Hattendorff's theorem and Thiele's differential equation generalized

by
Reinhardt Messerschmidt

Submitted in partial fulfillment of
the requirements for the degree

Magister Scientiae

in the Department of Insurance & Actuarial
Science
in the Faculty of Natural & Agricultural Sciences

University of Pretoria
Pretoria

February 2005

Declaration

I, the undersigned, hereby declare that the dissertation submitted herewith for the degree Magister Scientiae to the University of Pretoria contains my own, independent work and has not been submitted for any degree at any other university.

Signature	
Name	Reinhardt Messerschmidt
Date	2005-02-21

Summary

Title	Hattendorff's theorem and Thiele's differential equation generalized
Candidate	Reinhardt Messerschmidt
Supervisor	Prof. Johan Swart
Department	Insurance & Actuarial Science
Degree	M.Sc. (Actuarial science)

Hattendorff's theorem on the zero means and uncorrelatedness of losses in disjoint time periods on a life insurance policy is derived for payment streams, discount functions and time periods that are all stochastic. Thiele's differential equation, describing the development of life insurance policy reserves over the contract period, is derived for stochastic payment streams generated by point processes with intensities. The development follows that in [8] and [7].

In pursuit of these aims, the basic properties of Lebesgue-Stieltjes integration are spelled out in detail. An axiomatic approach to the discounting of payment streams is presented, and a characterization in terms of the integral of a discount function is derived, following the development in [9]. The required concepts and tools from the theory of continuous time stochastic processes, in particular point processes, are surveyed.

Acknowledgements

I would like to express my gratitude to:

Prof. Johan Swart, for his willingness to supervise a topic not strictly within his area of expertise. This shows him to be a true scientist.

Sanlam, for its financial support during my university studies.

My parents and brothers, for their love.

Contents

Declaration **i**

Summary **ii**

Acknowledgements **iii**

1 Introduction **1**

2 Lebesgue-Stieltjes integration **2**

3 Discounting **16**

4 Stochastic processes **23**

5 Hattendorff's theorem **34**

6 Point processes **36**

7 Thiele's differential equation **45**

Chapter 1

Introduction

Hattendorff's theorem states that the losses in disjoint time periods on a life insurance policy have zero means and are uncorrelated, so that the variance of the total loss is the sum of the variances of the periodic losses. The original formulation by K. Hattendorff in 1868 was an approximation based on limit arguments and the normal distribution, and the first rigorous proofs were provided by J.F. Steffensen in 1929. It is a theoretically important result that brought together the theories of statistics and life contingencies (see [5], p. xxxviii and [6], chapter 10).

Thiele's differential equation describes the development of a life insurance policy reserve over the contract period. A simple version (in standard actuarial notation),

$$\frac{d}{dt}({}_tV_x) = \bar{P}_x + \delta({}_tV_x) - \mu_{x+t}(1 - {}_tV_x),$$

was shown by T.N. Thiele to colleagues in 1875, but it was only printed in his obituary by J.P. Gram in 1910. The theoretical importance of Thiele's equation arises from the insight it gives into the dynamics of a life insurance policy, and its practical importance from its use in the design of policies with reserve dependent payments (see [5], p. xxxix and [6], chapter 15).

These classical results have been revived through the application of the modern theory of stochastic processes.

Basic measure theory, as covered in chapters 1-5 of [2], and basic probability theory, as covered in chapters 1-3 of [3], are considered to be prerequisite. Propositions stated without proof are direct consequences of the relevant definitions.

Chapter 2

Lebesgue-Stieltjes integration

Functions of finite variation

Suppose $x \in \mathbb{R}^{\mathbb{R}_+}$, $s, t \in \mathbb{R}_+$, and $s \leq t$. Let

$$\operatorname{Var}(x, s, t) = \sup \left\{ \sum_{i=1}^{n} |x(t_i) - x(t_{i-1})| : n \in \mathbb{N}, s = t_0 < t_1 < \ldots < t_n = t \right\}$$

if $s < t$, and $\operatorname{Var}(x, s, t) = 0$ if $s = t$. $\operatorname{Var}(x, s, t)$ is called the *variation of* x *over* $[s, t]$.

Proposition 2.1. *(a)* $\operatorname{Var}(x + y, s, t) \leq \operatorname{Var}(x, s, t) + \operatorname{Var}(y, s, t)$.

(b) If $c \in \mathbb{R}$, *then* $\operatorname{Var}(cx, s, t) = |c| \operatorname{Var}(x, s, t)$.

(c) If $s \leq t \leq u$, *then* $\operatorname{Var}(x, s, u) = \operatorname{Var}(x, s, t) + \operatorname{Var}(x, t, u)$.

(d) If x *is nondecreasing, then* $\operatorname{Var}(x, s, t) = x(t) - x(s)$.

(e) If x *is right continuous, then*

$$\operatorname{Var}(x, s, t) = \sup \left\{ \sum_{i=1}^{2^n} |x(t_i) - x(t_{i-1})| : n \in \mathbb{N}, t_i = s + \frac{i}{2^n}(t - s) \right\}.$$

Proof. ***(e).*** The result is clear if $s = t$, so suppose $s < t$. Clearly,

$$\sup \left\{ \sum_{i=1}^{2^n} |x(t_i) - x(t_{i-1})| : n \in \mathbb{N}, t_i = s + \frac{i}{2^n}(t - s) \right\} \leq \operatorname{Var}(x, s, t).$$

Suppose $\epsilon > 0$ and $s = t_0 < t_1 < \ldots < t_n = t$. x is right continuous, therefore for every $i = 1, 2, \ldots, n$ there exists $\delta_i > 0$ such that if $u \in [t_i, t_i + \delta_i)$, then

$$|x(u) - x(t_i)| < \frac{\epsilon}{2n}.$$

Choose $n_0 \in \mathbb{N}$ such that $n_0 \geq n$ and $(1/2^{n_0})(t-s) < \min\{\delta_1, \delta_2, \ldots, \delta_n\}$, then for every $i = 1, 2, \ldots, n$ there exists $m_i \in \mathbb{N}$ such that

$$s + \frac{m_i}{2^{n_0}}(t-s) \in [t_i, t_i + \delta_i).$$

For every $j = 1, 2, \ldots, 2^{n_0}$, let

$$t'_j = s + \frac{j}{2^{n_0}}(t-s),$$

then

$$\begin{aligned}
&\sum_{i=1}^{n} |x(t_i) - x(t_{i-1})| \\
&\leq \sum_{i=1}^{n} \left|x(t_i) - x(t'_{m_i})\right| + \sum_{i=1}^{n} \left|x(t'_{m_i}) - x(t'_{m_{i-1}})\right| + \sum_{i=1}^{n} \left|x(t'_{m_{i-1}}) - x(t_{i-1})\right| \\
&\leq \sum_{j=1}^{2^{n_0}} \left|x(t'_j) - x(t'_{j-1})\right| + \epsilon \\
&\leq \sup \left\{ \sum_{i=1}^{2^n} |x(t_i) - x(t_{i-1})| : n \in \mathbb{N}, t_i = s + \frac{i}{2^n}(t-s) \right\} + \epsilon,
\end{aligned}$$

from which the opposite inequality follows. □

$x \in \mathbb{R}^{\mathbb{R}_+}$ is said to have *finite variation* if for every $t \in \mathbb{R}_+$, $\mathrm{Var}(x, 0, t) < \infty$, and to have *bounded variation* if there exists $M \in \mathbb{R}_+$ such that for every $t \in \mathbb{R}_+$, $\mathrm{Var}(x, 0, t) \leq M$. If x has finite variation, define $[x], x^{\oplus}, x^{\ominus} \in \mathbb{R}^{\mathbb{R}_+}$ by

$$[x](t) = |x(0)| + \mathrm{Var}(x, 0, t), \qquad x^{\oplus} = \frac{[x] + x}{2}, \qquad x^{\ominus} = \frac{[x] - x}{2}.$$

$x \in \mathbb{R}^{\mathbb{R}_+}$ is said to have *left limits* if for every $t > 0$, $\lim_{s \uparrow t} x(s)$ exists. If x has left limits, define $x_-, \Delta x \in \mathbb{R}^{\mathbb{R}_+}$ by

$$x_-(t) = \begin{cases} 0 & \text{if } t = 0 \\ \lim_{s \uparrow t} x(s) & \text{if } t > 0 \end{cases}, \qquad \Delta x = x - x_-.$$

Define the relation $\preccurlyeq$ on $\mathbb{R}^{\mathbb{R}_+}$ by $x \preccurlyeq y$ if $x \leq y$ and for every $s, t \in \mathbb{R}_+$ such that $s < t$, $x(t) - x(s) \leq y(t) - y(s)$.

If $x \in \mathbb{R}^{\mathbb{R}_+}$ and $t \in \mathbb{R}_+$, define $x^{(t)} \in \mathbb{R}^{\mathbb{R}_+}$ by $x^{(t)}(s) = x(s \wedge t)$.

Proposition 2.2. *Suppose x is a function of finite variation.*

(a) $x = x^{\oplus} - x^{\ominus}$ *and* $[x] = x^{\oplus} + x^{\ominus}$.

(b) $[x], x^{\oplus}, x^{\ominus}$ *are nonnegative and nondecreasing.*

(c) x *has left limits, and* x_- *is left continuous.*

(d) If x *is right continuous, then so are* $[x], x^{\oplus}, x^{\ominus}$.

(e) If x *is nonnegative and nondecreasing, then* $x = [x] = x^{\oplus}$ *and* $x^{\ominus} = 0$.

(f) $[x + y] \preccurlyeq [x] + [y]$.

(g) If $a \in \mathbb{R}$*, then* $[ax] = |a|[x]$.

(h) $[x^{(t)}] = [x]^{(t)} \preccurlyeq [x]$.

Proof. (c). $x^{\oplus}, x^{\ominus}$ are nondecreasing functions, and therefore have left limits, so $x = x^{\oplus} - x^{\ominus}$ has left limits.

For the rest, it is sufficient (by parts (a) and (b)) to show that if x is non-negative and nondecreasing, then x_- is left continuous.

For every $m, n \in \mathbb{N}$, let

$$A_{m,n} = \{s \in [0, m] : (\Delta x)(s) > 1/n\},$$
$$A = \{s \in \mathbb{R}_+ : (\Delta x)(s) > 0\}.$$

$\frac{1}{n} |A_{m,n}| \leq x(m)$, i.e. $A_{m,n}$ is a finite set, therefore $A = \cup_{m=1}^{\infty} \cup_{n=1}^{\infty} A_{m,n}$ is at most countable, so there exists a strictly increasing sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_+$ such that $A \subseteq \{t_n : n \in \mathbb{N}\}$.

For every $n \in \mathbb{N}$, let

$$x_n = \sum_{i=1}^{n} (\Delta x)(t_i) 1([t_i, \infty)), \quad x_d = \sum_{i=1}^{\infty} (\Delta x)(t_i) 1([t_i, \infty)), \quad x_c = x - x_d.$$

These functions are all nonnegative and nondecreasing. Now

$$(x_c)_- = x_c + \Delta(x_c) = x_c + \Delta x - \Delta(x_d) = x_c,$$

i.e. x_c is left-continuous. Furthermore, for every $n \in \mathbb{N}$,

$$(x_n)_- = \sum_{i=1}^{n} (\Delta x)(t_i) 1((t_i, \infty)), \quad (x_d)_- = \sum_{i=1}^{\infty} (\Delta x)(t_i) 1((t_i, \infty)),$$

and for every $t \in \mathbb{R}_+$, $((x_n)_-)$ converges uniformly on $[0, t]$ to $(x_d)_-$, therefore $(x_d)_-$ is left continuous. It follows that $x_- = (x_c)_- + (x_d)_-$ is left continuous. □

(d). Suppose $t \in \mathbb{R}_+$ and $\epsilon > 0$. Because x is right continuous, there exists $\delta' \in (0, 1)$ such that if $s \in [t, t + \delta')$, then $|x(s) - x(t)| < \epsilon/2$. There also exist $t = t'_0 < t'_1 < \ldots < t'_n = t + 1$ such that

$$\mathrm{Var}(x, t, t + 1) - \epsilon/2 < \sum_{i=1}^{n} \left| x(t'_i) - x(t'_{i-1}) \right|.$$

Add $t + \delta'/2$ to this partition to obtain a new partition $t = t_0 < t_1 < \ldots < t_m = t + 1$, then $t < t_1 \leq t + \delta'/2 < t + \delta'$, therefore

$$\begin{aligned}
\mathrm{Var}(x, t, t+1) - \epsilon/2 &< \sum_{i=1}^{n} \left|x(t_i') - x(t_{i-1}')\right| \\
&\leq \sum_{i=1}^{m} |x(t_i) - x(t_{i-1})| \\
&= |x(t_1) - x(t)| + \sum_{i=2}^{m} |x(t_i) - x(t_{i-1})| \\
&< \epsilon/2 + \mathrm{Var}(x, t_1, t+1),
\end{aligned}$$

i.e. $\mathrm{Var}(x, t, t_1) < \epsilon$. Let $\delta = t_1 - t$ and suppose $s \in [t, t + \delta)$, then

$$|[x](s) - [x](t)| = \mathrm{Var}(x, t, s) \leq \mathrm{Var}(x, t, t_1) < \epsilon,$$

i.e. $[x]$ is right continuous. It follows that $x^{\oplus}, x^{\ominus}$ are also right continuous. □

$x \in \mathbb{R}^{\mathbb{R}_+}$ is called:

(i) an *integrator function* if it has finite variation and is right continuous.

(ii) a *distribution function* if it is nonnegative, nondecreasing, and right continuous. Let $\mathbb{D}$ be the set of all distribution functions.

(iii) a *signed distribution function* if it has bounded variation and is right continuous. Let $\mathbb{D}_s$ be the set of all signed distribution functions.

Proposition 2.3. *(a) If x is a distribution function or a signed distribution function, then x is also an integrator function.*

(b) If x is an integrator function, then $[x], x^{\oplus}, x^{\ominus}$ are distribution functions.

(c) If x is a signed distribution function, then $[x], x^{\oplus}, x^{\ominus}$ are bounded distribution functions.

(d) If x is an integrator function, then $x^{(t)}$ is a signed distribution function.

Distribution functions and measures

Let $\mathbb{M}$ be the set of all measures μ on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ such that for every $t \in \mathbb{R}_+$, $\mu([0, t]) < \infty$, and let $\mathbb{M}_s$ be the set of all finite signed measures on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$.

Suppose x is a distribution function. Define the extension x' of x to $\mathbb{R}$ by $x'(t) = 0$ if $t < 0$, and note that x' is nondecreasing and right continuous. Define $\mu^* \in \bar{\mathbb{R}}_+^{(2^{\mathbb{R}})}$ by

$$\mu^*(A) = \inf\left\{\sum_{i=1}^{\infty}(x'(t_i) - x'(s_i)) : s_i, t_i \in \mathbb{R}, s_i < t_i, A \subseteq \bigcup_{i=1}^{\infty}(s_i, t_i]\right\}.$$

μ^* is an outer measure. Let μ be its restriction to $\mathcal{B}(\mathbb{R}_+)$, then μ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}_+))$ such that for every $t \in \mathbb{R}_+$, $\mu([0,t]) = x(t)$, i.e. $\mu \in \mathbb{M}$. Furthermore, if $\nu \in \mathbb{M}$ is such that for every $t \in \mathbb{R}_+$, $\nu([0,t]) = x(t)$, then $\mu = \nu$.

