$. Assume $x_o \sim N(o, P_o)$. To get a physical model, suppose we write (2) as

$$\frac{dy_t}{dt} = Hx_t + \frac{dW_t}{dt}$$

then $\dfrac{dW_t}{dt}$ corresponds to white noise and $\dfrac{dy_t}{dt}$ is the " physical" observation.

The filtering problem is to calculate 'best estimate' of $x_t$ given $(y_s, s \leq t)$. There are two formulations for Kalman filter.

**(a) Strict Sense:** If $(V_t, W_t)$ are Brownian motions then $(x_t, y_t)$ is a Guassian process. Then $\hat{x}_t = E[x_t | y_s, s \leq t]$ is the "best estimate" in the sense of minimizing $E(x_t - z)^2$ over all $y_t$-measurable, square integrable random variables $z$, where

$$y_t = \sigma\{Y_s, s \leq t\}.$$

Because of normality, $\hat{x}_t$ is a liner function of $(y_s, s \leq t)$.

68 **(b) Wide Sense Formulation:** Do not suppose $V_t, W_t$ are normally distributed. Just suppose that the $i^{th}$ coordinates $V^i_t, W^i_t$ are uncorrelated and $EV^i_t V^i_s = t \wedge s$; $EW^i_t W^i_s = t \wedge s$, i.e., $V^i, W^i$ are orthogonal increment processes. Now look for the best linear estimate of $x_t$ given $(y_s, s \leq t)$. This will coincide with $E(x_t | y_t)$ in strict sense case.

Calculating $\hat{x}_t$ is a Hilbert space projection problem. The random variables we consider belong to $L^o_2(\Omega, F, \mathbb{P})$ which is a Hilbert space with inner product $(X, Y) = EXY$, where $o$ denotes the elements are of zero mean. For any process, say $y_t$ define $H^y = L(y_t, t \geq o)$, the linear span of $y_t$; this is a linear subspace. Then if $\hat{z}$ denotes the projection of $z$ onto $H^y$, then

$$\|z - \hat{z}\| = \min_{U \in H^y} \|z - U\|.$$

Let $\hat{x}_t$ be projection of $x_t$ onto $H^y_t = L(y_s, s \leq t)$. Then the "Innovations process" $\nu_t$ is defined by

$$d\nu_t = dy_t - H\hat{x}_t dt. \tag{3}$$

The Innovations process $\nu_t$ has the following properties:

(i) $\nu_t$ is an orthogonal increments process.

(ii) $H^y_t = H^\nu_t$.

(iii) $H^\nu_t = \left\{ \int\limits_o^t g(s) d\nu_s,\ g \in L_2[o, t] \right\}$.

69 Then $\hat{x}_t$ satisfies the linear equation

$$\begin{aligned} d\hat{x}_t &= A\hat{x}_t dt + P(t)H' d\nu_t \\ \hat{x}_o &= 0 \end{aligned} \tag{4}$$

where the error covariance $P(t) = E(x_t - \hat{x}_t)(x_t - \hat{x}_t)'$, (' denotes the transpose).

$$\begin{aligned} &P(t) \text{ satisfies the "Riccati equation"} \\ &\frac{d}{dt}P(t) = AP(t) + P(t)A' + CC' - P'(t)HH'P(t) \\ &P(o) = P_o = Cov(x_o). \end{aligned}$$

The above equation (4) is the *Kalman Filter*.

**Derivation of Kalman Filter equation:** From properties (ii), (iii) we know

$$\hat{x}_t = \int_o^t g(t,s) d\nu_s$$

for some $g$ such that

$$\int_o^t g^2(t,s) ds < \infty.$$

Now using projection, $x_t - \hat{x}_t \perp \nu_s, s \leq t$. So

$$\begin{aligned} E x_t \nu'_s &= E\hat{x}_t \nu'_s \\ &= E\left(\int_o^t g(t,u) d\nu_u\right)\nu'_s \\ &= \int_o^t g(t,u) du. \end{aligned}$$

Hence

$$g(t,s) = \frac{d}{ds} E x_t \nu'_s.$$

Write innovations process as 

$$d\nu_t = H\tilde{x}_t dt + dW_t \text{ where } \tilde{x}_t = x_t - \hat{x}_t.$$

$$Ex_t\nu_s' = \int_o^s E(x_t\tilde{x'}_u))H'du.$$

Now

$$x_t = \phi(t,u)x_u + \int_u^t \phi(t,r)CdV_r$$

where $\phi$ is the transition matrix of $A$. So

$$\begin{aligned} Ex_t\nu_s' &= \int_o^s \phi(t,u)(\tilde{x}_u\tilde{x'}_u)H'du \\ &= \int_o^s \phi(t,u)P(u)H'du \\ g(t,s) &= \phi(t,s)P(s)H'. \end{aligned}$$

So

$$\hat{x}_t = \int_o^t \phi(t,s)P(s)H'd\nu_s.$$

But this is the unique solutions of (4).

**Important Points:**

(1) It is a recursive estimator.

(2) In the strict sense version $\hat{x}_t$ is a sufficient statistic for the conditional distribution of $x_t$ given $(y_s, s \leq t)$, since this distribution is $N(\hat{x}_t, P(t))$ and $P(t)$ is nonrandom.

**Exercise 1.1** (Constant Signal)**.** *Let* $x_t = \theta$ *with* $E(\theta) = 0$*,* $Var\ (\theta) = \sigma^2$ *and*

$$dy_t = \theta dt + dW_t$$

71 *with $\theta$ independent of $W_t$. Show directly by projection that*

$$\hat{\theta}_t = \frac{1}{t + \frac{1}{\sigma^2}} y_t.$$

Now show that the Kalman filter gives the same result.

**Nonlinear Filtering**

Suppose "signal" $x_t$ is a Markov process and "observation" $y_t$ is given by

$$dy_t = h(x_t)dt + dW_t,$$

generally $h$ is a bounded measurable function (extra smoothness condition will be added later). Assume that for each $t$, $x_t$ and $(W_u - W_v), u, v \geq t$ are independent, which allows for the "feedback" case. Our objective is to calculate in recursive from the "estimates" of $x_t$. to do this, it is necessary to compute the condition of $x_t$ given

$$y_y = \sigma\{y_s, s \leq t\}.$$

**The Innovations Approach to Nonlinear Filtering**

This approach was originally suggested by Kailath for the linear case and by Kailath and Frost for nonlinear filtering. The definitive formulation of the filtering problem from the innovations standpoint was given by Fujiskki, Kallianpur and Kunita [18].

**Innovations Processes:** Consider process $y_t$ satisfying 72

$$dy_t = z_t dt + dW_t, t \in [o, T] \tag{5}$$

where $W_t$ is Brownian motion and assume

$$E \int\limits_o^T z_s^2 ds < \infty \tag{6}$$

and the "feedback" condition is satisfied. Let

$$\hat{z}_t = E[z_t | y_t].$$

More precisely $\hat{z}_t$ is the "predictable projection" of $z_t$ onto $y_t$. The innovations process is then

$$d\nu_t = dy_t - \hat{z}_y dt. \tag{7}$$

**Note (i):** $\nu_t$ is a Brownian motion w.r.t. $y_t$, i.e., $\nu_t$ is a $y_t$ martingale and $< \nu >_t$. If $F_t^\nu = \sigma\{\nu_s, s \leq t\}$, the question is whether $F_t^\nu = F_t^y$. It has been shown that in general, this is not true. But if (i) holds and $(z_t), (W_t)$ are independent, then Allinger-Mitter proved that $F_t^\nu = F_t^y$.

**Note (ii):** All $y_t$-martingales are stochastic integrals $w.r.t.$ $(\nu_t)$, i.e., if $M_t$ is a $y_t$-martingale, then there is a $g$ such that

$$\int_o^T g_s^2 ds < \infty \ a.s.$$

and

$$M_t = \int_o^t g_s d\nu_S.$$

This is true even if $F_t^\nu \neq F_t^y$, but note that $(g_s)$ is adapted to $F_t^y$, not necessarily to $F_t^\nu$.