These remarks allow one to define $\phi \in \mathbb{M}^{\mathbb{D}}$ by letting $\phi(x)$ be the element μ of $\mathbb{M}$ such that for every $t \in \mathbb{R}_+$, $\mu([0,t]) = x(t)$.

If $\mu \in \mathbb{M}$ ($\mu \in \mathbb{M}_s$) and $t \in \mathbb{R}_+$, define $\mu^{(t)} \in \bar{\mathbb{R}}_+{}^{\mathcal{B}(\mathbb{R}_+)}$ ($\mu^{(t)} \in \mathbb{R}^{\mathcal{B}(\mathbb{R}_+)}$) by $\mu^{(t)}(A) = \mu(A \cap [0,t])$. Note that $\mu^{(t)} \in \mathbb{M}$ ($\mu^{(t)} \in \mathbb{M}_s$).

Proposition 2.4. *(a) If $s, t \in \mathbb{R}_+$ and $s \leq t$, then $\phi(x)([s,t]) = x(t) - x_-(s)$.*

(b) ϕ is a bijection.

(c) $\phi(x+y) = \phi(x) + \phi(y)$.

(d) If $a \in \mathbb{R}_+$, then $\phi(ax) = a\phi(x)$.

(e) $\phi(x) \leq \phi(y)$ if and only if $x \preccurlyeq y$.

(f) $\phi(x^{(t)}) = \phi(x)^{(t)}$.

(g) $\phi(x)$ is a finite measure if and only if x is bounded.

Proof. ***(e).*** Suppose $\phi(x) \leq \phi(y)$. For every $s, t \in \mathbb{R}_+$ such that $s < t$,

$$x(t) - x(s) = \phi(x)((s,t]) \leq \phi(y)((s,t]) = y(t) - y(s).$$

Similarly, $x \leq y$, i.e. $x \preccurlyeq y$.

Suppose $x \preccurlyeq y$. Define the extensions x', y' of x, y to $\mathbb{R}$ by $x'(t) = y'(t) = 0$ if $t < 0$. Because $x \preccurlyeq y$, $y'(t) - y'(s) \leq x'(t) - x'(s)$ for every $s, t \in \mathbb{R}$ such that $s < t$. It follows that for every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\begin{aligned}\phi(x)(A) &= \inf\left\{\sum_{i=1}^{\infty}(x'(t_i) - x'(s_i)) : s_i, t_i \in \mathbb{R}, s_i < t_i, A \subseteq \bigcup_{i=1}^{\infty}(s_i, t_i]\right\}\\ &\leq \inf\left\{\sum_{i=1}^{\infty}(y'(t_i) - y'(s_i)) : s_i, t_i \in \mathbb{R}, s_i < t_i, A \subseteq \bigcup_{i=1}^{\infty}(s_i, t_i]\right\}\\ &= \phi(y)(A).\end{aligned}$$

□

Propositions 2.3(c) and 2.4(g) allow one to define $\psi \in \mathbb{M}_s{}^{\mathbb{D}_s}$ by $\psi(x) = \phi(x^{\oplus}) - \phi(x^{\ominus})$.

Proposition 2.5. *(a) If $s, t \in \mathbb{R}_+$ and $s \leq t$, then $\psi(x)([s,t]) = x(t) - x_-(s)$.*

(b) $|\psi(x)| = \phi([x])$, $\psi(x)^+ = \phi(x^{\oplus})$ and $\psi(x)^- = \phi(x^{\ominus})$.

(c) $\psi(x^{(t)}) = \psi(x)^{(t)}$.

Proof. (b). For every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$|\psi(x)(A)| \leq \phi(x^{\oplus})(A) + \phi(x^{\ominus})(A) = \phi([x])(A),$$

therefore $|\psi(x)| \leq \phi([x])$.

For every $s, t \in \mathbb{R}_+$ such that $s < t$,

$$\begin{aligned}
&[x](t) - [x](s) \\
&= \sup \left\{ \sum_{i=1}^{n} |x(t_i) - x(t_{i-1})| : s = t_0 < t_1 < \ldots < t_n = t \right\} \\
&= \sup \left\{ \sum_{i=1}^{n} |\psi(x)((t_{i-1}, t_i])| : s = t_0 < t_1 < \ldots < t_n = t \right\} \\
&\leq \sup \left\{ \sum_{i=1}^{n} |\psi(x)(A_i)| : A_1, A_2, \ldots, A_n \text{ disjoint}, (s,t] = \bigcup_{i=1}^{n} A_i \right\} \\
&= |\psi(x)|((s,t]) \\
&= |\psi(x)|([0,t]) - |\psi(x)|([0,s]) \\
&= \phi^{-1}(|\psi(x)|)(t) - \phi^{-1}(|\psi(x)|)(s).
\end{aligned}$$

Similarly, $[x] \leq \phi^{-1}(|\psi(x)|)$, i.e. $[x] \preccurlyeq \phi^{-1}(|\psi(x)|)$, therefore $\phi([x]) \leq |\psi(x)|$.

It follows that

$$\psi(x)^{+} = \frac{|\psi(x)| + \psi(x)}{2} = \frac{\phi([x]) + \phi(x^{\oplus}) - \phi(x^{\ominus})}{2} = \phi(x^{\oplus}).$$

Similarly, $\psi(x)^{-} = \phi(x^{\ominus})$. □

Lebesgue-Stieltjes integration

Suppose y is a distribution function. If $x \in \mathbb{R}^{\mathbb{R}_+}$ is measurable and nonnegative, let

$$\int x\, dy = \int x\, d\phi(y).$$

Suppose y is an integrator function. $x \in \mathbb{R}^{\mathbb{R}_+}$ is called *y-integrable* if x is simultaneously $\phi(y^{\oplus})$-integrable and $\phi(y^{\ominus})$-integrable. If x is y-integrable, let

$$\int x\, dy = \int x\, d\phi(y^{\oplus}) - \int x\, d\phi(y^{\ominus}).$$

Because $\phi([y]) = \phi(y^{\oplus}) + \phi(y^{\ominus})$, x is y-integrable if and only if x is $\phi([y])$-integrable. If y is a distribution function (signed distribution function), then x is y-integrable if and only if it is $\phi(y)$-integrable ($\psi(y)$-integrable), and if one of these conditions holds then

$$\int x\, dy = \int x\, d\phi(y) \qquad \left(\int x\, d\psi(y) \right).$$

Proposition 2.6. *(a) Suppose x is an integrator function, $s,t \in \mathbb{R}_+$ and $s \le t$. Then $1([s,t])$ is x-integrable and*

$$\int 1([s,t])\,dx = x(t) - x_-(s).$$

(b) If x, y are z-integrable, then $x+y$ is z-integrable and

$$\int (x+y)\,dz = \int x\,dz + \int y\,dz.$$

(c) If x is y-integrable and $a \in \mathbb{R}$, then ax is y-integrable and

$$\int ax\,dy = a\int x\,dy.$$

(d) Suppose z is y-integrable, and $x, x_1, \ldots$ are measurable functions such that $\{(x_n) \longrightarrow x\}$ and $\{|x_n| \le z\}$ are $\phi([y])$-conegligible. Then $x, x_1, \ldots$ are y-integrable, and

$$\left(\int x_n\,dy\right) \longrightarrow \int x\,dy.$$

(e) If x is y, z-integrable, then x is $(y+z)$-integrable and

$$\int x\,d(y+z) = \int x\,dy + \int x\,dz.$$

(f) If x is y-integrable and $a \in \mathbb{R}$, then x is (ay)-integrable and

$$\int x\,d(ay) = a\int x\,dy.$$

Proof. ***(e).*** $[y+z] \preccurlyeq [y]+[z]$, therefore $\phi([y+z]) \le \phi([y]) + \phi([z])$, so x is $(y+z)$-integrable.

For the rest, firstly note that for every $n \in \mathbb{N}$, $x1([0,n])$ is y,z-integrable. It is sufficient (by part (d)) to show that

$$\int_{[0,n]} x\,d(y+z) = \int_{[0,n]} x\,dy + \int_{[0,n]} x\,dz.$$

In turn, it is sufficient (by standard arguments) to show that for every $A \in \mathcal{B}(\mathbb{R}_+)$, $1(A)1([0,n])$ is y,z-integrable and

$$\int_{[0,n]} 1(A)\,d(y+z) = \int_{[0,n]} 1(A)\,dy + \int_{[0,n]} 1(A)\,dz.$$

Now $1(A)1([0,n])$ is measurable and

$$\int |1(A)1([0,n])|\,d[y] \le [y](n) < \infty,$$

i.e. $1(A)1([0,n])$ is y-integrable. Similarly, it is also z-integrable.

Let $\mathcal{H}_\pi = \{[0,s] : s \in \mathbb{R}_+\}$ and

$$\mathcal{H}_d = \left\{ A \in \mathcal{B}(\mathbb{R}_+) : \int_{[0,n]} 1(A)\, d(y+z) = \int_{[0,n]} 1(A)\, dy + \int_{[0,n]} 1(A)\, dz \right\}.$$

$\mathcal{H}_\pi \subseteq \mathcal{H}_d$, $\mathcal{H}_\pi$ is a π-system, and $\mathcal{H}_d$ is a d-system, therefore $\mathcal{B}(\mathbb{R}_+) = \sigma(\mathcal{H}_\pi) \subseteq \mathcal{H}_d$. □

(f). $\phi([ay]) = |a|\phi([y])$, so x is (ay)-integrable.

Suppose $a \geq 0$, then $(ay)^\oplus = ay^\oplus$ and $(ay)^\ominus = ay^\ominus$, therefore

$$\begin{aligned}\int x\, d(ay) &= \int x\, d(ay^\oplus) - \int x\, d(ay^\ominus)\\ &= a\int x\, dy.\end{aligned}$$

Suppose $a < 0$, then $(ay)^\oplus = (-a)y^\ominus$ and $(ay)^\ominus = (-a)y^\oplus$, therefore

$$\begin{aligned}\int x\, d(ay) &= \int x\, d((-a)y^\ominus) - \int x\, d((-a)y^\oplus)\\ &= (-a)\left(-\int x\, dy\right)\\ &= a\int xdy.\end{aligned}$$

□

The operation ”·”

Suppose y is an integrator function. $x \in \mathbb{R}^{\mathbb{R}_+}$ is called *locally y-integrable* if for every $t \in \mathbb{R}_+$, $x1([0,t])$ is y-integrable. If x is locally y-integrable, define $x \cdot y \in \mathbb{R}^{\mathbb{R}_+}$ by

$$(x\cdot y)(t) = \int_{[0,t]} x\, dy.$$

In the order of operations, ”·” ranks above addition but below multiplication, i.e. $x+y\cdot z = x+(y\cdot z)$ and $xy\cdot z = (xy)\cdot z$. Note that $(x\cdot y)^{(t)} = x1([0,t])\cdot y = x\cdot y^{(t)}$.

Proposition 2.7. *Suppose y is a distribution function and x is nonnegative and locally y-integrable.*

(a) $x \cdot y$ is a distribution function.

(b) For every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\phi(x\cdot y)(A) = \int_A x\, dy,$$

i.e. x is a Radon-Nikodym derivative of $\phi(x \cdot y)$ with respect to $\phi(y)$.

(c) $\Delta(x \cdot y) = x\Delta y$.

Proof. (a). $x \cdot y$ is clearly nonnegative and nondecreasing.

Suppose $t \in \mathbb{R}_+$ and $\epsilon > 0$. For every $n \in \mathbb{N}$, let $t_n = t + 1/n$, then $(x1([0,t_n])) \longrightarrow x1([0,t])$, therefore $((x \cdot y)(t_n)) \longrightarrow (x \cdot y)(t)$, so there exists $n_0 \in \mathbb{N}$ such that

$$|(x \cdot y)(t) - (x \cdot y)(t_{n_0})| = (x \cdot y)(t_{n_0}) - (x \cdot y)(t) < \epsilon.$$

Let $\delta = 1/n_0$ and suppose $s \in [t, t+\delta)$, then

$$|(x \cdot y)(t) - (x \cdot y)(s)| = (x \cdot y)(s) - (x \cdot y)(t) \leq (x \cdot y)(t_{n_0}) - (x \cdot y)(t) < \epsilon. \quad \square$$

(b). It is sufficient (by the monotone convergence theorem) to show that for every $n \in \mathbb{N}$,

$$\phi(x \cdot y)(A \cap [0,n]) = \int_{A \cap [0,n]} x \, dy.$$

Let $\mathcal{H}_\pi = \{[0,s] : s \in \mathbb{R}_+\}$ and

$$\mathcal{H}_d = \left\{ A \in \mathcal{B}(\mathbb{R}_+) : \phi(x \cdot y)(A \cap [0,n]) = \int_{A \cap [0,n]} x \, dy \right\}$$

$\mathcal{H}_\pi \subseteq \mathcal{H}_d$, $\mathcal{H}_\pi$ is a π-system, and $\mathcal{H}_d$ is a d-system, therefore $\mathcal{B}(\mathbb{R}_+) = \sigma(\mathcal{H}_\pi) \subseteq \mathcal{H}_d$. $\square$

(c). Suppose $t \in \mathbb{R}_+$ and $\epsilon > 0$. For every $n \in \mathbb{N}$, let $t_n = t - 1/n$, then $(x1([0,t_n])) \longrightarrow x1([0,t))$, therefore

$$((x \cdot y)(t_n)) \longrightarrow \int_{[0,t)} x \, dy,$$

so there exists $n_0 \in \mathbb{N}$ such that

$$\left| \int_{[0,t)} x \, dy - (x \cdot y)(t_{n_0}) \right| = \int_{[0,t)} x \, dy - (x \cdot y)(t_{n_0}) < \epsilon.$$

Let $\delta = 1/n_0$, and suppose $s \in (t-\delta, t)$, then

$$\left| \int_{[0,t)} x \, dy - (x \cdot y)(s) \right| = \int_{[0,t)} x \, dy - (x \cdot y)(s) \leq \int_{[0,t)} x \, dy - (x \cdot y)(t_{n_0}) < \epsilon.$$

It follows that for every $t \in \mathbb{R}_+$,

$$(x \cdot y)_-(t) = \int_{[0,t)} x \, dy,$$

$$\Delta(x \cdot y)(t) = \int_{[t,t]} x \, dy = x(t)(\Delta y)(t). \quad \square$$

Proposition 2.8. *Suppose y is a distribution function and x is y-integrable.*

(a) $x \cdot y$ is a signed distribution function.

(b) For every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\psi(x \cdot y)(A) = \int_A x\,dy,$$

i.e. x is a Radon-Nikodym derivative of $\psi(x \cdot y)$ with respect to $\phi(y)$.

(c) $[x \cdot y] = |x| \cdot y$.

Proof. ***(a).*** Suppose $t \in \mathbb{R}_+$ and $0 = t_0 < t_1 < \ldots < t_n = t$, then

$$\begin{aligned}
\sum_{i=1}^{n} |(x \cdot y)(t_i) - (x \cdot y)(t_{i-1})| &= \sum_{i=1}^{n} \left| \int_{(t_{i-1},t_i]} x\,dy^{\oplus} - \int_{(t_{i-1},t_i]} x\,dy^{\ominus} \right| \\
&\leq \sum_{i=1}^{n} \left(\int_{(t_{i-1},t_i]} |x|\,dy^{\oplus} + \int_{(t_{i-1},t_i]} |x|\,dy^{\ominus} \right) \\
&= \int_{(0,t]} |x|\,d[y] \\
&\leq \int |x|\,d[y],
\end{aligned}$$

i.e. $(x \cdot y)$ has bounded variation.

x^+, x^- are y-integrable and $x = x^+ - x^-$, therefore

$$x \cdot y = x^+ \cdot y - x^- \cdot y,$$

so $x \cdot y$ is right continuous (proposition 2.7(a)). □

(b). Let $\mathcal{H}_\pi = \{[0, s] : s \in \mathbb{R}_+\}$ and

$$\mathcal{H}_d = \left\{ A \in \mathcal{B}(\mathbb{R}_+) : \psi(x \cdot y)(A) = \int_A x\,dy \right\}.$$

$\mathcal{H}_\pi \subseteq \mathcal{H}_d$, $\mathcal{H}_\pi$ is a π-system, and $\mathcal{H}_d$ is a d-system, therefore $\mathcal{B}(\mathbb{R}_+) = \sigma(\mathcal{H}_\pi) \subseteq \mathcal{H}_d$. □

(c). For every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\phi([x \cdot y])(A) = |\psi(x \cdot y)|(A) = \int_A |x|\,dy \qquad \text{(see [2], proposition 4.2.4)},$$

therefore for every $t \in \mathbb{R}_+$,

$$[x \cdot y](t) = \phi([x \cdot y])([0, t])) = \int_{[0,t]} |x|\,dy = (|x| \cdot y)(t).$$ □

Suppose x is an integrator function and y is a distribution function. x is called *absolutely continuous with respect to* y if $\phi([x])$ is absolutely continuous with respect to $\phi(y)$. $z \in \mathbb{R}^{\mathbb{R}_+}$ is called a *Radon-Nikodym derivative of* x *with respect to* y if z is locally y-integrable and $x = z \cdot y$.

Proposition 2.9. *Suppose x is an integrator function and y is a distribution function.*

(a) If x is absolutely continuous with respect to y, then there exists a Radon-Nikodym derivative of x with respect to y.

(b) If w, z are Radon-Nikodym derivatives of x with respect to y, then $\{w = z\}$ is $\phi(y)$-conegligible.

(c) If z is a Radon-Nikodym derivative of x with respect to y, and w is locally x-integrable, then wz is locally y-integrable and $w \cdot x = wz \cdot y$.

Proof. (a). For every $t \in \mathbb{R}_+$, $x^{(t)}$ is a signed distribution function. $[x^{(t)}] \preccurlyeq [x]$, therefore

$$|\psi(x^{(t)})| = \phi([x^{(t)}]) \le \phi([x]),$$

so $\psi(x^{(t)})$ is absolutely continuous with respect to $\phi(y)$. Let z_t be a Radon-Nikodym derivative of $\psi(x^{(t)})$ with respect to $\phi(y)$.

Suppose $s \le t$, then

$$\psi(x^{(s)}) = \psi((x^{(t)})^{(s)}) = \psi(x^{(t)})^{(s)}.$$

For every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\psi(x^{(s)})(A) = \psi(x^{(t)})^{(s)}(A) = \int_{A \cap [0,s]} z_t \, dy = \int_A z_t 1([0,s]) \, dy,$$

therefore $\{z_t 1([0,s]) = z_s\}$ is $\phi(y)$-conegligible. For every $m, n \in \mathbb{N}$ such that $m \le n$, let

$$A_{m,n} = \{z_n 1([0,m]) = z_m\}.$$

Note that $\cap_{m=1}^{\infty} \cap_{n=m}^{\infty} A_{m,n}$ is $\phi(y)$-conegligible. Suppose $t \in \cap_{m=1}^{\infty} \cap_{n=m}^{\infty} A_{m,n}$ and choose $n_0 \ge t$, then for every $n \ge n_0$,

$$z_n(t) = z_n(t) 1([0, n_0])(t) = z_{n_0}(t),$$

i.e. $(z_n(t)) \longrightarrow z_{n_0}(t)$. It follows that $\cap_{m=1}^{\infty} \cap_{n=m}^{\infty} A_{m,n} \subseteq \{(z_n) \text{ converges}\}$.

Define $z \in \mathbb{R}^{\mathbb{R}_+}$ by

$$z(t) = \begin{cases} \lim_{n\to\infty} z_n(t) & \text{if } t \in \{(z_n) \text{ converges}\} \\ 0 & \text{if } t \notin \{(z_n) \text{ converges}\}. \end{cases}$$

Because $\{(z_n) \text{ converges}\} \in \mathcal{B}(\mathbb{R}_+)$, z is measurable.

Suppose $t \in \mathbb{R}_+$ and $s \in \cap_{m=1}^{\infty} \cap_{n=m}^{\infty} A_{m,n}$. Choose $n_0 \ge t$, then

$$z(s) 1([0,t])(s) = z_{n_0}(s) 1([0,t])(s) = z_t(s),$$

therefore $\{z1([0,t]) = z_t\}$ is $\phi(y)$-conegligible. It follows that

$$\int |z1([0,t])|\,dy = \int |z_t|\,dy < \infty,$$

i.e. z is locally y-integrable.

Furthermore, for every $t \in \mathbb{R}_+$,

$$(z \cdot y)(t) = (z \cdot y)^{(t)}(t) = (z1([0,t]) \cdot y)(t) = (z_t \cdot y)(t) = x^{(t)}(t) = x(t). \quad \square$$

(b). For every $t \in \mathbb{R}_+$, $\psi((w \cdot y)^{(t)}) = \psi((z \cdot y)^{(t)})$, therefore for every $A \in \mathcal{B}(\mathbb{R}_+)$,

$$\int_A w1([0,t])\,dy = \int_A z1([0,t])\,dy,$$

so $\{w1([0,t]) = z1([0,t])\}$ is $\phi(y)$-conegligible. Let

$$A_n = \{w1([0,n]) = z1([0,n])\}.$$

Note that $\cap_{n=1}^{\infty} A_n$ is $\phi(y)$-conegligible. Suppose $s \in \cap_{n=1}^{\infty} A_n$ and choose $n_0 \geq s$, then

$$w(s) = w(s)1([0,n_0])(s) = z(s)1([0,n_0])(s) = z(s),$$

therefore $\{w = z\}$ is $\phi(y)$-conegligible. $\square$

(c). For every $t \in \mathbb{R}_+$, $x^{(t)}$ is a signed distribution function and

$$x^{(t)} = (z \cdot y)^{(t)} = z1([0,t]) \cdot y,$$

therefore $z1([0,t])$ is a Radon-Nikodym derivative of $\psi(x^{(t)})$ with respect to $\phi(y)$ (proposition 2.8(b)). Because w is locally x-integrable, w is $\phi([x^{(t)}]) = |\psi(x^{(t)})|$-integrable, therefore $wz1([0,t])$ is $\phi(y)$-integrable, i.e. wz is locally y-integrable.

Furthermore, for every $t \in \mathbb{R}_+$,

$$(w \cdot x)(t) = \int w\,dx^{(t)} = \int wz1([0,t])\,dy = (wz \cdot y)(t). \quad \square$$

Proposition 2.10. *Suppose y is an integrator function and x is locally y-integrable.*

(a) $x \cdot y$ *is an integrator function.*

(b) $[x \cdot y] = |x| \cdot [y]$.

(c) $\Delta(x \cdot y) = x\Delta y$.

(d) z *is* $x \cdot y$*-integrable if and only if* xz *is* y*-integrable. If one of these conditions holds, then*

$$\int z\,d(x \cdot y) = \int zx\,dy.$$

Proof. ***(a).*** Suppose $t \in \mathbb{R}_+$ and $0 = t_0 < t_1 < \ldots < t_n = t$, then

$$\sum_{i=1}^{n} |(x \cdot y)(t_i) - (x \cdot y)(t_{i-1})| \leq \int_{(0,t]} |x| \, d[y],$$

i.e. $x \cdot y$ has finite variation.

x^+, x^- are locally y-integrable and $x = x^+ - x^-$, therefore

$$x \cdot y = x^+ \cdot y^{\oplus} - x^- \cdot y^{\oplus} - x^+ \cdot y^{\ominus} + x^- \cdot y^{\ominus},$$

so $x \cdot y$ is right continuous (proposition 2.7(a)). □

(b). Clearly, y is absolutely continuous with respect to $[y]$. Let z be a Radon-Nikodym derivative of y with respect to $[y]$. Because x is locally y-integrable, xz is locally $[y]$-integrable and $x \cdot y = xz \cdot [y]$ (proposition 2.9(c)).

For every $t \in \mathbb{R}_+$, $y^{(t)}$ is a signed distribution function and

$$\psi(y^{(t)}) = \psi((z \cdot [y])^{(t)}) = \psi(z \cdot [y^{(t)}]),$$

therefore z is a Radon-Nikodym derivative of $\psi(y^{(t)})$ with respect to

$$\phi([y^{(t)}]) = |\psi(y^{(t)})|$$

(proposition 2.8(b)). It follows that $\{|z| = 1\}$ is $\phi([y^{(t)}])$-conegligible (see [2], corollary 4.2.5), and because x is $y^{(t)}$-integrable, xz is $[y^{(t)}]$-integrable. For every $t \in \mathbb{R}_+$,

$$[x \cdot y](t) = [xz \cdot [y]]^{(t)}(t) = \left[xz \cdot [y^{(t)}]\right](t) = \left(|xz| \cdot [y^{(t)}]\right)(t) = (|x| \cdot [y])(t)$$

(proposition 2.8(c)). □

(c). x^+, x^- are locally y-integrable and $x = x^+ - x^-$, therefore

$$\begin{aligned}\Delta(x \cdot y) &= \Delta(x^+ \cdot y^{\oplus} - x^- \cdot y^{\oplus} - x^+ \cdot y^{\ominus} + x^- \cdot y^{\ominus}) \\ &= x^+\Delta(y^{\oplus}) - x^-\Delta(y^{\oplus}) - x^+\Delta(y^{\ominus}) + x^-\Delta(y^{\ominus}) \\ &= x\Delta y.\end{aligned}$$

(proposition 2.7(c)). □

(d). Suppose $z \in \mathbb{R}^{\mathbb{R}_+}$ is measurable, then

$$\int |z| \, d[x \cdot y] = \int |z| \, d(|x| \cdot [y]) = \int |zx| \, d[y],$$

(part (b) and proposition 2.7(b)), i.e. z is $x \cdot y$-integrable if and only if zx is y-integrable.

For the rest, firstly note that x^+, x^- are locally y-integrable, $x = x^+ - x^-$ and $|x| = x^+ + x^-$. Now

$$(x \cdot y)^{\oplus} = \frac{|x| \cdot [y] + x \cdot y}{2} = \frac{x^+ \cdot ([y] + y) + x^- \cdot ([y] - y)}{2} = x^+ \cdot y^{\oplus} + x^- \cdot y^{\ominus}.$$

Similarly, $(x \cdot y)^{\ominus} = x^{+} \cdot y^{\ominus} + x^{-} \cdot y^{\oplus}$. Note that z is $x^{+} \cdot y^{\oplus}$, $x^{-} \cdot y^{\ominus}$, $x^{+} \cdot y^{\ominus}$, $x^{-} \cdot y^{\oplus}$-integrable, therefore

$$\begin{aligned}\int z\, d(x \cdot y) &= \int z\, d(x^{+} \cdot y^{\oplus} + x^{-} \cdot y^{\ominus}) - \int z\, d(x^{+} \cdot y^{\ominus} + x^{-} \cdot y^{\oplus}) \\ &= \int zx^{+}\, dy^{\oplus} + \int zx^{-}\, dy^{\ominus} - \int zx^{+}\, dy^{\ominus} - \int zx^{-}\, dy^{\oplus} \\ &= \int xz\, dy.\end{aligned}$$ □

In particular, part (d) implies that z is locally $x\cdot y$-integrable if and only if zx is locally y-integrable, and if one of these conditions holds then $z \cdot (x \cdot y) = zx \cdot y$.

Proposition 2.11 (**Integration by parts**). *Suppose x, y are integrator functions. Then x is locally y-integrable, y_{-} is locally x-integrable, and $xy = x \cdot y + y_{-} \cdot x$.*

Proof. Because $x^{\oplus}$ is nondecreasing and right continuous, it is measurable. For every $t \in \mathbb{R}_{+}$,

$$\int |x^{\oplus} 1([0,t])|\, d[y] \leq x^{\oplus}(t) \int 1([0,t])\, d[y] = x^{\oplus}(t)[y](t) < \infty,$$

i.e. $x^{\oplus}$ is locally y-integrable. Similarly $x^{\ominus}$ is locally y-integrable, therefore $x = x^{\oplus} - x^{\ominus}$ is locally y-integrable. By applying the same argument to $y_{-} = (y^{\oplus})_{-} - (y^{\ominus})_{-}$ it follows that y_{-} is locally x-integrable.

Suppose x, y are distribution functions and $t \in \mathbb{R}_{+}$. Let

$$A = [0,t] \times [0,t], \quad B = \{(s,u) \in A : s \leq u\}, \quad C = \{(s,u) \in A : s > u\}.$$

By Fubini's theorem,

$$\iint_A d(\phi(x) \times \phi(y)) = x(t)y(t),$$

$$\iint_B d(\phi(x) \times \phi(y)) = \int_{[0,t]} \left(\int_{[0,u]} dx(s) \right) dy(u) = \int_{[0,t]} x\, dy = (x \cdot y)(t),$$

$$\iint_C d(\phi(x) \times \phi(y)) = \int_{[0,t]} \left(\int_{[0,s)} dy(u) \right) dx(s) = \int_{[0,t]} y_{-}\, dx = (y_{-} \cdot x)(t).$$

Because B, C are disjoint and $A = B \cup C$, it follows that $xy = x \cdot y + y_{-} \cdot x$.

For the case where x, y are arbitrary integrator functions, first decompose xy into

$$xy = x^{\oplus}y^{\oplus} - x^{\ominus}y^{\oplus} - x^{\oplus}y^{\ominus} + x^{\oplus}y^{\oplus},$$

then apply the result for distribution functions, and finally recombine. □

Chapter 3

Discounting

Valuation functions

A common economic problem is to assign a single amount at a particular time to a stream of payments flowing from one participant to another. This process is called *discounting*, and is of particular importance to a life insurance company due to the long term nature of its business. It is influenced primarily by the company's investment policy.