73 **A General Filtering Formula:** Take an $F_t$-martingale $n_t$, process $(\alpha_t)$ satisfying

$$E \int_o^T |\alpha_s|^2 ds < \infty$$

and $F_o$ measurable random variable $\xi_o$ with $E\xi_o^2 < \infty$.
Now define an $F_t$ semi-martingale $\xi_t$ by

$$\xi_t = \xi_o + \int_o^t \alpha_s ds + n_t. \tag{8}$$

Since $< W, W >_t = t$, we have

$$< n, W >_t = \int_o^t \beta_s ds$$

for some $\beta_t$ and for any martingale $n_t$. Let

$$\hat{\xi}_t = E[\xi_t | y_t].$$

Then $\hat{\xi}_t$ satisfies the following stochastic differential equation

$$\hat{\xi}_t = \hat{\xi}_o + \int_o^t \hat{\alpha}_s ds + \int_o^t [\widehat{\xi_s z_s} + \hat{\xi}_s \hat{z}_s + \hat{\beta}_s] d\nu_s. \tag{9}$$

*Proof.* Define

$$\mu_t = \hat{\xi}_t - \hat{\xi}_o = \int_o^t \hat{\alpha}_s ds$$

□

Then $\mu_t$ is a $y_t$- martingale. So there is some integrable function $\eta$ such that

$$\mu_t = \int_o^t \eta_s d\nu_s. \tag{10}$$

Now we will identity the form of $n_t$, using ideas of Wong [28]. Using **74**
(5) and (8) and I to formula,

$$\xi_t y_t = \xi_o y_o + \int_o^t \xi_s (z_s ds + dW_s) + \int_o^t y_s (a_s ds + dn_s)$$

Now calculate $\hat{\xi}_t y_t$ using (7) and (10),

$$\hat{\xi}_t y_t = \hat{\xi}_o y_o + \int_o^t \hat{\xi}_s (\hat{z}_s ds + d\nu_s) + \int_o^t y_s (\hat{a}_s ds + \eta_s d\nu_s) + \int_o^t \eta_s ds.$$

Now for $t \geq s$,

$$E\left[\xi_t y_t - \hat{\xi}_t y_t | y_s\right] = 0.$$

So

$$E\left[\int\limits_s^t \left((\widehat{\xi_u z_u}) - \hat{\xi}_u \hat{z}_u + \hat{\beta}_u - \eta_u\right) du | y_s\right] = 0.$$

Let

$$V(u) = \widehat{\xi_u z_u} - \hat{\xi}_u \hat{z}_u + \hat{\beta}_u - \eta_u.$$

Then $V(u)$ is predictable process and

$$E\left[\int_s^t V(u) du | F_s\right] = 0.$$

This is,

$$\int\limits_{A \times [s,t]} V(u) du dP = 0 \forall s, t \geq s, A \in F_s.$$

The class of sets $A \times [s, t]$ generates $P$, the predictable $\sigma$-field. Hence $V(u, \omega) = 0$ a.e. $dt * dP$. Hence the result.

75 Formula (9) is not a recursive equation for $\hat{\xi}_t$. Still we can use it to obtain more explicit results for filtering of Markov processes. Let $(x_t)$ be a Markov process and $A, \mathcal{D}(A)$ be generator, i.e., for $f \in \mathcal{D}(A)$ then

$$C_t^f = f(x_t) - f(x_s) - \int\limits_s^t A f(x_u) du$$

is a martingale. Suppose

$$< C^f, W >_t = \int\limits_o^t Z f(x_s) ds$$

for some function $Zf$. Introduce the notation

$$\prod_t (f) = E[f(x_t) | y_t]$$

Now apply (9) with $\xi_t = f(x_t); Af(x_s) = \alpha_s, C^f_t = n_t$ and $z_t = h(x_t)$ to get Fujisaki-Kallianpur-Kunita filtering formula

$$\Pi_t(f) = \Pi_o(f) + \int_o^t \Pi_s(Af)ds + \int_o^t [\Pi_s(Df) - \Pi_s(h)\Pi_s(f)]\, dv_s \quad (11)$$

where

$$Df(x) = Zf(x) + h(x)f(x).$$

If we interpret $\Pi_t$ as the conditional distribution of $x_t$ given $y_t$, so that

$$\Pi_t(f) = \int f(x)\Pi_t(dx) = E[f(x_t)|y_t],$$

then (11) is a measure-valued stochastic differential equation, and gives an infinite-dimensional recursive equation for filtering.

**Exercise 1.2.** *Derive the Kalman filter from the Fujisaki-Kallianpur-Kunita equation.* 76

**The Unnormalized (Zakai) Equations:**

Introduce a new probability measure $P_o$ on $(\Omega, F)$ with $t \in [o, T]$ by

$$\frac{dP_o}{dP} = \exp\left(-\int_o^T h(x_s)dW_s - \frac{1}{2}\int_o^T h^2(x_s)ds\right).$$

Since $h$ is bounded, $P_o$ is probability measure and $(y_t)$ is a $P_o$-Brownian motion. Also

$$\begin{aligned} < C^f_t, y >_t &= < C^f_t, w >_t \\ &= \int_o^t Zf(x_s)ds. \end{aligned}$$

Note that, in general, $C^f_t$ is a semi-martingale under $P_o$ but $< ., . >$ is invariant under absolutely continuous change of measure. Also if $Z = o$, then $x_t$ has the same distribution under either measure. Let

$$\wedge_T = \frac{dP}{dP_o} = \exp\left( \int_o^T h(x_s) dy_s - \frac{1}{2} \int_o^T h^2(x_s) ds \right).$$

Let $E_o$ denote the expectation under $P_o$. Then it can be calculated that under measure $P_o$,

$$\begin{aligned} \Pi_t(f) &= E[f(x_t) | y_t] \\ &= \frac{E_o[f(x_t) A_t | y_t]}{E_o[A_t | y_t]} \\ &=: \frac{\sigma_t(f)}{\sigma_t(1)}. \end{aligned}$$

77 Then $\sigma_t(f)$ is an unnormalized conditional distribution since $\sigma_t(1)$ does not depend on $f$. To obtain a¡n equation satisfied by $\sigma_t$, we need a semi-martingale representation for $\sigma_t(1)$. First we have

$$d\Lambda_t = h(x_t)\Lambda_t dy_t \tag{12}$$

i.e.,

$$\Lambda_t = 1 + \int_o^t h(x_s)\Lambda_s dy_s$$

also $\Lambda_t$ is a $(F_t, P_o)$ martingale. Then as before

$$\hat{\Lambda}_t = E_o[\Lambda_t | y_t]$$

is a $y_t$-martingale, so there exists some $y_t$-adapted integrand $\eta_t$ such that

$$\hat{\Lambda}_t = 1 + \int_o^t \eta_s dy_s \tag{13}$$

To identify $\eta_t$, we use the same technique as in deriving the $FKK$ equation. Calculate using (12) and I to’s rule,

$$\Lambda_t y_t = \int_o^t \Lambda_t dy_s + \int_o^t y_s \Lambda_s h(x_s) dy_s + \int_o^t \Lambda_s h(x_s) ds.$$

Calculating using (13) and Ito's rule,

$$\hat{\Lambda}_t y_t = \int_o^t \hat{\Lambda}_s dy_s + \int_o^t y_s \eta_s dy_s + \int_o^t \eta_s ds.$$

Now

$$E_o[\Lambda_t y_t - \hat{\Lambda}_t y_t | y_s] = o \text{ for } t \geq s,$$

so we get

$$\eta_t = \Lambda\widehat{(t)h(}x_t) := E_o[\Lambda_t h(x_t) | y_t].$$

So (13) becomes ag **78**

$$\hat{\Lambda}_t = 1 + \int \hat{\Lambda}_s \Pi_s(h) dy_s \tag{14}$$

This has a unique solution

$$\hat{\Lambda}_t = \exp\left(\int_o^t \Pi_s(h) dy_s - \frac{1}{2}\int_o^t \Pi_s^2(h) ds\right)$$
$$= \sigma_t(1)$$

**Theorem 1.1.** *$\sigma_t(f)$ satisfies the "Zakai equation"*

$$d\sigma_t(f) = \sigma_t(Af)dt + \sigma_t(Df)dy_t \tag{15}$$
$$\sigma_o(f) = \Pi_o(f) = E[f(x_t)].$$

*Proof.* Direct calculation using (11), (14) and the fact that

$$\sigma_t(f) = \hat{\Lambda}_t \Pi_t(f).$$

□

**Corollary 1.1.** *There is a one-to-one relation between Zakai equation and FKK equation, in that whenever $\sigma_t$ satisfies Zakai equation, $\sigma_f(f)/\sigma_t(1)$ satisfies (11), and whenever $\prod_t(f)$ satisfies (11),*

$$\Pi_t(f) \exp\left(\int_o^t \Pi_s(h) dy_s - \frac{1}{2}\int_o^t \Pi_s(h) ds\right)$$

*satisfies Zakai equation.*

**The Zakai Equation with Stratonovich Integrals:** 79

Recall

$$\int\limits_o^t u_s \circ dv_s := \int\limits_o^t u_s dv_s + \frac{1}{2} < u, v >_t$$

where $o$ denotes a Stratonovich Stochastic integral and $u$ and $v$ are continuous semi-martingales. We have to calculate $< \sigma.(Df), y >_t$. From Zakai equation

$$d\sigma_t(Df) = \sigma_t(ADf)dt + \sigma_t(D^2 f)dy_t.$$

So

$$d < \sigma.(Df), y >_t = \sigma_t(D^2 f)dt.$$

So the Stratonovich version of the Zakai equation is

$$\begin{aligned} d\sigma_t(f) &= \sigma_t(Af)dt + \sigma_t(Df)ody_t - \frac{1}{2}\sigma_t(D^2 f)dt \\ &= \sigma_t(Lf)dt + \sigma_t(Df)ody_t \end{aligned}$$

where

$$Lf(x) = Af(x) - \frac{1}{2}D^2 f(x).$$

**Application to Diffusion Process:**

Consider a process $x_t \in \mathbb{R}^d$ satisfying

$$df(x_t) = X_o f(x_t)dt + X_j f(x_t)odB_t^j \tag{16}$$

for arbitrary smooth $f$, where $X_o, \ldots, X_r$ are vector fields on $\mathbb{R}^d$ we suppose that $< B^j, W >_t = \alpha^j t$ for some constants $\alpha^1, \ldots, \alpha^r$

80 Note that $A$ is the generator of $x_t$ under measure $P$ (not $P_o$). This is given by

$$Af(x) = X_o f(x) + \frac{1}{2}\sum_j X_j^2 f(x).$$

*Proof.* Rewrite (16) in Ito form. Replace $f$ by $X_k f$ in (16).

$$dX_k f(x_t) = X_o X_k f(x_t)dt + X_i X_k f(x_t)odB_t^j.$$

□

Then
$$d < X_k f, B^k >_t = X_k^2 f(x_t)dt.$$

Then Ito version of (16) is

$$df(x_t) = \left(X_o + \frac{1}{2}\sum X_j^2\right) f(x_t)dt + X_j f(x_t)dB_t^j.$$

So
$$A = X_O + \frac{1}{2}\sum X_j^2$$

and[2]
$$C_t^f = \int_o^t X_j f(x_s)dB_s^j.$$

**Proposition 1.1.** *For Z given by*

$$< C^f, W >_t = \int Z f(x_s)ds.$$

*with* $Zf = \sigma^j X_j, Z$ *is a vector field.*

*Proof.*

$$\begin{aligned} d < C^f, W > &= d < \int X_j f dB^j, W > \\ &= X_j f d < B^j, W > \\ &= \alpha^j X_j f(x_t)dt. \end{aligned}$$

So **81**

$$\begin{aligned} D &= Z + h \\ &= \alpha^i X_i + h. \end{aligned}$$

□

[2]We sometimes use the convention of implied summation over repeated indices.

**Proposition 1.2.** *There exist vector fields $Y_o Y_1, \ldots Y_r$ such that*

$$A - \frac{1}{2}D^2 = \frac{1}{2}\sum_j Y_j^2 + Y_o - \frac{1}{2}Dh.$$

*Proof.*

$$\begin{aligned} D^2 f &= (\alpha^i X_i + h)(\alpha^j X_j f + hf) \\ &= \alpha^i \alpha^j X_i X_j f + \alpha^i X_i(hf) + \alpha^j h X_j f + h^2 f \\ &= \alpha^i \alpha^j X_i X_j f + hZf + Dhf. \end{aligned}$$

□

Let $\alpha = (\alpha^1 \alpha^2, \ldots \alpha^r)'$ and suppose $I - \alpha\alpha'$ is nonnegative definite. Write $(I - \alpha\alpha') = \Delta\Delta'$ and let $X = (X_1, \ldots, X_r)'$.

$$X'\Delta\Delta' X = \sum X_i^2 - \alpha^i \alpha^j X_i X_j.$$

So define $Y = \Delta' X$. Then $Y_i' s$ are vector fields and

$$\begin{aligned} A - \frac{1}{2}D^2 &= \frac{1}{2}\sum_i Y_i^2 - hZ + X_o - \frac{1}{2}Dh \\ &= \frac{1}{2}\sum_i Y_i^2 + Y_o - \frac{1}{2}Dh \end{aligned}$$

where $Y_o = X_o - hZ$. It remains to check that $I - \alpha\alpha' \geq o$. Take $\xi \in \mathbb{R}^r$ with $|\xi| = 1$. Then

$$\begin{aligned} \xi'(I - \alpha\alpha')\xi &= 1 - (\alpha'\xi)^2 \\ &\geq 1|\alpha|^2 \\ &\geq 0. \end{aligned}$$

**82** Since $|\alpha|^2 \leq 1$, for

$$\begin{aligned} < \sum_i \alpha^i B^i, W >_t &= \sum_i \alpha_i < B^i, W >_t \\ &= |\alpha|^2 t \end{aligned}$$

So $\alpha^2 t = EUW_t$, where $U = \sum_i \alpha^i B^i$.

$$|\alpha|^4 t^2 \leq EU^2 t = |\alpha|^2 t^2.$$

So we have Zakai equation in Ito's form

$$d\sigma_t(f) = \sigma_t(Af)dt + \sigma_t(Df)dy_t$$

With vector fields $X_0, X_1, \ldots X_r$ and in Stratonovich form

$$d\sigma_t(f) = \sigma_t(Lf) + \sigma_t(Df)ody_t$$

With vector Fields $Y_0, Y_1, \ldots Y_r$ plus a "$0^{\text{th}}$-order" term. Now we investigate what sort of process is $x_t$ under the measure $P_o$.

**Proposition 1.3.** *Under $P_o, x_t$ satisfies the equation*

$$df(x_t) = Y_o f(x_t)dt + Zf(x_t)ody_t + Y_j f(x_t)odb_t^j \tag{17}$$

*where $b^1, \ldots b^r$ are independent standard Brownian motions independent of $y_t$*

*Proof.* Recall the Girsanov transformation. Under $P, B^1, \ldots B^r$ are independent and $< B^j, W >_t = \alpha^j t$. Now

$$\frac{dP_o}{dP} = \exp\left(M_t - \frac{1}{2} < M, M >_t\right)$$

where $M$ is a $P$-martingale. Under $P_o, B_t^j - < B^j, M >_t$ is a martingale **83**
and hence a Brownian motion. Here

$$M_t = -\int_o^t h(x_s)dW_s;$$
$$d < B^j, M >= -\alpha^j h(x_t)dt.$$

□

So under $P_o$,

$$dV_t^j = dB_t^j + \alpha^j h(x_t)dt$$

are independent Brownian motions, but $V^j$ is not independent of $y_t$, in fact $< V^j, y >_t = \alpha^j t$.