Distribution functions are convenient descriptions of payment streams. If P is a distribution function, then $P(t)$ represents the total amount that has been paid up to time t.

Note that if P is a distribution function, then $\lim_{t\to\infty} P(t)$ exists in $\bar{\mathbb{R}}_+$. Let $P(\infty) = \lim_{t\to\infty} P(t)$.

If $t \in \mathbb{R}_+$, let $u_t = 1([t, \infty))$. u_t represents the payment stream consisting of a single payment of 1 made at time t.

$W \in \bar{\mathbb{R}}_+^{\mathbb{R}_+ \times \mathbb{D}}$ is called a *valuation function* if it satisfies the following axioms:

Axiom 1. $W(t, 0) = 0$.

Axiom 2. (a) If $P \leq Q$, then $W(t, P) \leq W(t, Q)$.

(b) If $P \leq Q$ and $P(\infty) < Q(\infty)$, then $W(t, P) < W(t, Q)$.

Axiom 3. If (P_n) is a sequence of distribution functions such that $\sum_{i=1}^{\infty} P_i$ is a distribution function, then

$$W\left(t, \sum_{i=1}^{\infty} P_i\right) = \sum_{i=1}^{\infty} W(t, P_i).$$

Axiom 4. If P is bounded, then $W(t, P) < \infty$.

Axiom 5. If $W(t, P) < \infty$, then for every $s \in \mathbb{R}_+$, $W(s, P) = W(s, W(t, P)u_t)$.

Axiom 6. For every distribution function P, $W(\cdot, P)$ is continuous.

$W(t, P)$ represents the amount assigned at time t to the payment stream represented by P. Let $\mathbb{W}$ be the set of all valuation functions.

Axiom 1 is a kind of "no arbitrage" assumption. Axiom 2 ensures that more is preferred to less, e.g. $W(0, u_1) < W(0, 2u_1)$, and that amounts received earlier are preferred to amounts received later, e.g. $W(0, u_2) \leq W(0, u_1)$. An interpretation of axioms 3 and 5 is that nothing is gained or lost by subdividing or cashing payment streams. Axiom 6 is arguably the least self-evident of the axioms.

Proposition 3.1. *(a) If $P(\infty) > 0$, then $W(t, P) > 0$.*

(b) If $P_1, P_2, \ldots, P_n$ are distribution functions, then $\sum_{i=1}^n P_i$ is a distribution function and

$$W\left(t, \sum_{i=1}^{n} P_i\right) = \sum_{i=1}^{n} W(t, P_i).$$

(c) If P is a distribution function, $a \in \mathbb{R}_+$, and $W(t, P) < \infty$, then aP is a distribution function and $W(t, aP) = aW(t, P)$.

(d) $0 < W(s, u_t) < \infty$ and

$$W(s, u_t) = \frac{W(0, u_t)}{W(0, u_s)}$$

Proof. Part (a) follows from axioms 1 and 2(b).

Part (b) follows from axioms 1 and 3.

Part (c) is clear if $a = 0$ or $P = 0$, so it may be assumed that $a > 0$ and $P(\infty) > 0$, therefore $W(t, P) > 0$ (part (a)). Suppose $a \in \mathbb{N}$, then by part (b),

$$W(t, aP) = W(t, \sum_{i=1}^{a} P) = \sum_{i=1}^{a} W(t, P) = aW(t, P).$$

Suppose there exists $b \in \mathbb{N}$ such that $a = 1/b$, then by part (b),

$$W(t, P) = W(t, \sum_{i=1}^{b} \frac{1}{b} P) = \sum_{i=1}^{b} W(t, aP) = bW(t, aP),$$

i.e. $W(t, aP) = aW(t, P)$. It follows that for every $a \in \mathbb{Q}$, $W(t, aP) = aW(t, P)$. Suppose there exists $a_0 \in \mathbb{R}_+$ such that $W(t, a_0 P) < a_0 W(t, P)$. Choose $c \in \mathbb{Q}$ such that

$$\frac{W(t, a_0 P)}{W(t, P)} < c < a_0.$$

then $W(t, a_0 P) < cW(t, P) = W(t, cP)$, contradicting $W(t, cP) \leq W(t, a_0 P)$ (axiom 2). If there exists $a_0 \in \mathbb{R}_+$ such that $W(t, a_0 P) > a_0 W(t, P)$, then a contradiction can be derived in a similar way.

Part (d) follows from axioms 1, 4, 5 and part (c). $\square$

Discount functions

$v \in (0, \infty)^{\mathbb{R}_+}$ is called a *discount function* if it is nonincreasing, continuous, and $v(0) = 1$. Let $\mathbb{V}$ be the set of all discount functions.

Note that if v is a discount function, then $\lim_{t\to\infty} v(t)$ exists in $[0, 1]$. Let $v(\infty) = \lim_{t\to\infty} v(t)$.

If v is a discount function, let $v' \in \mathbb{R}^{\mathbb{R}_+}$ be defined by $v'(t) = 1 - v(t)$. Note that v' is a continuous distribution function.

Proposition 3.2. *If P is a distribution function and v is a P-integrable discount function, then*

$$\int v\,dP = \int P\,dv' + \int v(\infty)\,dP.$$

Proof. By integration by parts and the fact that v' is continuous,

$$v'P = P \cdot v' + v' \cdot P,$$

therefore for every $n \in \mathbb{N}$,

$$\int_{[0,n]} v\,dP = \int_{[0,n]} P\,dv' + v(n)P(n) = \int_{[0,n]} P\,dv' + \int_{[0,n]} v(n)\,dP.$$

Now $|v(n)1([0, n])| \leq v$, v is P-integrable, and $(v(n)1([0, n])) \longrightarrow v(\infty)1(\mathbb{R}_+)$, therefore by Lebesgue's dominated convergence theorem,

$$\left(\int_{[0,n]} v(n)\,dP\right) \longrightarrow \int v(\infty)\,dP.$$

The result follows by applying the monotone convergence theorem to the other two integrals. □

Characterization

Define $\theta \in \left(\bar{\mathbb{R}}_+^{\mathbb{R}_+\times\mathbb{D}}\right)^{\mathbb{V}}$ by

$$\theta(v)(t, P) = \frac{1}{v(t)} \int v\,dP.$$

Proposition 3.3. *$\theta(v)$ is a valuation function.*

Proof. Only the verification of axiom 2 is nontrivial.

Suppose $P \leq Q$. For every $n \in \mathbb{N}$, the functions $v, P^{(n)}, Q^{(n)}$ are bounded, therefore v is $P^{(n)}, Q^{(n)}$-integrable. It follows that

$$\begin{aligned}\int v\, dP^{(n)} &= \int P^{(n)}\, dv' + \int v(\infty)\, dP^{(n)} \\ &= \int P^{(n)}\, dv' + v(\infty)P(n) \\ &\leq \int Q^{(n)}\, dv' + v(\infty)Q(n) \\ &= \int v\, dQ^{(n)},\end{aligned}$$

therefore

$$\int v\, dP = \lim_{n\to\infty} \int v\, dP^{(n)} \leq \lim_{n\to\infty} \int v\, dQ^{(n)} = \int v\, dQ.$$

Suppose $P \leq Q$ and $P(\infty) < Q(\infty)$. There exists $t_0 \in \mathbb{R}_+$ such that $P \leq Q^{(t_0)}$ and $P(\infty) < Q^{(t_0)}(\infty)$. $v, P, Q^{(t_0)}$ are bounded, therefore v is $P, Q^{(t_0)}$-integrable.

Suppose $v(\infty) > 0$, then

$$\begin{aligned}\int v\, dP &= \int P\, dv' + \int v(\infty)\, dP \\ &= \int P\, dv' + v(\infty)P(\infty) \\ &< \int Q^{(t_0)}\, dv' + v(\infty)Q^{(t_0)}(\infty) \\ &= \int v\, dQ^{(t_0)} \\ &\leq \int v\, dQ.\end{aligned}$$

Suppose $v(\infty) = 0$. $P, Q^{(t_0)}$ are v'-integrable, nonnegative, and

$$\int P\, dv' \leq \int Q^{(t_0)}\, dv'.$$

Suppose that

$$\int P\, dv' = \int Q^{(t_0)}\, dv',$$

then $\left\{P = Q^{(t_0)}\right\}$ is $\phi(v')$-conegligible. Now $P(\infty) < Q^{(t_0)}(\infty)$, therefore there exists $s_0 \in \mathbb{R}_+$ such that for every $t > s_0$, $P(t) \leq P(\infty) < Q^{(t_0)}(t)$, i.e.

$$(s_0, \infty) \subseteq \mathbb{R}_+ \setminus \left\{P = Q^{(t_0)}\right\};$$

but

$$\phi(v')((s_0, \infty)) = v'(\infty) - v'(s_0) = v(s_0) > 0,$$

which is a contradiction. It follows that

$$\int P\,dv' < \int Q^{(t_0)}\,dv',$$

therefore

$$\int v\,dP = \int P\,dv' < \int Q^{(t_0)}\,dv' = \int v\,dQ^{(t_0)} \le \int v\,dQ. \qquad \square$$

θ may therefore be regarded as an element of $\mathbb{W}^{\mathbb{V}}$.

Proposition 3.4. *θ is a bijection.*

Proof. Suppose $\theta(v) = \theta(w)$, then for every $t \in \mathbb{R}_+$,

$$v(t) = \frac{1}{v(0)} \int v\,du_t = \theta(v)(0, u_t) = \theta(w)(0, u_t) = w(t),$$

i.e. θ is injective.

Suppose W is a valuation function. Define $v \in \bar{\mathbb{R}}_+^{\mathbb{R}_+}$ by $v(t) = W(0, u_t)$. It must first be verified that v is a discount function.

By proposition 3.1(d), $v \in (0,\infty)^{\mathbb{R}_+}$. Suppose $s \le t$, then $u_t \le u_s$, therefore $v(t) \le v(s)$ (axiom 2), i.e. v is nonincreasing. For every $t \in \mathbb{R}_+$,

$$W(t, u_0) = \frac{W(0, u_0)}{W(0, u_t)}$$

(proposition 3.1(d)), therefore

$$v(0) = \frac{W(0, u_0)}{W(0, u_0)} = 1,$$

and

$$v(t) = \frac{W(0, u_0)}{W(t, u_0)} = \frac{1}{W(t, u_0)},$$

so v is continuous (axiom 6).

It remains to be shown that for every $(t, P) \in \mathbb{R}_+ \times \mathbb{D}$,

$$W(t, P) = \frac{1}{v(t)} \int v\,dP.$$

Firstly, suppose that there exist sequences $(a_n)_{n\in\mathbb{N}}$ and $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_+$ such that

$$P = \sum_{i=1}^{\infty} a_i u_{t_i},$$

then

$$\begin{aligned}
W(t,P) &= \sum_{i=1}^{\infty} a_i W(t, u_{t_i}) \qquad \text{(axiom 3, proposition 3.1(c))} \\
&= \frac{1}{W(0,u_t)} \sum_{i=1}^{\infty} a_i W(0, u_{t_i}) \qquad \text{(proposition 3.1(d))} \\
&= \frac{1}{v(t)} \int v \, dP.
\end{aligned}$$

Now suppose that P is bounded. For every $n \in \mathbb{N}$, define $\bar{v}_n, \bar{P}_n \in {\mathbb{R}_+}^{\mathbb{R}_+}$ by

$$\begin{aligned}
\bar{v}_n(s) &= \begin{cases} v(0) & \text{if } s = 0 \\ v\left(\frac{\lceil 2^n s \rceil - 1}{2^n}\right) & \text{if } s > 0 \end{cases} \\
&= v(0)1(0)(s) + \sum_{i=0}^{\infty} v\left(\frac{i}{2^n}\right) 1\left(\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right]\right)(s), \\
\bar{P}_n(s) &= P\left(\frac{\lfloor 2^n s \rfloor + 1}{2^n}\right) \\
&= \sum_{i=0}^{\infty} P\left(\frac{i+1}{2^n}\right) 1\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right)(s) \\
&= P\left(\frac{1}{2^n}\right) + \sum_{i=1}^{\infty} \left[P\left(\frac{i+1}{2^n}\right) - P\left(\frac{i}{2^n}\right)\right] 1\left(\left[\frac{i}{2^n}, \infty\right)\right)(s).
\end{aligned}$$

Note that

$$\begin{aligned}
\int v \, d\bar{P}_n &= v(0)P\left(\frac{1}{2^n}\right) + \sum_{i=1}^{\infty} v\left(\frac{i}{2^n}\right)\left[P\left(\frac{i+1}{2^n}\right) - P\left(\frac{i}{2^n}\right)\right] \\
&= v(0)P(0) + \sum_{i=0}^{\infty} v\left(\frac{i}{2^n}\right)\left[P\left(\frac{i+1}{2^n}\right) - P\left(\frac{i}{2^n}\right)\right] \\
&= \int \bar{v}_n \, dP.
\end{aligned}$$

Now $|\bar{v}_n| \leq 1$, and 1 is P-integrable. Because $\left(\frac{\lceil 2^n s \rceil - 1}{2^n}\right) \uparrow s$, it follows that $(\bar{v}_n) \downarrow v$, therefore by Lebesgue's dominated convergence theorem

$$\left(\int \bar{v}_n \, dP\right) \longrightarrow \int v \, dP.$$

$s \leq \frac{\lfloor 2^n s \rfloor + 1}{2^n}$, therefore $P(s) \leq \bar{P}_n(s)$. By axiom 2,

$$W(t,P) \leq W(t, \bar{P}_n) = \frac{1}{v(t)} \int v \, d\bar{P}_n = \frac{1}{v(t)} \int \bar{v}_n \, dP \longrightarrow \frac{1}{v(t)} \int v \, dP.$$

The opposite inequality can be derived in a similar way.

Finally, suppose that P is an arbitrary distribution function. Let $P_1 = P^{(1)}$ and for every $n \in \mathbb{N}$ such that $n \geq 2$, let $P_n = P^{(n)} - P^{(n-1)}$. (P_n) is a sequence of bounded distribution functions such that $P = \sum_{i=1}^{\infty} P_i$, therefore by axiom 3,

$$W(t,P) = \sum_{i=1}^{\infty} W(t,P_i) = \sum_{i=1}^{\infty} \frac{1}{v(t)} \int v\,dP_i = \frac{1}{v(t)} \int v\,dP. \qquad \square$$

Proposition 3.4 establishes that assigning amounts to payment streams in accordance with the axioms given at the start of this chapter is equivalent to integrating the associated discount function.

Chapter 4

Stochastic processes

Suppose a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given. In what follows, the technicalities arising from dealing with $\mathbb{P}$-negligible sets will be largely ignored.

Filtrations

A collection $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ of sub-σ-algebras of $\mathcal{F}$ is called a *filtration* if for every $s, t \in \mathbb{R}_+$ such that $s \leq t$, $\mathcal{F}_s \subseteq \mathcal{F}_t$. Suppose $(\mathcal{F}_t)$ is a filtration. Let

$$\mathcal{F}_\infty = \sigma\left(\bigcup_{t\in\mathbb{R}_+} \mathcal{F}_t\right).$$

$(\mathcal{F}_t)$ is said to be *right continuous* if for every $t \in \mathbb{R}_+$,

$$\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s,$$

and is said to be *complete* if $\mathcal{F}_0$ contains all sets $A \in \mathcal{F}$ such that $\mathbb{P}[A] = 0$. If $(\mathcal{F}_t)$ is both right continuous and complete, then it is said to satisfy the *usual assumptions*.

For the remainder of this chapter, suppose a filtration $(\mathcal{F}_t)$ satisfying the usual assumptions is given. Most of the definitions in this chapter depend on the particular filtration used — this will not be mentioned explicitly on each occasion.

Stochastic processes

An element of $\mathbb{R}^{\mathbb{R}_+\times\Omega}$ is called a *process*. Suppose X is a process. For every $t \in \mathbb{R}_+$, $X(t, \cdot)$ is denoted by $X(t)$. X is said to be *adapted* if for every $t \in \mathbb{R}_+$, $X(t)$ is $\mathcal{F}_t$-measurable, and to be *deterministic* if for every $t \in \mathbb{R}_+$, $X(t)$ is constant. X is called a *nondecreasing* (*left continuous*, *right continuous*, *continuous*, *finite* *variation, distribution, integrator, discount) process* if for almost every $\omega \in \Omega$, $X(\cdot, \omega)$ is a nondecreasing (left-continuous, right continuous, continuous, finite variation, distribution, integrator, discount) function.

If X, Y are processes, then they are said to be:

(i) *indistinguishable* if for almost every $\omega \in \Omega$, $X(\cdot, \omega) = Y(\cdot, \omega)$.

(ii) *modifications* of each other if for every $t \in \mathbb{R}_+$, $X(t) = Y(t)$ almost surely.

Note that if two processes are indistinguishable, then they are modifications of each other.

Stopping times

An element of $\bar{\mathbb{R}}_+{}^\Omega$ is called a *random time*. A random time T is called *finite* if $T < \infty$.

If S, T are random times, let

$$(S, T] = \{(t, \omega) \in \mathbb{R}_+ \times \Omega : S(\omega) < t \leq T(\omega)\}.$$

$[S, T], [S, T), (S, T)$ are defined similarly.

Suppose X is a process and T is a random time. Define $X^{(T)} \in \mathbb{R}^{\mathbb{R}_+ \times \Omega}$ by

$$X^{(T)}(t, \omega) = X(t \wedge T(\omega), \omega).$$

If T is finite, or $X \in \mathbb{R}^{\bar{\mathbb{R}}_+ \times \Omega}$, define $X(T) \in \mathbb{R}^\Omega$ by

$$X(T)(\omega) = X(T(\omega), \omega).$$

A random time T is called a *stopping time* if for every $t \in \mathbb{R}_+$, $\{T \leq t\} \in \mathcal{F}_t$. If T is a stopping time, let

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \text{For every } t \in \mathbb{R}_+,\ A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

If $A \in \mathcal{F}_T$, define the random time T_A by

$$T_A(\omega) = \begin{cases} T(\omega) & \text{if } \omega \in A \\ \infty & \text{if } \omega \in \Omega \setminus A. \end{cases}$$

Proposition 4.1. *(a) A random time T is a stopping time if and only if for every $t \in \mathbb{R}_+$, $\{T < t\} \in \mathcal{F}_t$.*

(b) Every constant random time is a stopping time.

(c) If S, T are stopping times, then $S \wedge T$ is also a stopping time.

(d) $\mathcal{F}_T$ is a σ-algebra, and T is $\mathcal{F}_T$-measurable.

(e) T_A is a stopping time.

(f) If T is a constant stopping time, equal to $t_0 \in \mathbb{R}$, then $\mathcal{F}_T = \mathcal{F}_{t_0}$.

(g) If $S \leq T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.

(h) Suppose X is left continuous or right continuous, and adapted, and $a \in \mathbb{R}$. If the random time T is defined by

$$T(\omega) = \inf\{t \in \mathbb{R}_+ : X(t,\omega) > a\}$$

(with the infimum taken in $\bar{\mathbb{R}}$), then T is a stopping time.

Proof. (a). Suppose T is a stopping time, then for every $t \in \mathbb{R}_+$,

$$\{T < t\} = \bigcup_{n=1}^{\infty} \{T \leq t - 1/n\} \in \mathcal{F}_t.$$

Suppose that for every $t \in \mathbb{R}_+$, $\{T < t\} \in \mathcal{F}_t$, then for every $m \in \mathbb{N}$,

$$\bigcap_{n=1}^{\infty} \{T < t + 1/n\} = \bigcap_{n=m}^{\infty} \{T < t + 1/n\} \in \mathcal{F}_{t+1/m},$$

therefore

$$\{T \leq t\} = \bigcap_{n=1}^{\infty} \{T < t + 1/n\} \in \bigcap_{m=1}^{\infty} \mathcal{F}_{t+1/m} = \mathcal{F}_t. \qquad \square$$

(h). For every $t \in \mathbb{R}_+$,

$$\{T < t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t)} \{X(s) > a\} \in \mathcal{F}_t. \qquad \square$$

σ-algebras on $\mathbb{R}_+ \times \Omega$ associated with $(\mathcal{F}_t)$

Let $\mathcal{P}$ be the σ-algebra on $\mathbb{R}_+ \times \Omega$ generated by the left continuous adapted processes. $\mathcal{P}$ is called the *predictable σ-algebra*, and a process that is $\mathcal{P}$-measurable is said to be *predictable*. Let

$$\mathcal{P}_1 = \{(s,t] \times A : s,t \in \mathbb{R}_+, s \leq t, A \in \mathcal{F}_s\} \cup \{\{0\} \times A : A \in \mathcal{F}_0\}$$
$$\mathcal{P}_2 = \{[0,T] : T \text{ is a stopping time}\} \cup \{\{0\} \times A : A \in \mathcal{F}_0\}.$$

Note that $\mathcal{P}_1$ is a π-system.

Proposition 4.2. *(a) $\mathcal{P} = \sigma(\mathcal{P}_1) = \sigma(\mathcal{P}_2)$.*

(b) Every deterministic process is predictable.

(c) Suppose S, T are stopping times such that $S \leq T$. If X is a $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_S$-measurable process, then $X1((S,T])$ is predictable.

(d) *If $A \in \mathcal{F}_0$ and X is a $\mathcal{F}_0$-measurable random variable, then $X1(\{0\} \times A)$ is predictable.*

Proof. ***(a).*** For every $A \in \mathcal{P}_2$, $1(A)$ is a left continuous adapted process, therefore $A \in \mathcal{P}$, so $\sigma(\mathcal{P}_2) \subseteq \mathcal{P}$.

Suppose X is a left continuous adapted process. For every $n \in \mathbb{N}$, define $X_n \in \mathbb{R}^{\mathbb{R}_+ \times \Omega}$ by

$$X_n(s,\omega) = \begin{cases} X(0,\omega) & \text{if } s = 0 \\ X\left(\frac{\lceil 2^n s \rceil - 1}{2^n}\right) & \text{if } s > 0 \end{cases}$$
$$= X(0,\omega)1(0)(s) + \sum_{i=0}^{\infty} X\left(\frac{i}{2^n},\omega\right) 1\left(\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right]\right)(s).$$

For every $B \in \mathcal{B}(\mathbb{R})$,

$$\{X_n \in B\} = \Big[\{0\} \times \{X(0) \in B\}\Big] \cup \bigcup_{i=1}^{\infty} \left[\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right] \times \left\{X\left(\frac{i}{2^n},\omega\right) \in B\right\}\right],$$

therefore X_n is $\sigma(\mathcal{P}_1)$-measurable. But $\left(\frac{\lceil 2^n s \rceil - 1}{2^n}\right) \uparrow s$, therefore $X_n \longrightarrow X$, so X is $\sigma(\mathcal{P}_1)$-measurable. It follows that $\mathcal{P} \subseteq \sigma(\mathcal{P}_1)$.

Suppose $s, t \in \mathbb{R}_+, s \leq t$ and $A \in \mathcal{F}_s$, then by proposition 4.1(b),(e) and (f),

$$1((s,t] \times A) = 1((s_A, t_A]) = 1([0,t_A]) - 1([0,s_A]),$$

therefore $\sigma(\mathcal{P}_1) \subseteq \sigma(\mathcal{P}_2)$. □

(b). It is sufficient (by standard arguments) to show that for every $B \in \mathcal{B}(\mathbb{R}_+)$, $1(B)1(\Omega) = 1(B \times \Omega)$ is predictable. Let $\mathcal{H}_\pi = \{(s,t] : s,t \in \mathbb{R}_+, s \leq t\} \cup \{0\}$ and let $\mathcal{H}_d = \{A \in \mathcal{B}(\mathbb{R}_+) : 1(A \times \Omega) \text{ is predictable}\}$. $\mathcal{H}_\pi \subseteq \mathcal{H}_d$, $\mathcal{H}_\pi$ is a π-system, and $\mathcal{H}_d$ is a d-system, therefore $\mathcal{B}(\mathbb{R}_+) = \sigma(\mathcal{H}_\pi) \subseteq \mathcal{H}_d$. □

(c). It is sufficient (by standard arguments) to show that for every $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_S$, $1(A)1((S,T])$ is predictable. Let

$$\mathcal{H}_\pi = \{B \times C : B \in \mathcal{B}(\mathbb{R}_+), C \in \mathcal{F}_S\},$$
$$\mathcal{H}_d = \{A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_S : 1(A)1((S,T]) \text{ is predictable}\}.$$

For every $B \in \mathcal{B}(\mathbb{R}_+)$ and $C \in \mathcal{F}_S$,

$$1(B \times C)1((S,T]) = 1(B \times \Omega)1((S_C,T_C]) = 1(B \times \Omega)\Big(1([0,T_C]) - 1([0,S_C])\Big)$$

is predictable, so $\mathcal{H}_\pi \subseteq \mathcal{H}_d$. $\mathcal{H}_\pi$ is a π-system and $\mathcal{H}_d$ is a d-system, therefore $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_S = \sigma(\mathcal{H}_\pi) \subseteq \mathcal{H}_d$. □

(d). It is sufficient (by standard arguments) to show that for every $B \in \mathcal{F}_0$, $1(B)1(\{0\} \times A)$ is predictable. This follows immediately from

$$1(B)1(\{0\} \times A) = 1(\{0\} \times (A \cap B)).$$ □

In particular, proposition 4.2(c) implies that if $s, t \in \mathbb{R},\ s \leq t,\ A \in \mathcal{F}_s$ and X is a $\mathcal{F}_s$-measurable random variable, then $X1((s,t] \times A) = X1((s_A, t_A])$ is predictable.

Let $\mathcal{O}$ be the σ-algebra on $\mathbb{R}_+ \times \Omega$ generated by the right continuous adapted processes. $\mathcal{O}$ is called the *optional σ-algebra*, and a process that is $\mathcal{O}$-measurable is said to be *optional.* Let

$$\mathcal{O}_1 = \{[s,t) \times A : s, t \in \mathbb{R}_+, s \leq t, A \in \mathcal{F}_s\}$$
$$\mathcal{O}_2 = \{[0,T) : T \text{ is a stopping time}\}.$$

Proposition 4.3. *(a)* $\mathcal{O} = \sigma(\mathcal{O}_1) = \sigma(\mathcal{O}_2)$.

(b) $\mathcal{P} \subseteq \mathcal{O}$.

Proof. (a). Similar to the proof of proposition 4.2(a). □

(b). Suppose T is a stopping time, then

$$[0,T] = \bigcap_{n=1}^{\infty} [0, T + 1/n) \in \mathcal{O}.$$

Suppose $A \in \mathcal{F}_0$. For every $n \in \mathbb{N}$, define the random time T_n by

$$T_n(\omega) = \begin{cases} 1/n & \text{if } \omega \in A \\ 0 & \text{if } \omega \in \Omega \setminus A, \end{cases}$$

then T_n is a stopping time, and

$$\{0\} \times A = \bigcap_{n=1}^{\infty} [0, T_n) \in \mathcal{O}.$$

It follows that $\mathcal{P} = \sigma(\mathcal{P}_2) \subseteq \mathcal{O}$. □

A process is called *progressive* if for every $t \in \mathbb{R}$, its restriction to $[0,t] \times \Omega$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable. Let $\mathcal{Q}$ be the set of all subsets A of $\mathbb{R}_+ \times \Omega$ such that $1(A)$ is a progressive process. $\mathcal{Q}$ is easily verified to be a σ-algebra, called the *progressive σ-algebra.* Note that a process is progressive if and only if it is $\mathcal{Q}$-measurable.

Proposition 4.4. *(a) If a process is left continuous or right continuous, and adapted, then it is progressive.*

(b) $\mathcal{O} \subseteq \mathcal{Q} \subseteq \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty$.

(c) Every progressive process is adapted.

(d) If X is a progressive process and T is a finite stopping time, then $X(T)$ is $\mathcal{F}_T$-measurable.

Proof. (a). Suppose X is a left continuous adapted process and $t \in \mathbb{R}_+$. For every $n \in \mathbb{N}$, define X_n as in the proof of proposition 4.2(a). For every $B \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}&\{X_n\restriction_{[0,t]\times\Omega} \in B\}\\ &= \Big[\{0\} \times \{X(0) \in B\}\Big] \cup \bigcup_{i=1}^{\infty} \left[\left(\frac{i}{2^n} \wedge t, \frac{i+1}{2^n} \wedge t\right] \times \left\{X\left(\frac{i}{2^n} \wedge t, \omega\right) \in B\right\}\right],\end{aligned}$$

therefore $X_n\restriction_{[0,t]\times\Omega}$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable. But $\left(\frac{\lceil 2^n s\rceil - 1}{2^n}\right) \uparrow s$, therefore $X_n\restriction_{[0,t]\times\Omega} \longrightarrow X\restriction_{[0,t]\times\Omega}$, so $X\restriction_{[0,t]\times\Omega}$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable.

Similarly, if X is right continuous and adapted, then it is progressive. □

(b). It follows by part (a) that $\mathcal{O} \subseteq \mathcal{Q}$.

Suppose X is progressive, then for every $B \in \mathcal{B}(\mathbb{R})$,

$$\{X \in B\} = \bigcup_{n\in\mathbb{N}} \{X\restriction_{[0,n]\times\Omega} \in B\}.$$

For every $n \in \mathbb{N}$, $\{X\restriction_{[0,n]\times\Omega} \in B\} \in \mathcal{B}([0,n]) \otimes \mathcal{F}_n \subseteq \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty$, therefore X is $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty$-measurable. □

(c). Suppose X is progressive and $t \in \mathbb{R}_+$. $X\restriction_{[0,t]\times\Omega}$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable, therefore $X(t) = X\restriction_{[0,t]\times\Omega}(t)$ is $\mathcal{F}_t$-measurable. □

(d). It must first be shown that for every $t \in \mathbb{R}_+$, $X(T \wedge t)$ is $\mathcal{F}_t$-measurable. Define $F \in ([0,t] \times \Omega)^\Omega$ by $F(\omega) = (T(\omega) \wedge t, \omega)$. Because T is a stopping time, F is $\mathcal{F}_t/\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable. Now $X(T \wedge t) = X\restriction_{[0,t]\times\Omega} \circ F$, therefore $X(T \wedge t)$ is $\mathcal{F}_t$-measurable.