Now define $\tilde{b}_t^j = V_t^j - \alpha_j y_t$, then $< \tilde{b}^j, y >= 0$, and this implies $\tilde{b}^j, y$ are independent. But the $\tilde{b}^j$ are now not independent. In fact,

$$< b^j, b^k >_t = \begin{cases} -\alpha^j \alpha^k t & \text{for } k \neq j \\ 1 - (\alpha^j)^2 t & \text{for } k = j \end{cases}$$

So

$$\begin{aligned} < \tilde{b}^j, \tilde{b}^k >_t &= \left[< \tilde{b}^j, \tilde{b}^k >_t\right] \\ &= (I - \alpha\alpha')t. \end{aligned}$$

Let $(I - \alpha\alpha') = \Delta\Delta'$ as before and define $b_t = \Delta'^{-1}\tilde{b}_t$.
Then $\tilde{b}_t = \Delta'\tilde{b}_y$ and $< b >_t =$ It. So

$$\begin{aligned} df(x_t) &= X_o f(x_t)dt + X_j f(x_t) o dB^i \\ &= X_o f(x_t)dt + X_j f(x_t) o(-\alpha^j h(x_t)dt + \alpha^j dy_t + d\tilde{b}_t^j) \\ &= (X_o f(x_t) - hZf(x_t))dt + Zf(x_t) o dy_t + Y_j f(x_t) o d\tilde{b}_t^j \\ &= Y_o f(x_t)dt + Zf(x_t) o dy_t + Y_j f(x_t) o \, \mathrm{d}\, \tilde{b}_t^j \end{aligned}$$

84 where $Y_o f = X_o f - hZf$ and $Y = \Delta' X$.

The so called *Kallianpur- Striebel Formula* gives the solution $\sigma_t$ of the Zakai equation as a function space integral in the following form, where $x_t$ is functional of $(y, b^1, \ldots b^r)$.

$$\sigma_t(f) = E_o\left[f(x_t)\exp\left(\int_o^t h(x_s)dy_s - \frac{1}{2}\int_o^t h^2(x_s)ds\right) y_t\right]$$

i.e., as a function of $y$, we have

$$\sigma_t(f)(y) = \int_{C^r[o,t]}\left[f(x_t)\exp\left(\int_o^t h(x_s)dy_s - \frac{1}{2}\int_o^t h^2(x_s)ds\right)\right]\mu_w(db^1)\ldots\mu_w(db^r)$$

where $\mu_w(db)$ is the Wiener measure on $C[\circ, T]$.

# 2 Pathwise Solutions of Differential Equations

Consider the Doss- Sussman construction for the equation

$$\begin{aligned} \dot{x} &= b(x) + g(x)\dot{w} \\ x(o) &= x \end{aligned} \tag{1}$$

where $w \in C'(\mathbb{R}_+)$. Let $\phi(t, x)$ be the "flow" of $g$, i.e.,

$$\begin{aligned} \frac{\partial}{\partial^t}\phi(t, x) &= g(\phi(t, x)) \\ \phi(o, x) &= x. \end{aligned}$$

If $b = o$, then it is immediate that the solution of (1) is **85**

$$x_t = \phi(w(t), x).$$

If $b \neq o$, then the solution of (1) is of the form

$$x_t = \phi(w(t), \eta(t)) \tag{2}$$

where $\eta(t)$ satisfies some *ODE*. With $x(t)$ defined by (2),

$$\dot{x}(t) = g(x(t))\dot{w}(t) + \phi_x(w(t), \eta(t))\dot{\eta}(t)$$

and we require that

$$\phi_x(w(t), \eta(t))\dot{\eta}(t) = b(\phi(w(t), \eta(t))).$$

So $x(t)$ satisfies (1) if $\eta(t)$ satisfies

$$\begin{aligned} \dot{\eta} &= (\phi_x(w(t), \eta(t)))^{-1} b(\phi(w(t), \eta(t))) \\ \eta(o) &= x \end{aligned}$$

**Coordinate-free form:** Let

$$\begin{aligned} X_o f(x) &= b(x)\frac{df}{dx} \\ X_1 f(x) &= g(x)\frac{df}{dx} \end{aligned}$$

$$\xi_t(x) = \phi(w(t), x).$$

Define
$$(\xi_{t*}^{-1} X_o) f(x) = X_o(fo\xi_t^{-1})(\xi_t(x)).$$
Then the equation for $\eta(t)$ can be expressed as

$$\frac{d}{dt} f(\eta_t) = (\xi_{t*}^{-1} X_o) f(\eta_t)$$
$$\eta_o = x,$$

86 for

$$\begin{aligned}(\xi_{t*}^{-1} X_o) f(\eta_t) &= b(\xi_t(x)) \frac{d}{dx} f(\xi_t^{-1}(x))|_{\xi_t(x)} \\ &= b(\xi_t(x)) \frac{d}{dx} f(x) \frac{d}{dx}(\xi_t^{-1}(x))|_{\xi_t(x)} \\ &= b(\xi_t(x)) \frac{d}{dx} f(x)(\xi_x(x))^{-1}. \qquad (*)\end{aligned}$$

Since $\xi^{-1}(\xi(x)) = x$ and so

$$\frac{d}{dx}(\xi^{-1}(\xi(x))) \frac{d}{dx}\xi(x) = 1.$$

When $x \in \mathbb{R}^d$, then

$$X_o f(x) = \sum_{i=1}^{d} b^i(x) \frac{\partial f(x)}{\partial x_i}$$
$$X_1 f(x) = \sum_{i=1}^{d} g^i(x) \frac{\partial f(x)}{\partial x_i}$$

Then $(*)$ is of the form

$$(\xi_{t*}^{-1} X_o) f(\eta_t) = \sum_{i=1}^{d} b^i(x) \frac{\partial}{\partial x_i}\{f \circ \xi_t^{-1}\}|_{\xi_t(x)}$$

$$= \sum_{j=1}^{d} \sum_{i=1}^{d} \frac{\partial f(x)}{\partial x_i} \frac{\partial (\xi_t^{-1})^j}{\partial x_i}(x).$$

So the same results apply for $x_t \in \mathbb{R}^d$, but generally *not* for more than one "input", i.e., for vector $w(t)$.

**Interpretation:** $x_t$ defined by (2) makes sense for any $w(.) \in C(\mathbb{R}_+)$. In particular, if $w(t)$ is a sample path of Brownian motion, then what equation does $x_t$ satisfy? **87**
Answer: the Stratonovich equation

$$dx_t = b(x_t)dt + g(x_t)odw_t \qquad (3)$$

**Exercise 2.1.** *Expand $x_t$ given by (2) using Ito's calculus and show that it satisfies (3).*

*The following examples show that the pathwise solution idea cannot generally be extended to "multi-input" equations.*

**Example 2.1.** Let

$$\dot{x} = g^1(x)\dot{w}^1 + g^2(x)\dot{w}^2$$
$$x(o) = x.$$

The solution should be of form

$$x_t = h(w_t^1, w_t^2). \qquad (4)$$

Then with $h_1(w^1, w^2) = \dfrac{\partial}{\partial w^1} h(w^1, w^2)$ etc., we have

$$\dot{x}_t = h_1\dot{w}^1 + h_2\dot{w}^2$$

$$h_1(w_t^1, w_t^2) = g^1 oh(w_t^1, w_t^2)$$
$$h_2(w_t^1, w_t^2) = g^2 oh(w_t^1, w_t^2)$$

and

$$h_{12}(w_t^1, w_t^1) = g_x^1 oh.h_2 = (g_x^1 oh)(g^2 oh)$$

$$h_{21}(w_t^1, w_t^2) = (g_x^2 oh)(g^1 oh).$$

So we must have

$$g^1 g_x^2 = g^2 g_x^1.$$

88 Define the Lie bracket $[X_1, X_2] = X_1X_2 - X_2X_1$. Now

$$X_1X_2 f = g^1 \frac{d}{dx}\left(g^2 \frac{df}{dx}\right)$$
$$g^1 g_x^2 f_x + g^1 g^2 f_{xx}.$$

Therefore

$$[X_1, X_2] f = (g^1 g_x^2 - g^2 g_x^1) f_x.$$

So a necessary condition for (4) to hold is that

$$[X_1, X_2] = 0$$

i.e.,
$$X_1X_2 = X_2X_1.$$

**Exercise 2.2.** *Consider*

$$\dot{X} = \sum_{i=1}^{n} g^i(x)\dot{w}^i.$$

*Let $\phi^i(t, x)$ be the flow of $g^i$ and $\xi_t^i(x) = \phi^i(w_t^i, x)$. Then show that*

$$x_t = \xi_t^1 o \xi_t^2 o \dots o \xi_t^n$$

*if $[X^i, X^j] = 0\ \forall i, j$.*

With one input, $||w^n - w|| \to 0$ implies $x_t^n \to x_t$, where $\|\cdot\|$ is the sup norm. But with inputs $w^1, w^2$, the solution map generally is not continuous.