For every $B \in \mathcal{B}(\mathbb{R})$ and $t \in \mathbb{R}_+$,

$$\{X(T) \in B\} \cap \{T \leq t\} = \{X(T \wedge t) \in B\} \cap \{T \leq t\} \in \mathcal{F}_t. \qquad \square$$

Stochastic Lebesgue-Stieltjes integration

If X is a finite variation process, define the processes $[X], X^\oplus, X^\ominus, X_-, \Delta X$ by

$$\begin{aligned}[X](t,\omega) &= [X(\cdot,\omega)](t),\\ X^\oplus &= \frac{[X]+X}{2},\\ X^\ominus &= \frac{[X]-X}{2},\\ X_-(t,\omega) &= X(\cdot,\omega)_-(t),\\ (\Delta X)(t,\omega) &= (\Delta X(\cdot,\omega))(t).\end{aligned}$$

Proposition 4.5. *If X is a right continuous adapted process, then so are $[X], X^\oplus, X^\ominus$.*

Proof. Suppose $t \in \mathbb{R}_+$, then by proposition 2.1(e),

$$[X](t) = \lim_{n\to\infty} \sum_{i=1}^{2^n} \left| X\left(\frac{i}{2^n}t\right) - X\left(\frac{i-1}{2^n}t\right) \right|,$$

therefore $[X](t)$ is $\mathcal{F}_t$-measurable. It follows that $X^{\oplus}(t), X^{\ominus}(t)$ are also $\mathcal{F}_t$-measurable. □

Suppose Y is an integrator process. A process X is called Y-integrable if for almost every $\omega \in \Omega$, $X(\cdot,\omega)$ is $Y(\cdot,\omega)$-integrable. If X is Y-integrable, define the random variable $\int X\,dY$ by

$$\left(\int X\,dY\right)(\omega) = \int X(\cdot,\omega)\,dY(\cdot,\omega).$$

A process X is called locally Y-integrable if for every $t \in \mathbb{R}_+$, $X1([0,t])$ is Y-integrable. If X is locally Y integrable, define the process $X \cdot Y$ by

$$(X \cdot Y)(t,\omega) = (X(\cdot,\omega) \cdot Y(\cdot,\omega))(t) = \int_{[0,t]} X(\cdot,\omega)\,dY(\cdot,\omega).$$

Proposition 4.6. *Suppose Y is an integrator process and X is a predictable locally Y-integrable process.*

(a) If Y is adapted, then $X \cdot Y$ is adapted.

(b) If Y is predictable, then $X \cdot Y$ is predictable.

Proof. (a). Let $\mathcal{H}$ be the set of all processes X such that X is predictable, locally Y-integrable and $X \cdot Y$ is adapted. The following are easy to verify:

(i) $\mathcal{H}$ is a vector subspace of $\mathbb{R}^{\mathbb{R}_+ \times \Omega}$.

(ii) $1 \in \mathcal{H}$.

(iii) If X is a bounded process and (X_n) is a sequence of nonnegative elements of $\mathcal{H}$ such that $X_n \uparrow X$, then $X \in \mathcal{H}$.

Suppose $s, u \in \mathbb{R}_+$, $s \le u$, $A \in \mathcal{F}_s$ and $X = 1((s,u] \times A)$, then X is clearly predictable and locally Y-integrable. Furthermore,

$$\begin{aligned}(X \cdot Y)(t,\omega) &= \int_{[0,t]} 1(A)(\omega)1((s,u])\,dY(\cdot,\omega) \\ &= 1(A)(\omega)\left(Y^{(u)}(t,\omega) - Y^{(s)}(t,\omega)\right).\end{aligned}$$

Suppose $t \in \mathbb{R}_+$. If $t \le s$, then $(X \cdot Y)(t) = 0$ is $\mathcal{F}_t$-measurable. If $t \in (s,u]$, then

$$(X \cdot Y)(t) = 1(A)(Y(t) - Y(s))$$

is $\mathcal{F}_t$-measurable because $\mathcal{F}_s \subseteq \mathcal{F}_t$. If $t \geq u$, then

$$(X \cdot Y)(t) = 1(A)(Y(u) - Y(s))$$

is $\mathcal{F}_t$-measurable because $\mathcal{F}_s \subseteq \mathcal{F}_u \subseteq \mathcal{F}_t$.

Suppose $A \in \mathcal{F}_0$ and $X = 1(\{0\} \times A)$, then X is clearly predictable and locally Y-integrable. Furthermore, for every $t \in \mathbb{R}_+$, $(X \cdot Y)(t) = 1(A)Y(0)$ is $\mathcal{F}_0$-measurable, and therefore also $\mathcal{F}_t$-measurable.

It follows that for every $A \in \mathcal{P}_1$, $1(A) \in \mathcal{H}$, therefore $\mathcal{H}$ contains all bounded predictable processes (see [1], appendix A1, theorem 4, or [10], chapter I, theorem 8).

Finally, suppose that X is an arbitrary predictable locally Y-integrable process, and that $t \in \mathbb{R}_+$. For every $n \in \mathbb{N}$, let $X_n = (X \wedge n) \vee (-n)$. X_n is a bounded predictable process, therefore $X_n \in \mathcal{H}$. By applying 2.6(d) for every $\omega \in \Omega$ to the functions

$$X(\cdot, \omega)1([0,t]),\ X_1(\cdot, \omega)1([0,t]),\ X_2(\cdot, \omega)1([0,t]),\ \ldots,$$

it follows that $(X_n \cdot Y)(t) \longrightarrow (X \cdot Y)(t)$, so $(X \cdot Y)(t)$ is $\mathcal{F}_t$-measurable. □

(b). Let $\mathcal{H}$ be the set of all processes X such that X is predictable, locally Y-integrable and $X \cdot Y$ is predictable. It is sufficient (by using the same technique as in part (a)) to show that for every $A \in \mathcal{P}_1$, $1(A) \in \mathcal{H}$.

Suppose $s, u \in \mathbb{R}_+$, $s \leq u$, $A \in \mathcal{F}_s$ and $X = 1((s, u] \times A)$, then X is clearly predictable and locally Y-integrable. Furthermore,

$$\begin{aligned}(X \cdot Y)(t, \omega) &= 1(A)(\omega)\left(Y^{(u)}(t, \omega) - Y^{(s)}(t, \omega)\right)\\ &= \begin{cases} Y(t, \omega) - Y(s, \omega) & \text{if } (t, \omega) \in (s, u] \times A \\ Y(u, \omega) - Y(s, \omega) & \text{if } (t, \omega) \in (u, \infty) \times A\ , \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

i.e.

$$\begin{aligned}X \cdot Y &= \Big(Y - Y(s)\Big)1((s, u] \times A) + \Big(Y(u) - Y(s)\Big)1((u, \infty) \times A)\\ &= Y1((s, u] \times A) - Y(s)1((s, u] \times A) + \lim_{n\to\infty}\Big(Y(u) - Y(s)\Big)1((u, n] \times A),\end{aligned}$$

so $X \cdot Y$ is predictable.

Suppose $A \in \mathcal{F}_0$ and $X = 1(\{0\} \times A)$, then X is clearly predictable and locally Y-integrable. Furthermore, $X \cdot Y = 1(A)Y(0)$ is left continuous and adapted, so $X \cdot Y$ is predictable. □

Martingales

A set $\mathcal{H}$ of integrable random variables is said to be *uniformly integrable* if for every $\epsilon > 0$ there exists $M \in \mathbb{R}_+$ such that for every $X \in \mathcal{H}$,

$$\int_{\{|X|>M\}} |X|\ d\mathbb{P} < \epsilon.$$

Proposition 4.7. *(a) If $X_1, X_2, \ldots, X_n$ are integrable random variables, then $\{X_1, X_2, \ldots, X_n\}$ is uniformly integrable.*

(b) If $\mathcal{H}$ is uniformly integrable, and $\mathcal{H}'$ is a set of random variables such that for every $X \in \mathcal{H}'$ there exists $Y \in \mathcal{H}$ such that $|X| \leq Y$, then $\mathcal{H}'$ is uniformly integrable.

(c) Suppose $\mathcal{H}$ is a set of integrable random variables, then the following statements are equivalent:

(i) $\mathcal{H}$ is uniformly integrable.

(ii) There exists $f \in {\mathbb{R}_+}^{\mathbb{R}_+}$ such that f is nondecreasing, convex,

$$\lim_{t\to\infty} \frac{f(t)}{t} = \infty$$

and there exists $M \in \mathbb{R}_+$ such that for every $X \in \mathcal{H}$, $\mathbb{E}[f \circ |X|] \leq M$.

(d) If $\mathcal{H}$ is uniformly integrable and $\mathbb{G}$ is a set of sub-σ-algebras of $\mathcal{F}$, then

$$\{\mathbb{E}[X|\mathcal{G}] : X \in \mathcal{H}, \mathcal{G} \in \mathbb{G}\}$$

is uniformly integrable.

Proof. ***(c).*** See [3], paragraph II-22. □

(d). Note that $\{\mathbb{E}[X|\mathcal{G}] : X \in \mathcal{H}, \mathcal{G} \in \mathbb{G}\}$ is a set of integrable random variables. $\mathcal{H}$ is uniformly integrable, so let f be as in part (c). Suppose $X \in \mathcal{H}$ and $\mathcal{G} \in \mathbb{G}$, then by Jensen's inequality and the fact that f is nondecreasing and convex,

$$\begin{aligned}\mathbb{E}\Big[f \circ \Big|\mathbb{E}[X|\mathcal{G}]\Big|\Big] &\leq \mathbb{E}\Big[f \circ \mathbb{E}\Big[|X| \;\Big|\; \mathcal{G}\Big]\Big] \\ &\leq \mathbb{E}\Big[\mathbb{E}\Big[f \circ |X| \;\Big|\; \mathcal{G}\Big]\Big] \\ &= \mathbb{E}[f \circ |X|]\end{aligned}$$

The result follows by part (c). □

Suppose X is a process. X is said to be *uniformly integrable* if

$$\{X(t) : t \in \mathbb{R}_+\}$$

is uniformly integrable, and to be of *class D* if

$$\{X(T) : T \text{ is a finite stopping time}\}$$

is uniformly integrable.

A process X is called a *submartingale* (*martingale*) if:

(i) X is adapted.

(ii) For every $t \in \mathbb{R}_+$, $X(t)$ is integrable.

(iii) For every $s, t \in \mathbb{R}_+$ such that $s \leq t$, $X(s) \leq \mathbb{E}[X(t)|\mathcal{F}_s]$ $(=)$.

Note that every martingale is also a submartingale. A martingale X is said to be *square integrable* if there exists $M \in \mathbb{R}_+$ such that for every $t \in \mathbb{R}_+$, $\mathbb{E}[X^2(t)] \leq M$.

Proposition 4.8. *Suppose X is a submartingale, and define $f \in \mathbb{R}^{\mathbb{R}_+}$ by $f(t) = \mathbb{E}[X(t)]$. If f is right continuous, then X has a right continuous modification.*

Proof. See [4], paragraph VI-4. □

In particular, if X is a martingale, then f is constant, therefore X has a right continuous modification.

If X is a submartingale (martingale) and Y is a integrable random variable, then X is said to be *subclosed* (*closed*) by Y if for every $t \in \mathbb{R}_+$, $X(t) \leq \mathbb{E}[Y|\mathcal{F}_t]$ $(=)$.

By proposition 4.7(a) and (d), a closed martingale is uniformly integrable.

Proposition 4.9 (Submartingale convergence theorem). *If X is a right continuous uniformly integrable submartingale (martingale), then $\lim_{t\to\infty} X(t)$ exists, is integrable and subcloses (closes) X. If Y also subcloses (closes) X, then $\lim_{t\to\infty} X(t) \leq \mathbb{E}[Y|\mathcal{F}_\infty]$ $(=)$.*

Proof. See [4], paragraph VI-6. □

If X is a right continuous uniformly integrable submartingale, then the submartingale convergence theorem can be used in an obvious way to extend X to $\bar{\mathbb{R}}_+ \times \Omega$, so that $X(T)$ is defined for stopping times T that are not necessarily finite.

Proposition 4.10 (Doob's optional sampling theorem). *Suppose X is a right continuous submartingale (martingale).*

(a) If S, T are bounded stopping times such that $S \leq T$, then $X(S), X(T)$ are integrable, and $X(S) \leq \mathbb{E}[X(T)|\mathcal{F}_S]$ $(=)$.

(b) If X is uniformly integrable and S, T are stopping times such that $S \leq T$, then $X(S), X(T)$ are integrable, and $X(S) \leq \mathbb{E}[X(T)|\mathcal{F}_S]$ $(=)$.

Proof. See [4], paragraph VI-10. □

Proposition 4.11. *(a) If X is a right continuous submartingale (martingale) and T is a stopping time, then $X^{(T)}$ is also a submartingale (martingale).*

(b) Every nonnegative right continuous subclosed submartingale is uniformly integrable, and is of class D.

(c) Every uniformly integrable martingale is of class D.

Proof. ***(a).*** See [4], paragraph VI-12. □

(b). Suppose X is a nonnegative right continuous subclosed submartingale. By proposition 4.7(a),(b) and (d), X is uniformly integrable. For every finite stopping time T, $X(T) \leq \mathbb{E}[X(\infty)|\mathcal{F}_T]$ (Doob's optional sampling theorem), therefore by proposition 4.7(a),(b) and (d) again, X is of class D. □

(c). Suppose X is a right continuous uniformly integrable martingale. Note that X is of class D if and only if $|X|$ is. Now by Jensen's inequality, $|X|$ is a nonnegative right continuous subclosed submartingale, therefore it is of class D. □

A process X is called a *local submartingale* (*martingale*, *square integrable martingale*) if there exists a sequence (T_n) of stopping times such that $T_n \uparrow \infty$, and for every $n \in \mathbb{N}$, $X^{(T_n)}$ is a submartingale (martingale, square integrable martingale). The sequence (T_n) is called a *fundamental sequence for* X. Note that every submartingale (martingale, square integrable martingale) is also a local submartingale (martingale, square integrable martingale) (use the sequence $T_n = \infty$).

A process X is said to be *zero at zero* if $X(0) = 0$.

Proposition 4.12. *Every finite variation predictable process that is also a zero at zero local martingale is indistinguishable from 0.*

Proof. See [4], paragraph VI-80. □

Proposition 4.13 (Doob-Meyer decomposition). *(a) If X is a right continuous submartingale of class D, then there exists a nondecreasing right continuous predictable process Y such that $X - Y$ is a zero at zero uniformly integrable martingale. Y is unique up to indistinguishability.*

(b) If X is a right continuous local submartingale, then there exists a nondecreasing right continuous predictable process Y such that $X - Y$ is a zero at zero local martingale. Y is unique up to indistinguishability.

Proof. (a). See [4], paragraph VII-9(b). □

(b). See [4], paragraph VII-12. □

If X is a local square integrable martingale, then by Jensen's inequality, X^2 is a local submartingale. Let $\langle X \rangle$ be the nondecreasing right continuous predictable process in the Doob-Meyer decomposition of X^2. Because X^2 is nonnegative, $\langle X \rangle$ is also nonnegative and therefore a distribution process. $\langle X \rangle$ is called the *predictable variation process of* X.

Chapter 5

Hattendorff's theorem

Consider a mathematical model of a life insurance policy with the following elements:

(i) A complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $(\mathcal{F}_t)$ satisfying the usual assumptions. These decribe all the uncertainties associated with the policy.

(ii) An adapted integrator process P. $P^{\oplus}, P^{\ominus}$ are then adapted distribution processes representing the company's outgo payment stream (including e.g. policy benefits), and its income payment stream (including e.g. premiums) respectively.

(iii) An adapted P-integrable discount process v, such that for almost every $\omega \in \Omega$, $v(\cdot, \omega)$ is the discount function used to assign values to both $P^{\oplus}(\cdot, \omega)$ and $P^{\ominus}(\cdot, \omega)$.

(iv) v and P are such that $\int v\, dP$ is a $\mathcal{F}_\infty$-measurable, square integrable random variable.

Let

$$W = \int v\, dP = \int v\, dP^{\oplus} - \int v\, dP^{\ominus}.$$

W represents the policy's *net present value.*

It is characteristic of a life insurance policy that the company first receives premiums and then pays benefits only much later. The company cannot, however, in the meantime declare all of the received premiums as profit and distribute it to its shareholders, because when the time then comes to pay benefits there will be nothing available. It therefore has to set part of the received premiums aside as a reserve, which is entered into its balance sheet as a liability. The reserve at a particular time should be such that it, together with the investment returns to be earned on it, is sufficient to meet the estimated future net outgo, given the latest information. Define the process V by

$$V(t,\omega) = \frac{1}{v(t)}\mathbb{E}\left[\int_{(t,\infty)} v\,dP \,\middle|\, \mathcal{F}_t\right](\omega).$$

$V(t)$ represents the *reserve at time t*.

Accounting convention states that the net loss over a period is the net outgo plus the net increase in liabilities over that period. If S, T are finite stopping times such that $S \leq T$, let

$$L_{(S,T]} = \int_{(S,T]} v\,dP + v(T)V(T) - v(S)V(S).$$

$L_{(S,T]}$ represents the *net present value of the loss over* $(S,T]$.

Define the process M by

$$M(t,\omega) = \mathbb{E}[W|\mathcal{F}_t](\omega).$$

M is a uniformly integrable martingale, and may therefore be assumed to be right continuous. W is $\mathcal{F}_\infty$-measurable, therefore $M(t) \longrightarrow W$ (submartingale convergence theorem). Because v is a continuous adapted process and P is adapted, $v \cdot P$ is adapted. It follows that for every $t \in \mathbb{R}_+$,

$$M(t) = \mathbb{E}\left[\int_{[0,t]} v\,dP + \int_{(t,\infty)} v\,dP \,\middle|\, \mathcal{F}_t\right] = (v \cdot P)(t) + v(t)V(t),$$

therefore $L_{(S,T]} = M(T) - M(S)$.

Proposition 5.1 (Hattendorff's theorem). *(i) If S, T are finite stopping times such that $S \leq T$, and $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}_S$, then*

$$\mathbb{E}[L_{(S,T]}|\mathcal{G}] = 0.$$

(ii) If S, T, U, V are finite stopping times such that $S \leq T \leq U \leq V$, and $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}_S$, then

$$\operatorname{cov}[L_{(S,T]}, L_{(U,V]}|\mathcal{G}] = 0.$$

Proof. ***(a).*** By Doob's optional sampling theorem,

$$\mathbb{E}[L_{(S,T]}|\mathcal{G}] = \mathbb{E}\Big[\mathbb{E}[L_{(S,T]}|\mathcal{F}_S] \,\Big|\, \mathcal{G}\Big] = \mathbb{E}\Big[\mathbb{E}[M(T) - M(S)|\mathcal{F}_S] \,\Big|\, \mathcal{G}\Big] = 0. \quad \square$$

(b). By part (a),

$$\operatorname{cov}[L_{(S,T]}, L_{(U,V]}|\mathcal{G}] = \mathbb{E}[L_{(S,T]}L_{(U,V]}|\mathcal{G}] = \mathbb{E}\Big[L_{(S,T]}\mathbb{E}[L_{(U,V]}|\mathcal{F}_T] \,\Big|\, \mathcal{G}\Big] = 0. \quad \square$$

Although Hattendorff's theorem is treated here in a life insurance context, it can be applied to any financial operation using the same definition of loss.

Chapter 6

Point processes

Once again, suppose that a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given.

Definitions

A system $((Z_n)_{n\in\mathbb{N}_0}, (T_n)_{n\in\mathbb{N}_0})$ is called a *point process* if

(i) (Z_n) is a sequence of $\mathcal{F}$-measurable elements of ${\mathbb{N}_0}^{\Omega}$ such that $\bigcup_{i=0}^{\infty} Z_i(\Omega)$ is a finite set, $Z_0 = 0$ and for every $n \in \mathbb{N}$, $Z_n \neq Z_{n-1}$.

(ii) (T_n) is a sequence of $\mathcal{F}$-measurable finite random times such that $T_0 = 0$ and (T_n) increases strictly to ∞.

For the remainder of this chapter, suppose that a point process $((Z_n), (T_n))$ is given.

Let

$$\mathcal{S} = \bigcup_{i=0}^{\infty} Z_i(\Omega), \quad \mathcal{T} = \{(k,l) : k, l \in \mathcal{S}, k \neq l\}.$$

$\mathcal{S}$ is called the *state space*, and $\mathcal{T}$ the *transition space*.

For every $(k,l) \in \mathcal{T}$, let

$$N_{kl} = \sum_{i=1}^{\infty} 1(Z_{i-1} = k) 1(Z_i = l) 1([T_i, \infty)).$$

N_{kl} is called the *counting process of the transition* (k,l).

Also, let

$$N = \sum_{(k,l)\in\mathcal{T}} N_{kl}, \quad Z = \sum_{i=0}^{\infty} Z_i 1([T_i, T_{i+1})), \quad T = \sum_{i=0}^{\infty} T_i 1([T_i, T_{i+1})).$$

Z is called the *current state* and T the *time of last transition*.

For every $k \in \mathcal{S}$, let $I_k = 1(Z = k)$; and finally, define $U \in (\mathcal{S}^{\mathbb{R}_+})^{\mathbb{R}_+\times\Omega}$ by

$$U(t,\omega)(s) = Z(s \wedge T_-(t,\omega), \omega).$$

Proposition 6.1. *For every $k \in \mathcal{S}$,*

$$I_k = \delta_{0k} + \sum_{l \in \mathcal{S}, l \neq k} (N_{lk} - N_{kl}).$$

Proof.

$$\begin{aligned} I_k &= \sum_{i=0}^{\infty} 1(Z_i = k)\Big(1([T_i, \infty)) - 1([T_{i+1}, \infty))\Big) \\ &= 1(Z_0 = k) + \sum_{i=1}^{\infty} 1(Z_i = k) 1([T_i, \infty)) - \sum_{i=1}^{\infty} 1(Z_{i-1} = k) 1([T_i, \infty)) \\ &= \delta_{0k} + \sum_{l \in \mathcal{S}, l \neq k} (N_{lk} - N_{kl}). \end{aligned}$$

□

The filtration generated by $((Z_n), (T_n))$

For every $t \in \mathbb{R}_+$, let

$$\mathcal{F}_t^0 = \sigma(N_{kl}(s) : (k, l) \in \mathcal{T}, s \in [0, t]),$$

and let $\mathcal{F}_t$ be the set of all subsets A of Ω for which there exist $B \in \mathcal{F}_t^0, N \in \mathcal{F}$ such that $\mathbb{P}[N] = 0$ and $B \setminus N \subseteq A \subseteq B \cup N$. $\mathcal{F}_t$ is a σ-algebra containing $\mathcal{F}_t^0$, and $(\mathcal{F}_t)$ is a complete filtration.

Proposition 6.2. *$(\mathcal{F}_t)$ is right continuous.*

Proof. It must first be shown that the filtration $(\mathcal{F}_t^0)$ is right continuous. The crucial property to be used is that for every $(t, \omega) \in \mathbb{R}_+ \times \Omega$ there exists $\epsilon > 0$ such that for every $(s, (k, l)) \in [t, t + \epsilon) \times \mathcal{T}$,

$$N_{kl}(s, \omega) = N_{kl}(t, \omega).$$

For every $t \in \mathbb{R}_+$, define $X_t \in ({\mathbb{R}_+}^{\mathbb{R}_+ \times \mathcal{T}})^\Omega$ by

$$X_t(\omega)(s, (k, l)) = N_{kl}(s \wedge t, \omega),$$

and note that $\mathcal{F}_t^0 = \sigma(X_t)$.

Suppose $t \in \mathbb{R}_+$ and

$$A \in \bigcap_{s > t} \mathcal{F}_s^0 = \bigcap_{n=1}^{\infty} \mathcal{F}_{t+1/n}^0,$$

i.e. for every $n \in \mathbb{N}$, there exists

$$B_n \in \bigotimes_{\mathbb{R}_+ \times \mathcal{T}} \mathcal{B}(\mathbb{R}_+)$$

such that $A = \{X_{t+1/n} \in B_n\}$.

For every $n \in \mathbb{N}$, let $C_n = \{X_t = X_{t+1/n}\}$. (C_n) is a nondecreasing sequence, and by the crucial property noted above, its union is Ω, therefore

$$\begin{aligned} A &= \liminf(A \cap C_n) \\ &= \liminf(\{X_{t+1/n} \in B_n\} \cap C_n) \\ &= \liminf(\{X_t \in B_n\} \cap C_n) \\ &= \liminf(\{X_t \in B_n\}) \in \mathcal{F}_t^0. \end{aligned}$$

Finally, suppose that

$$A \in \bigcap_{s>t} \mathcal{F}_s = \bigcap_{n=1}^{\infty} \mathcal{F}_{t+1/n},$$

i.e. for every $n \in \mathbb{N}$ there exist $B_n \in \mathcal{F}_{t+1/n}^0, N_n \in \mathcal{F}$ such that $\mathbb{P}[N_n] = 0$ and

$$B_n \setminus N_n \subseteq A \subseteq B_n \cup N_n.$$

Now

$$\limsup(B_n) \in \bigcap_{n=1}^{\infty} \mathcal{F}_{t+1/n}^0 = \mathcal{F}_t^0, \quad \limsup(N_n) \in \mathcal{F}, \quad \mathbb{P}[\limsup(N_n)] = 0$$

and $\limsup(B_n) \setminus \limsup(N_n) \subseteq A \subseteq \limsup(B_n) \cup \limsup(N_n)$, i.e. $A \in \mathcal{F}_t$. □

It follows that $(\mathcal{F}_t)$ satisfies the usual assumptions. This filtration is used for the remainder of this chapter.

Proposition 6.3. *(a) For every $n \in \mathbb{N}_0$, T_n is a stopping time, and*

$$\sigma(T_0, Z_0, \ldots, T_n, Z_n) \subseteq \mathcal{F}_{T_n}.$$

(b) U is $\mathcal{P}$-measurable.

Proof. (a). For every $t \in \mathbb{R}_+$, $\{T_n \le t\} = \{N(t) \ge n\} \in \mathcal{F}_t$, i.e. T_n is a stopping time.

Suppose $i = 0, 1, \ldots, n$. T_i is $\mathcal{F}_{T_i}$-measurable and $\mathcal{F}_{T_i} \subseteq \mathcal{F}_{T_n}$, so T_i is $\mathcal{F}_{T_n}$-measurable. If $i = 0$, then Z_i is constant and therefore $\mathcal{F}_{T_n}$-measurable. If $i > 0$, then for every $l \in \mathcal{S}$,

$$1(Z_i = l) = \sum_{k \in \mathcal{S}, k \ne l} 1(Z_{i-1} = k) 1(Z_i = l) = \sum_{k \in \mathcal{S}, k \ne l} (N_{kl}(T_i) - N_{kl}(T_{i-1})).$$

Now for every $(k, l) \in \mathcal{T}$, N_{kl} is a right continuous adapted process, and therefore progressive. Furthermore, for every $j = 0, 1, \ldots, n$, T_j is a finite stopping time, therefore $N_{kl}(T_j)$ is $\mathcal{F}_{T_j}$-measurable (proposition 4.4(d)), and $\mathcal{F}_{T_j} \subseteq \mathcal{F}_{T_n}$. It follows that Z_i is $\mathcal{F}_{T_n}$-measurable. □

(b). It is sufficient to show that for every $s \in \mathbb{R}_+$, the process $U(\cdot,\cdot)(s)$ is predictable. Suppose $s \in \mathbb{R}_+$, then for every $t \in \mathbb{R}_+$ and almost every $\omega \in \Omega$,

$$\begin{aligned}
&U(t,\omega)(s)\\
&= \sum_{i=0}^{\infty} Z_i(\omega).1([T_i, T_{i+1}))(s \wedge T_-(t,\omega),\omega))\\
&= Z_0(\omega)1(0)(t) + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} Z_i(\omega).1([T_i,T_{i+1}))(s\wedge T_j(\omega),\omega).1((T_j,T_{j+1}])(t,\omega)\\
&= Z_0(\omega)1(0)(t) + \sum_{j=0}^{\infty}\left[\sum_{i=0}^{\infty} Z_i(\omega).1([T_i,T_{i+1}))(s\wedge T_j(\omega),\omega)\right]1((T_j,T_{j+1}])(t,\omega)\\
&= Z_0(\omega)1(\{0\}\times\Omega)(t,\omega) + \sum_{j=0}^{\infty} X_{s,j}(t,\omega)1((T_j,T_{j+1}])(t,\omega),
\end{aligned}$$

where $X_{s,j}$ is the process defined by

$$X_{s,j}(t,\omega) = \begin{cases} Z_0(\omega)1([T_0,\infty))(s,\omega) & \text{if } j=0\\ \sum_{i=0}^{j-1} Z_i(\omega)1([T_i,T_{i+1}))(s,\omega) + Z_j(\omega)1([T_j,\infty))(s,\omega) & \text{if } j>0. \end{cases}$$

By part (a), $X_{s,j}$ is $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{F}_{T_j}$-measurable, therefore $U(\cdot,\cdot)(s)$ is predictable. □

Stochastic intensities

For this section, suppose that $(k,l) \in \mathcal{T}$ is given. Define the process τ by $\tau(t,\omega) = t$. Note that N_{kl} and τ are adapted distribution processes.

A process λ_{kl} is called a (k,l)-*intensity* if it is nonnegative, predictable, locally τ-integrable and for every nonnegative predictable process X

$$\mathbb{E}\left[\int X\, dN_{kl}\right] = \mathbb{E}\left[\int X\, d(\lambda_{kl}\cdot\tau)\right].$$

For the remainder of this section, suppose there exists a (k,l)-intensity λ_{kl}. Let $M_{kl} = N_{kl} - \lambda_{kl}\cdot\tau$. Because N_{kl} and $\lambda_{kl}\cdot\tau$ are adapted distribution processes, M_{kl} is an adapted integrator process.