89 **Example 2.2** (Sussmann [23])**.** Let $t \in [0, 1]$ and

$$\dot{x}^n = Ax^n\dot{w}^{1,n} + Bx^n\dot{w}^{2,n}$$
$$x(o) = x_o$$

where $A, B$ are $n \times n$ matrices with

$$[A, B] = AB - BA \neq 0.$$

Partition $[0, 1]$ into $n$ equal intervals $I_j^n = \left[\frac{j-1}{n}, \frac{j}{n}\right]$, $j = 1, 2, \ldots, n$. Partition each $I_j^n$ into four equal intervals $I_{j,i}^n$, $i = 1, 2, 3, 4$. Define $w^{1,n}$ to be equal to $4n^{1/2}$ for $t \in I_{j,1}^n$ to $-4n^{1/2}$ for $I_{j,2}^n$, and to zero for all other $t$. Similarly, let $\dot{w}^{2,n}$ be equal to $4n^{1/2}$ for $t \in I_{j,2}^n$ to $-4n^{1/2}$ for $t \in I_{j,4}^n$, and to zero for all other $t$.

Then

$$w^{i,n}(t) = \int_o^t \dot{w}^{i,n}(s)ds, i. = 1, 2.$$

Clearly $\dot{w}^{i,n}$ converges to zero uniformly as $n \to \infty, i = 1, 2$. Let $s = n^{-1/2}$, then

$$\begin{aligned} x^n(1/n) &= e^{Bs}e^{-As}e^{Bs}e^{As}x_o \\ &= e^{\tau}x_o. \end{aligned}$$

We use the Baker-Campbell -Hausdorff formula $e^A e^B = e^C$ where

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}\{[[B, A], A] + [[B, A], B]\} + \cdots$$

we get

$$\tau = [B, A]\frac{1}{n} + o(1/n).$$

So **90**

$$\begin{aligned} x^n(1) &= e^{n\tau}x_o \\ &= e^{([B,A]+o(1/n))}x_o \end{aligned}$$

Hence

$$\lim_{n\to\infty} x_t^n = e^{t[B,A]}x_o.$$

# 3 Pathwise Solution of the Filter Equation

Consider the equation

$$df(x_t) = Y_o f(x_t)dt + Zf(x_t)ody_t + Y_i f(x_t)odb_t^i.$$

To express this pathwise in $(y_t)$, let $\phi(t,x)$ be the integral curve of $Z$ and $\xi_t(x) = \phi(y(t),x)$. Define $\eta_t$ as follows:

$$\begin{aligned} df(\eta_t) &= (\xi_t^{-1*} Y_o) f(\eta_t)dt + (\xi_{t*}^{-1} Y_j) f(\eta_t) \circ db_t^j \\ \eta_o &= x. \end{aligned}$$

Then

$$\begin{aligned} x_t &= \xi_t o \eta_t \\ &= \phi(y(t), \eta(t)). \end{aligned}$$

The generator of $\eta$ is

$$A_t^* = \xi_{t*}^{-1} Y_o + \frac{1}{2} \sum_j \left(\xi_{t*}^{-1} Y_j\right)^2.$$

91 **The Kallianpur-Striebel Formula:** Recall

$$\begin{aligned} \sigma_t(f) &= E^{(b)}\left[f(x_t)\exp\left(\int_o^t h(x_s)dy_s - \frac{1}{2}\int_o^t h^2(x_s)ds\right)\right] \\ &= E^{(b)}\left[f(x_t)\exp\left(\int_o^t h(x_s)ody_s - \frac{1}{2}\int_o^t Dh(x_s)ds\right)\right] \end{aligned}$$

where $D = Z + h$.

**Notation:** For any diffeomorphism $\psi : M \to M, \psi^* : C^\infty(M) \to C^\infty(M)$ is given by

$$\psi^* f(x) = f o \psi(x) = f(\psi(x)).$$

So

$$f(x_t) = \xi_t^* f(\eta_t),$$

and $$\sigma_t(f) = E^{(b)}\left[\xi_t^* f(\eta_t)\exp\left(\int_o^t \xi_s^* h(\eta_s) o dy_s - \frac{1}{2}\int_o^* Dh(\eta_s)ds\right)\right]$$

The next step is to remove "ody". Define

$$H(t,x) = \int_o^t \phi_s^* h(x)ds.$$

Calculate $H(y_t,\eta_t)$ using Stratonovich calculus

$$dH(y_t,\eta_t) = \xi_t^* h(\eta_t) o dy_t + \xi_{t*}^{-1} Y_o H y_s(\eta_s)ds + \xi_{t*}^{-1} Y_i H y_s(\eta_s) o db_s^i.$$

**Notation:**

$$\begin{aligned} g_s(x) &= H(y(s),x)\\ y_j^* &= \xi_{t*}^{-1} Y_j\\ B_s f(x) &= \phi_s^* f(x)\exp\left(\int_o^s \phi_u^* h(x)du\right). \end{aligned}$$

Finally, we get **92**

$$\sigma_t(f) = E^{(b)}[By_{(t)} f(\eta_t)\alpha_t^o(y)]$$

where the multiplicative functional $\alpha_t^s$ is given by

$$\begin{aligned} \alpha_t^s(y) = \exp\Bigg[\int_s^t Y_j^* g_u(\eta_u) db_u^j - \frac{1}{2}\int_s^t (Y_j^*)^2 g_u(\eta_u)du\\ - \int_s^t Y_o^* g_u(\eta_u)du - \frac{1}{2}\int_s^t \xi_u^* Dh(n_u)du\Bigg] \end{aligned}$$

So $\sigma_t(f)$ is now pathwise in $y$ with $\sigma_t(f): C[o,t] \to \mathbb{R}$ and $(\sigma_t(f)/\sigma_t(1))$ is a version of $E[f(x_t)|y_t]$. Now we want to compute $\sigma_t(f)$ recursively.

**(a) Multiplicative Functional Approach:**

Let $(x_t)$ be a Markov process with extended generator $(A, \mathcal{D}(A))$. The associated semigroup on $B(E)$ is

$$T_{s,t} f(x) = E_{s,x}[f(x_t)]$$

$\alpha_t^s (s \leq t)$ is a *multiplicative functional* $(m.f.)$ of $(x_t)$ if $\alpha_t^s$ is $\sigma\{x_u, s \leq u \leq t\}$-measurable and for $r \leq s \leq t$,

$$\alpha_t^r = \alpha_s^r \alpha_t^s.$$

Corresponding to $\alpha_t^s$ there is a semigroup defined by

$$T_{s,t}^{\alpha} f(x) = E_{s,x}\left[f(x_t)\alpha_t^s\right].$$

In particular,

$$T_{s,t}^{\alpha} 1 = E_{s,x}[\alpha_t^s].$$

93 It is a Markov (or Sub-Markov) semigroup when

$$E_{s,x}\ [\alpha_t^s] = 1\ (\leq 1).$$

If $(x_t)$ is a homogeneous Markov process, $\alpha_t^s$ is a homogeneous m.f. if

$$\alpha_t^s = \alpha_{t+r}^{s+r}\ o\ \theta_r$$

where $$\theta_r\ x_t = x_{t+r}.$$

Then

$$\alpha_t^s = \alpha_{t-s}^o\ o\ \theta_{-s}.$$

So denoting $\alpha_t = \alpha_t^o$, the $m.f.$ property is

$$\alpha_{t+s} = \alpha_t . \alpha_s\ o\ \theta_t.$$

Now we want to find the generator of

$$T_t^{\alpha}\ f(x) = E_x(f(x_t)\ \alpha_t).$$

Suppose for the moment that $\alpha_t \leq 1, \forall\ t$. Then $\alpha_t$ is monotone decreasing. In this cases $\alpha_t$ corresponds to "killing" at rate $(-d\alpha_t/\alpha_t)$. It is possible to construct an "$\alpha$-subprocess" which is a Markov process $x_t$ such that

$$E_x[f(x_t^{\alpha})] = T_t^{\alpha}\ f(x).$$

See Blumenthal and Getoor [1]. Define the extended generator of $T^\alpha$ to be the extended generator of $x_t^\alpha$, i.e.,

$$f(x_t^\alpha) - f(x_o^\alpha) - \int_o^t A^\alpha f(x_s^\alpha) ds$$

is a local Martingale if $f \in D(A^\alpha)$. This says (excluding stopping) 94

$$E\left[ f(x_t^\alpha) - f(x_s^\alpha) - \int_s^t A^\alpha f(x_u^\alpha) du | F_s \right] = 0$$

or
$$E_{x_s}[\alpha_{t-s}\, f(x_t)] - f(x_s) - E_{x_s} \int_s^t \alpha_{u-s} A^\alpha f(x_u) du = 0.$$

So equivalently, $f \in D(A^\alpha)$ if

$$\alpha_t \left(f(x_t) - f(x) - \int_o^t \alpha_s A^\alpha f(x_s) ds\right.$$

is a local Martingale $(P_x)$ for every $x$.

This characterizes $A^\alpha$ even when the condition $\alpha_t \leq 1$ is not satisfied, so we adopt it as our definition.

**Example 3.1.** Let $\gamma_t = \exp\left(-\int_o^t V(x_s) ds\right)$ where $V \in B(E)$. Take $f \in D(A)$ and compute

$$d(\gamma_t\, f(x_t)) = \gamma_t\, A\, f(x_t)\, dt + \gamma_t\, d\, M_t\, f - V(x_t)\, f(x_t)\, \gamma_t\, dt.$$

$$\gamma_t f(x_t) - f(x) = \int_o^t \gamma_s [A f(x_s) - V(x_s) f(x_s)]\, ds + \int_o^t \gamma_s\, dM_s f.$$

So

$$A^\gamma f(x) = A f(x) - V(x) f(x).$$

**Example 3.2.** Let $\beta_t = \dfrac{a(x_t)}{a(x_o)}$ where a $\in$ $D(A)$ and $a(x) > o\ \forall\ x$. Then 95

$$T_t^\beta\, g(x) = \frac{1}{a(x)}\, T_t(af)(x).$$

**Exercise 3.1.** *Show that*

$$A^\beta f(x) = \frac{1}{a(x)}\, A(af)(x).$$

*Now suppose $x_t$ satisfies*

$$df(x_t) = X_o f(x_t)dt + X_j\ f(x_t)\ o\ dw_t^j.$$

Take $g \in C_b^\infty(E)$ and define

$$\delta_t = \exp\left(-\int\limits_o^t X_j g(x_u) dw_u^j - \frac{1}{2}\int\limits_o^t \sum_j (X_j g(x_u))^2 du\right). \qquad (1)$$

If we define

$$\frac{dP_x^\delta}{dP_x} = \delta_t,$$

then

$$d\tilde{w}_t^j = dw_t^j + X_j\ g(x_t)dt$$

is a $P_x^\delta$ - Brownian motion. Thus

$$df(x_t) = (X_o f(x_t) - \sum_j X_j g(x_t) X_j f(x_t))dt + X_j f(x_t) o\ d\tilde{w}_t^{-j}.$$

Now $\delta_t$ is a $m.f.$ of $x_t$ (as will be verified below) and

$$E_x[f(x_t)\delta_t] = E_x^\delta[f(x_t)].$$

96 So

$$A^\delta f(x) = \left(X_o - \sum_j X_j g(x_t) X_j\right) f + \frac{1}{2}\sum_j X_j^2\ f.$$

The three examples here are related by

$$\int_o^t X_j\, g(x_s) dw_s^j = g(x_t) - g(x) - \int_o^t Ag(x_s) ds.$$

Using this in (1), we see that $\delta_t$ factors

$$\delta_t = \beta_t\, \gamma_t$$

with

$$V(x) = -Ag(x) + \frac{1}{2} \sum_j (X_j\, g(x))^2$$
$$a(x) = e^{-g(x)}.$$

So

$$A^\delta f(x) = e^g\, A(e^{-g} f) - \left(Ag - \frac{1}{2}\sum_j (X_j g)^2\right) f(x).$$

So

$$e^g A(e^{-g} f) = Af - \sum (X_j g) X_j f - \left(Ag + \frac{1}{2}\sum_j (X_j g)^2\right) f.$$

**Exercise 3.2.** *Verify that this result is correct by direct calculation of* $e^g A(e^{-g} f)$.

We have the unnormalized solution of filtering problem as

$$\sigma_t(f) = E[B_{y_t}\; f(\eta_t)\alpha_t^o(y)]$$

for $y \in C[o, T]$. Here $\eta_t$ is a diffusion, $\alpha_t^o(y)$ is a m.f. of $\eta_t$. Now

$$\alpha_t^s(y) = \exp\left(-\int_s^t Y_j^* g(\eta_u) db_u^j - \frac{1}{2}\int_s^t \sum_j (Y_j^* g_u(\eta_u))^2 du\right)$$
$$\times \exp\left(\frac{1}{2}\int_s^t \sum_j (Y_j^* g_u(\eta_u))^2 du - \frac{1}{2}\int_s^t (Y_j^*)^2 g_u(\eta_u) du\right.$$

$$-\int_s^t Y_o^* g_u(\eta_u)du - \frac{1}{2}\int_s^t \xi_u^* Dh(\eta_u)du\Bigg).$$

97 This factors $\alpha_t^s$ into product of a "Girsanov" $m.f.$ and a "Feynman-Kac" $m.f.$ Hence the corresponding generator is

$$A_t^y f = A_t^* f - \sum_j Y_j^*\, g_t Y_j^* f + \left[\frac{1}{2}\sum_j (Y_j^* g_t)^2 - A^* g_t - \frac{1}{2}\xi_t^* Dh\right] f.$$

**Proposition 3.1.** $A_t^y f = B_{y_t}\left(A - \frac{1}{2}D^2\right)B_{y_t}^{-1}$.

*Proof.* This can be verified by a straightforward but somewhat lengthy calculation, using the expansion for $e^g A e^{-g}$ obtained previously, once has obtained an expression for $B_t^{-1}$. Recall that $B_t$ is defined by

$$B_t f(x) = f(\xi(t,x)) \exp\int_o^t h(\xi(u,x))du.$$

□

It is a group of operators with generator $D = Z + h$. The inverse $B_t^{-1}$ is given as follows. Let $g(x) = B_t f(x)$, then

$$\begin{aligned} f(x) &= B_t^{-1} g(x) \\ &= g(\xi(-t,x)) \exp\left(-\int_o^t h(\xi(u,\xi(-t,x)))du\right) \\ &= g(\xi^{-1}(t,x)) \exp\left(-\int_o^t h(\xi^{-1}(s,x))ds\right). \end{aligned}$$

98 **Example 3.3** (Independent "signal" and "noise")**.** Take $Z = 0$, then $\xi(t,x) = x$ and

$$\begin{aligned} A_t^y f(x) &= e^{h(x)y(t)}\left(A - \frac{1}{2}h^2\right)(e^{-y(t)h(.)} f(.))(x) \\ &= e^{hy(t)} A e^{-hy(t)} f - \frac{1}{2}h^2 f. \end{aligned}$$

It is easy to see that this must be the right formula. The calculations have been carried out for arbitrary $y \in C[o, T]$ but $A_t^y$ depends only on $y(t)$. So $A_t^y = A_t^{\bar{y}}$ where $\bar{y}(s) \equiv y(t)$ ($t$ fixed). Now

$$\begin{aligned}
\sigma_t(f)(\bar{y}) &= E\left[f(x_t)e^{h(x_t)\bar{y}_t}\exp\left(-\int\limits_o^t \bar{y}(s)dh(x_s) - \frac{1}{2}\int\limits_o^t h^2(x_s)ds\right)\right] \\
&= E\left[f(x_t)e^{h(x_t)\bar{y}_t}\exp\left(-\bar{y}(t)h(x_s) + \bar{y}(t)h(x_o) - \frac{1}{2}\int\limits_o^t h^2(x_s)ds\right)\right] \\
&= E\left[f(x_t)e^{h(x_t)y_t}\frac{\exp\left(-y(t)h(x_t)\right)}{\exp(-y(t)h(x_o))}.\exp\left(-\frac{1}{2}\int\limits_o^t h^2(x_s)ds\right)\right].
\end{aligned}$$

So we have separated into two functionals and the result follows.

**Direct Solution of Zakai Equation:** We will consider a slight generalization from the original Zakai equation. Define

$$\begin{aligned}
L &= \frac{1}{2}\sum_j Y_j^2 + Y_o + h_o \\
D &= Z + h
\end{aligned}$$

where $Y_i, Z$ are smooth vector fields; $h, h_o$ are $C_b^\infty$ functions (Previously, **99**
we had $h_o = -\frac{1}{2}Dh$). Write $\langle f, \mu\rangle$ for $\int f\, d_\mu$ and consider the measure-valued equation

$$d\langle f, \sigma_t\rangle = \langle Lf, \sigma_t\rangle dt + \langle Df, \sigma\rangle \circ\ dy_t \tag{2}$$

where $\langle f, \sigma.\rangle = f(x)$ i.e., $\sigma_o = \delta_x$. The solution can be expressed as follows: Define $x_t$ by

$$df(x_t) = Zf(x_t).dy_t + Y_o f(x_t)dt + Y_j(x_t) \circ db_t^j, x_o = x,$$

where $b_j$ are Brownian motion independent of $y$. Then the solution is

$$\sigma_t(f) = E^b\left[f(x_t)\exp\left(\int_o^t h_o(x_s)ds + \int_o^t h(x_s) \circ dy_s\right)\right].$$

Kunita [17] show that this solution is unique if coefficients are smooth and bounded. Now the question is whether $\sigma_t$ has a density.

**Theorem 3.1.**
$$\langle f, \sigma_t \rangle = \langle B_{y_t} f, \nu_t \rangle \tag{3}$$

*where*
$$B_{y_t} f(x) = \langle f, \mu_t \rangle$$

*and $\mu_t, \nu_t$ satisfy the equations*

$$d < f, \mu_t > = < Df, \mu_t > \ ody_t \tag{4}$$
$$< f, \mu_0 > = f(x)$$
$$d < f, \nu_t > = < B_{y_t}\ L\ B_{y_t}^{-1}, \nu_t \ > dt. \tag{5}$$

100 *Proof.* If $L = o$, then (4) is the same as (2); so the solution of (4) is

$$< f, \mu_t >= f(x_t) \exp\left( \int\limits_o^t h(x_s) \circ dy_s \right)$$

where $x_t$ satisfies
$$df(x_t) = Zf(x_t) ody_t.$$

□

But this has pathwise solution $x_t = \xi(y_t, x)$. The previous definition of $B$ was
$$B_{y_t} f(x) = f(\xi(y_t, x)) \exp\left( \int\limits_o^{y_t} h(\xi(u, x)) du \right).$$

Now,
$$d\left( \int\limits_o^{y_t} h(\xi(u, x)) du \right) = h(\xi(y_t, x)) ody_t.$$

So (3) holds with $B_t$ defined as before. Now

$$d < B_{y_t} f, \nu_t > = d' < B_{y'_t} f, \nu_t > + < B_{y_t} f, \dot{\nu}_t > dt$$

$$= B_{y_t} L\ B_{y_t}^{-l}\ B_{y_t} f, \nu_t > dt + < B_{y_t} D\ f, \nu_t > ody_t$$
by (5) and (4).
$$=< B_{y_t} L\ f, \nu_t > dt + < B_{y_t}\ D\ f, \nu_t > ody_t.$$

This verifies (2).

**Proposition 3.2.** *Suppose $\nu_t$ has a density function $q_t(z, x)$ [3]. Then for* **101**
*$t > o.\sigma_t$ has density*

$$\rho_t(V) = q_t(\xi^{-1}(y_t, V)x) \exp\left(\int\limits_o^{y_t} h(\xi^{-1}(s, V))ds\right) \times \mid \frac{\partial}{\partial V} \xi^{-1}(y_t, V) \mid \quad (6)$$

*where $\mid \dfrac{d\xi^{-1}}{dV} \mid$ is the Jacobian of the map $V \to \xi^{-1}(y_t, V)$.*

*Proof.* If $\nu_t$ has a density $q_t$, then

$$\begin{aligned} < f, \sigma_t > &= \int B_{y_t} f(z) q_t(z, x) dz \\ &= \int f(\xi(y_t, z)) \exp(\int\limits_o^{y_t} h(\xi(u, z)) du) q_t(z, x) dz. \end{aligned}$$

Changing the variable to $V = \xi(y_t, z)$ gives (6). □

**Theorem 3.2** (Bismut [2])**.** *$\nu_t$ has $C^\infty$-density if the $Y_i$ are "smooth" vector fields, i.e., coefficients are bounded with bounded derivatives of all orders and $Y_1, \ldots, Y_n$ satisfy the "restricted Hörmander condition" $H$ :- Consider vector fields $Y_i$, $[Y_i, Y_j]$, $[[Y_i, Y_j], Y_k] \ldots$. At each $x$ the restrictions of these vector fields to $x$ span $T_x(M)$.*

In local coordinates $Y_i = \sum\limits_i b_i(x) \dfrac{\partial}{\partial x_i}, \ldots$ etc. So the condition says the vectors $b$ etc. span $\mathbb{R}^d$ at each $x$. Recall, $B_{y_t}\ LB_{y_t}^{-1} = A_t^* +$ (1[st] and

[3]Here $z$ is the "dummy variable" and $x$ refers of the initial condition in (2), i.e., $\sigma_0 = \delta_x$

$0^{\text{th}}$ order terms)

$$A^* = \sum_i (\xi_{t*}^{-1} Y_i)^2 + \cdots$$

Now

$$\xi_{t^*}^{-1}[Y_i, Y_j] = [\xi_{t^*}^{-1} Y_i, \xi_{t^*}^{-1} Y_j] \text{ etc .}$$

102 So if $Y_i$ satisfy the Hörmander condition, then $(\xi_{t*}^{-1} Y_i)$ satisfies it.

Hörmander's own result requires coefficients to be $C^\infty$ in $(t, x)$. Here the coefficients are continuous (but not even $C^1$) in $t$. Bismut's version of Malliavin calculus shows that the result still holds with this degree of smoothness in $t$.

In the filtering problem, the "signal process" involved vector fields $X_1, X_2, \ldots X_n, X_0$ and $Y = \Delta X$, where $\Delta$ is nonsingular if $| \alpha |< 1$. Then $X = \Delta^{-1} Y$. So

$$[\ldots[[X_{i_1}, X_{i_2}]X_{i_3}]\ldots X_{i_k}] = \sum_j c_j \left[\ldots[[Y_1\delta, Y_2\delta], Y_3\delta]\ldots Y_{in,j}\delta\right].$$

So if the "$X$" Lie brackets span $\mathbb{R}^d$ then there must be a collection of "$Y$" brackets which also span $\mathbb{R}^d$. The Hörmander condition for $X$ with $| \alpha |< 1$ implies the existence of density.

**The Case of Vector Observations:** Let $dy^i = h_i(x_t)dt + dW_t^{0,i}, W^{0,i}$ are independent Brownian motions. $\alpha$ will now be a matrix,

$$\alpha_{ij} t =< W^{0,i}, W^{0,j} > .$$

Consider the following cases. $(a)$ Independent signal and noise: Here $\alpha_{ij} = 0 \forall i, j$. Then whole theory goes through unchanged.

103 Then (2) becomes

$$d < f, \mu_t >= \sum_i < h_i f, \mu_t > ody_t^i$$

with solution

$$< f, \mu_t > = \exp\left(\sum_i y_i(t) h(x)\right) f(x)$$

$$= \prod_i \exp(y_i(t)h(x))f(x)$$

So this gives pathwise solution as before.

**Another Point of View:** The Kallianpur-Striebel formula is

$$\sigma_t(f) = E^{(x)}\left[f(x_t)\exp\left(\sum_i \int_i^t h_i(x_s)dy_s^i - \frac{1}{2}\sum_i \int_o^t h_i^2(x_s)ds\right)\right]$$

$$= E^{(x)}\left[f(x_t)\prod_i e^{y^i(t)h(x_t)}\exp\left(\sum_i \int_o^t y^i(s)dh_i(x_s)\right.\right.$$

$$= \left.\left.\frac{1}{2}\left(\sum_i \int_o^t h_i^2(x_s)ds\right)\right)\right].$$

**(b) The General Case:** Here we have no "pathwise" theory (except under very artificial conditions) but the same theory goes through a.s. (Wiener measure). There is no continuous extension to the whole of $C^p[o,T]$. In this case, equation (2) becomes

$$d < f, \mu_t >= \sum_i < D_i f, \mu_t > ody_t^i$$

where $D_i = Z_i + h$ and $Z_i$ is a vector field .

A pathwise solution only exists if $D_i's$ commute, which is very artificial. But, as before, the solution can be expressed as

$$< f, \mu_t >= f(x_t)\exp\left(\sum_i \int_0^t h_i(x_s)ody_s^i\right) \tag{7}$$

where $x_t$ satisfies **104**

$$df(x_t) = \sum_i D_i f(x_t)ody_t^i$$

$$x_o = x.$$

Regard $\mu$ as the operator mapping $f \to < f, \mu_t >$. Then "stochastic flow" theory (Elworthy [13], Kunita [17], Bismut [2] says that if $D_i' s$ have smooth coefficients then $x \to x_t(x, \omega)$ is a diffeomorphism a.a. $\omega$, and so the inverse map $x_t^{-1}(x)$ exists. We have to calculate $\mu_t^{-1}$. Generalize (7) slightly to

$$< f, \mu_{s,t} >= f(x_t, t) \exp\left(\sum_i \int_s^t h_i(x_r, r)\, o dy_r^i\right)$$

$$d f(x_t, t) = \sum_i D_i f(x_t, t)\ o\ dy_t^i\ t \geq s$$

$$x_s = x.$$

**Proposition 3.3** (Kunita [17])**.**

$$\mu_{s,t}^{-1}(f(x)) = f(x_{s,t}^{-1}(x)) \exp\left(-\sum_i \int_s^t h_i(x_{r,t}^{-1}(x)) \circ \hat{d}y_r^i\right) \tag{8}$$

*where "$o\hat{d}$" means backwards Stratonovich integral. Here define $\sigma-$fields $F_{r,t}$ $r \leq t$, by*

$$F_{r,t} = \sigma\left\{y_u^i - y_v^i,\ r \leq u, v \leq t, i = 1, 2, \ldots, d\right\}.$$

*Then $\int_s^t \phi_r \hat{d}y^i$ is a well defined backward I to integral if $\phi_r$ is a back-*
105 *ward semimartingale w.r.t. $(F_{r,t})_{r \leq t}$. Then the Stratonovich integral is defined as usual. If $\Phi_r$ is continuous, then*

$$\int \phi_r o \hat{d} y_r^i = \sum_k \Phi\left(\frac{t_k^n + t_{k+1}^n}{2}\right)\left(y_{k_{k+1}^n}^i - y_{t_k^n}^i\right)$$

*So $\mu_{s,t}^{-1}(x)$ is well defined by (8). Now verify that*

$$\int_s^t h_i\left(x_{r,t}^{-1}(p)\right) \circ \hat{d}y_r^i\Big|_{p = x_{s,t}(x)}$$

$$= \int_s^t h_i(x_{s,r}) o dy_r^i.$$

*This checks that* $\mu_{s,t}^{-1}(\mu_{s,t}(x)) = x$.

Now all remaining calculations go through as before *but only* a.s. (Wiener measure).

More general results on existence of densities have been obtained by Bismut and Michel [3]

# Bibliography

[1] R.M. BLUMENTHAL and R.K. GETOOR, Markov processes and Potential theory, Academic press, New York 1968. 106

[2] J.M. BISMUT, Martingales, The Malliavin calculus and Hypoellicity under general Hormander conditions: Z. Wahrscheinlichkeitstheorie 56 (1981) 469-505.

[3] J.M. BISMUT and D. MICHEL, Diffusions conditionelles I, II, J. Functional Analysis 44 (1981), 174-211, 45 (1982) 274-292.

[4] R. BOEL, P. VARAIYA and E. WONG, Martingales on jump processes. I Representation results, SIAM J. Control 13 (1975) 999-1021.

[5] J.M.C. CLARK, The design of robust approximations to the Stochastic differentials equations of nonlinear filtering, Communication Systems and Random Process Theory, ed. J.K. Skwirzynski, NATO Advanced Study Institute Series, Sijthoff and Noordhoff, Alphen aan den Rijn, 1978.

[6] M.H.A. DAVIS, The representation of Martingales of jump processes, SIAM J. Control and Optimization 14 (1976) 623-638.

[7] M.H.A. DAVIS, Piecewise-deterministic Markov Processes, a general class of non-diffusion Stochastic models, to appear in J. Royal Statist. Soc. (B).

[8] M.H.A. DAVIS and S.I. MARCUS, An Introduction to nonlinear filtering, Stochastic Systems: The Mathematics of Filtering

and Identification and Applications ed. M. Hazewinkel and J.C. Willems, 1981

107 [9] M.H.A. Davis, Pathwise nonlinear filtering, Stochastic Systems, The Mathematics of Filtering and Identification and applications ed. M. Hazewinkel and J.C. Willems, Reidel (1981) 505-528.

[10] M.H.A. Davis, On multiplicative functional transformation arising in nonlinear filtering theory, Z. Wahrscheinlichkeits theorie Verw. Geb. 54, (1980) 125-139.

[11] M.H.A. DAVIS, Linear Estimation and Stochastic Control, Champan & Hall, London 1977.

[12] C. DELLACHERIE and P.A. MEYER, Probablity and Potentials, North-Holland 1978.

[13] K.D.ELWORTHY, Stochastic dynamical systems and their flows, Stochastic Analysis, ed. A. Friedman and M. Pinsky, Acadenic Press, New York 1978.

[14] R.J. ELLOTT, Stochastic Calculus and Applications Springer-Verlag 1982.

[15] N. EL Karovi, Les aspectes probabilistes du controle stochastique, in Lecture Notes in Mathematics, 876, Springer-Verlag 1981

[16] B.V. GNEDENKO and I. I. KOVALENKO, Introduction to the theory of Mass Service[Russian] Moscow, 1966.

[17] H. KUNITA, Densities of a measure-valued process governed by a stochastic partial differential equation, Systems and Control letters 1 (1981) 100-104.

[18] H. KUNITA, Stochastic partial differential equations connected with nonlinear filtering, Nonlinear filtering and Stochastic Control, Lecture Notes in Mathematics 972, Springer-Verlag, Berlin 1982.

108 [19] G. KALLIANPUR, Stochastic Filtering theory, Springer-Verlag, 1980.

[20] E. PARDOUX, Equations du filtering non-linear, de la prediction et du lissage, Stochastics 6 (1982) 193-231.

[21] S. R. PLISKA, Controlled jump processes, Stochastic processes and their applications 3 (1979) 259-282.

[22] Z. ROSBERG, P.P. VARAIYA and J. C. WALRAND, Optimal Control of Service in Tandem Queues, IEEE Trans. Automatic Control 27 (1982) 600-610.

[23] H.J. SUSSMAN, On the gap between deterministic and Stochastic ordinary differential equations, The Annals of Probability 6 (1978) 19-41.

[24] D. VERMES, Optimal dynamic control of a useful class of randomly jumping processes, report pp-80-15. International Institute for Applied Systems Analysis, Laxenburg, Australia, 1980.

[25] D. VERMES, Optiomal Control of Piecewise-deterministic Processes, Preprint, Bolyai Institute, University of Szeged, Hungary 1983 (to appear in Stochastics)

[26] R.B. VINTER and R.M. LEWIS, A necessary and sufficient condition for optimality of Dynamic Programming type, Making no apriori assumption on the controls, SIAM J. Control and Optimization 16 (1978) 571-583.

[27] A.D. WENTZELL, A course in the theory of Stochastic processes, McGraw-Hill Inc. 1981.

[28] E. WONG, Recent Progress in Stochastic Processes-a survey. IEEE Trans. Inform. Theory, IT-19, 1973, 262-275.

[29] F.A.VAN DER DUYA SCHOUTON, Markov Decision Drift Pro- 109
cesses, Mathematics Centre Tract, Amsterdam 1983.

[30] A.A. YUSHKEVICH, Continuous-time Markov decision processes with interventions, Stochastics 9 (1983) 235-274.

[31] R. RISHEL, Dynamic programming and minimum principles for systems with jump Markov disturbances, SIAM J. Control 13 (1975) 338-371.