Proposition 6.4. *Suppose X is a predictable process.*

(a) If for every $t \in \mathbb{R}_+$,

$$\mathbb{E}\left[\int_{[0,t]} |X|\, d(\lambda_{kl}\cdot\tau)\right] < \infty,$$

then X is locally M_{kl}-integrable and $X\cdot M_{kl}$ is a martingale.

(b) If X is locally $M_{kl}, \lambda_{kl} \cdot \tau$-integrable, then $X \cdot M_{kl}$ is a local martingale.

(c) If X is locally M_{kl}-integrable and X^2 is locally $\lambda_{kl} \cdot \tau$-integrable, then $X \cdot M_{kl}$ is a local square integrable martingale and $\langle X \cdot M_{kl} \rangle = X^2 \cdot (\lambda_{kl} \cdot \tau)$.

Proof. ***(a).*** For every $t \in \mathbb{R}_+$,

$$\int_{[0,t]} |X| \; d(\lambda_{kl} \cdot \tau) < \infty,$$

i.e. X is locally $\lambda \cdot \tau$-integrable. Furthermore, because $|X|$ is a nonnegative predictable process,

$$\mathbb{E}\left[\int_{[0,t]} |X| \; dN_{kl}\right] = \mathbb{E}\left[\int_{[0,t]} |X| \; d(\lambda_{kl} \cdot \tau)\right] < \infty,$$

therefore X is locally N_{kl}-integrable. It follows that X is locally M_{kl}-integrable.

X is predictable and M_{kl} is adapted, therefore $X \cdot M_{kl}$ is adapted.

For every $t \in \mathbb{R}_+$,

$$\mathbb{E}\Big[|(X \cdot N_{kl})(t)|\Big] \leq \mathbb{E}\left[\int_{[0,t]} |X| \; dN_{kl}\right] < \infty,$$

$$\mathbb{E}\Big[|(X \cdot (\lambda_{kl} \cdot \tau))(t)|\Big] \leq \mathbb{E}\left[\int_{[0,t]} |X| \; d(\lambda_{kl} \cdot \tau)\right] < \infty,$$

i.e. $(X \cdot N_{kl})(t)$ and $(X \cdot (\lambda_{kl} \cdot \tau))(t)$ are integrable random variables, therefore $(X \cdot M_{kl})(t)$ is also an integrable random variable.

Suppose $s, t \in \mathbb{R}_+$, $s \leq t$ and $A \in \mathcal{F}_s$, then $X^+ 1((s,t] \times A)$ and $X^- 1((s,t] \times A)$ are nonnegative predictable processes, therefore by using $X = X^+ - X^-$ and the integrability of $(X \cdot N_{kl})(t)$ and $(X \cdot (\lambda_{kl} \cdot \tau))(t)$,

$$\begin{aligned}
&\mathbb{E}\Big[1(A)\Big((X \cdot N_{kl})(t) - (X \cdot N_{kl})(s)\Big)\Big] \\
&= \mathbb{E}\left[\int X 1((s,t] \times A) \, dN_{kl}\right] \\
&= \mathbb{E}\left[\int X 1((s,t] \times A) \, d(\lambda_{kl} \cdot \tau)\right] \\
&= \mathbb{E}\Big[1(A)\Big((X \cdot (\lambda_{kl} \cdot \tau))(t) - (X \cdot (\lambda_{kl} \cdot \tau))(s)\Big)\Big].
\end{aligned}$$

After rearranging,

$$\mathbb{E}[1(A)(X \cdot M_{kl})(s)] = \mathbb{E}[1(A)(X \cdot M_{kl})(t)]. \qquad \square$$

(b). For every $n \in \mathbb{N}$, define the random time S_n by

$$S_n(\omega) = \inf\left\{t \in \mathbb{R}_+ : \int_{[0,t]} |X(\cdot,\omega)| \; d(\lambda_{kl} \cdot \tau)(\cdot,\omega) > n\right\}.$$

S_n is a stopping time, and because X is locally $\lambda_{kl} \cdot \tau$-integrable, $(S_n)\uparrow\infty$. For every $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$,

$$\mathbb{E}\left[\int_{[0,t]} |X1([0,S_n])| \, d(\lambda_{kl} \cdot \tau))\right] \leq n < \infty,$$

therefore $X1([0,S_n]) \cdot M_{kl} = (X \cdot M_{kl})^{(S_n)}$ is a martingale. □

(c). Suppose $t \in \mathbb{R}_+$. By the Cauchy-Schwartz inequality,

$$\left(\int_{[0,t]} |X| \, d(\lambda_{kl} \cdot \tau)\right) \leq \left(\int_{[0,t]} X^2 \, d(\lambda_{kl} \cdot \tau)\right)^{1/2} \left(\int_{[0,t]} 1^2 \, d(\lambda_{kl} \cdot \tau)\right)^{1/2} < \infty,$$

i.e. X is locally $\lambda_{kl} \cdot \tau$-integrable, therefore $X \cdot M_{kl}$ is a local martingale with (S_n) (as defined in part (b)) as a fundamental sequence.

Let $M = X \cdot M_{kl}$. M is an integrator process, and

$$\Delta M = X\Delta(N_{kl} - \lambda_{kl} \cdot \tau) = X\Delta N_{kl}.$$

By integration by parts,

$$\begin{aligned} M^2 &= M_- \cdot M + M \cdot M \\ &= M_- \cdot M + (M_- + \Delta M) \cdot M \\ &= (2M_-) \cdot M + (\Delta M) \cdot M \end{aligned}$$

Now

$$(\Delta M) \cdot M = \sum (\Delta M)^2 = \sum X^2 (\Delta N_{kl})^2 = \sum X^2 \Delta N_{kl} = X^2 \cdot N_{kl},$$

therefore

$$\begin{aligned} M^2 &= (2M_- X) \cdot M_{kl} + X^2 \cdot M_{kl} + X^2 \cdot (\lambda_{kl} \cdot \tau) \\ &= (2M_- + X)X \cdot M_{kl} + X^2 \cdot (\lambda_{kl} \cdot \tau). \end{aligned}$$

For every $n \in \mathbb{N}$, define the random times U_n, V_n, W_n by

$$\begin{aligned} U_n(\omega) &= \inf\left\{t \in \mathbb{R}_+ : \int_{[0,t]} X^2(\cdot,\omega) \, d(\lambda_{kl} \cdot \tau)(\cdot,\omega) > n\right\}, \\ V_n(\omega) &= \inf\{t \in \mathbb{R}_+ : M_-(t,\omega) > n\}, \\ W_n &= S_n \wedge U_n \wedge V_n. \end{aligned}$$

U_n, V_n, W_n are stopping times. Because X^2 is locally $\lambda_{kl} \cdot \tau$-integrable, $(U_n)\uparrow\infty$, and because M_- is a process, $(V_n)\uparrow\infty$, therefore $(W_n)\uparrow\infty$. For every $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$,

$$\begin{aligned} &\mathbb{E}\left[\int_{[0,t]} |(2M_- + X)X| \, 1([0,W_n]) d\,(\lambda_{kl} \cdot \tau)\right] \\ &\leq \mathbb{E}\left[\int_{[0,t]} \left|2nX + X^2\right| 1([0,W_n]) d\,(\lambda_{kl} \cdot \tau)\right] \\ &\leq 2n^2 + n < \infty, \end{aligned}$$

therefore

$$(2M_- + X)X1([0, W_n]) \cdot (\lambda_{kl} \cdot \tau) = ((2M_- + X)X \cdot (\lambda_{kl} \cdot \tau))^{(W_n)}$$

is a zero at zero martingale. By proposition 4.11(a), (W_n) is a fundamental sequence for the local martingale M, and for every $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$,

$$\mathbb{E}\left[\left(M^{(W_n)}(t)\right)^2\right] = 0 + \mathbb{E}\left[(X^2 \cdot (\lambda_{kl} \cdot \tau))^{(W_n)}(t)\right] \leq n,$$

i.e. M is a local square integrable martingale. It follows from the uniqueness of the Doob-Meyer decomposition that $\langle X \cdot M_{kl}\rangle = X^2 \cdot (\lambda_{kl} \cdot \tau)$. □

In particular, 1 is a predictable locally $M_{kl}, \lambda_{kl} \cdot \tau$-integrable process, therefore $M_{kl} = 1 \cdot M_{kl}$ is a local martingale. Because N_{kl} is a distribution process such that for every $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$, $\mathbb{E}[N_{kl}^{(T_n)}(t)] \leq \mathbb{E}[n] < \infty$, it is a local submartingale, therefore $\lambda_{kl} \cdot \tau$ is the nondecreasing right continuous predictable process in the Doob-Meyer decomposition of N_{kl}. 1^2 is also $\lambda_{kl} \cdot \tau$-integrable, therefore M_{kl} is a local square integrable martingale and $\langle M_{kl}\rangle = 1^2 \cdot (\lambda_{kl} \cdot \tau) = \lambda_{kl} \cdot \tau$.

Martingale representation

For the remainder of this chapter, suppose that for every $(k, l) \in \mathcal{T}$, there exists a (k, l)-intensity λ_{kl}.

Proposition 6.5 (Martingale representation theorem)**.** *If X is a zero at zero uniformly integrable martingale, then for every $(k, l) \in \mathcal{T}$ there exists a predictable locally $M_{kl}, \lambda_{kl} \cdot \tau$-integrable process X_{kl} such that*

$$X = \sum_{(k,l)\in\mathcal{T}} X_{kl} \cdot M_{kl}.$$

Proof. See [1], chapter III, theorem 9. □

Miscellaneous results

Proposition 6.6. *(a) If X is a predictable integrator process, then for every $(k, l) \in \mathcal{T}$, ΔX is a predictable locally N_{kl}-integrable process and $(\Delta X) \cdot N_{kl} = 0$.*

(b) Suppose X is a predictable integrator process. Then for every $k \in \mathcal{S}$, I_k is an adapted integrator process, and $I_k \cdot X$ is predictable.

(c) Suppose X is an integrator process and the difference between two submartingale of class D. Then for every $k \in \mathcal{S}$, $(I_k)_-(I_k)$ is locally X-integrable and $(I_k)_-(I_k) \cdot X$ is predictable.

Proof. (a). X_- is a left continuous adapted process, so ΔX is predictable. For almost every $\omega \in \Omega$, $(\Delta X)(\cdot, \omega)$ is nonzero at a countable number of points, and $\lambda_{kl} \cdot \tau$ is continuous, therefore for every $t \in \mathbb{R}_+$,

$$\mathbb{E}\left[\int_{[0,t]} |\Delta X| \, d(\lambda_{kl} \cdot \tau)\right] = \mathbb{E}[0] < \infty.$$

It follows by proposition 6.4(a) that ΔX is locally $M_{kl}, \lambda_{kl} \cdot \tau$-integrable, so it is also locally N_{kl}-integrable.

For the same reason as above, for every $t \in \mathbb{R}_+$,

$$\int_{[0,t]} (\Delta X)^2 \, d(\lambda_{kl} \cdot \tau) = 0,$$

i.e. $(\Delta X)^2$ is locally $\lambda_{kl} \cdot \tau$-integrable, therefore $(\Delta X) \cdot M_{kl}$ is a local square integrable martingale (proposition 6.4(c)), and

$$\langle (\Delta X) \cdot M_{kl} \rangle = (\Delta X)^2 \cdot (\lambda_{kl} \cdot \tau) = 0.$$

By the definition of $\langle (\Delta X) \cdot M_{kl} \rangle$, it follows that $((\Delta X) \cdot M_{kl})^2$ is a nonnegative zero at zero local martingale, therefore $(\Delta X) \cdot M_{kl} = 0$, so

$$(\Delta X) \cdot N_{kl} = (\Delta X) \cdot M_{kl} + (\Delta X) \cdot (\lambda_{kl} \cdot \tau) = 0. \qquad \square$$

(b). For every $m, n \in \mathbb{N}_0$ such that $m \leq n$, T_n is a stopping time and Z_m is $\mathcal{F}_{T_n}$-measurable, therefore $(T_n)_{\{Z_m = k\}}$ is a stopping time. It follows that

$$I_k = \sum_{i=0}^{\infty} 1(Z_i = k) 1([T_i, T_{i+1})) = \sum_{i=0}^{\infty} 1\Big(\Big[(T_i)_{\{Z_i = k\}}, (T_{i+1})_{\{Z_i = k\}}\Big)\Big)$$

is adapted, as well as right continuous and of finite variation, i.e. an integrator process.

By integration by parts,

$$\begin{aligned} I_k \cdot X &= I_k X - X_- \cdot I_k \\ &= \Big((I_k)_- \cdot X + X \cdot I_k\Big) - X_- \cdot I_k \\ &= (I_k)_- \cdot X + (\Delta X) \cdot I_k \end{aligned}$$

Because $(I_k)_-$ is a left continuous adapted process, it is predictable, therefore $(I_k)_- \cdot X$ is predictable. Furthermore, by part (a),

$$(\Delta X) \cdot I_k = (\Delta X) \cdot \left(\delta_{0k} + \sum_{l \in \mathcal{S}, l \neq k} (N_{lk} - N_{kl})\right) = X(0)\delta_{0k},$$

therefore $I_k \cdot X$ is predictable. $\square$

(c). $(I_k)_-$ is an adapted left continuous process, and I_k is an adapted right continuous process, therefore both are $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty$-measurable, so for every $\omega \in \Omega$, $(I_k)_-(\cdot, \omega) I_k(\cdot, \omega)$ is $\mathcal{B}(\mathbb{R}_+)$-measurable. Furthermore, $(I_k)_-(\cdot, \omega) I_k(\cdot, \omega)$ is bounded, therefore it is locally $X(\cdot, \omega)$-integrable.

There exist a right continuous finite variation predictable process Y and a zero at zero uniformly integrable martingale Z such that $X = Y + Z$ (Doob-Meyer decomposition). In turn, for every $(k,l) \in \mathcal{T}$ there exists a predictable locally $M_{kl}, \lambda_{kl} \cdot \tau$-integrable process Z_{kl} such that $Z = \sum_{(k,l)\in\mathcal{T}} Z_{kl} \cdot M_{kl}$ (martingale representation theorem), therefore

$$
\begin{aligned}
&(I_k)_-(I_k) \cdot X \\
&= (I_k)_-(I_k) \cdot Y + (I_k)_-(I_k) \cdot \left(\sum_{(k,l)\in\mathcal{T}} Z_{kl} \cdot (N_{kl} - \lambda_{kl} \cdot \tau) \right) \\
&= I_k \cdot \Big((I_k)_- \cdot Y \Big) - I_k \cdot \left(\sum_{(k,l)\in\mathcal{T}} (I_k)_- Z_{kl} \cdot (\lambda_{kl} \cdot \tau) \right) \\
&\qquad + \sum_{(k,l)\in\mathcal{T}} Z_{kl} \cdot \Big((I_k)_-(I_k) \cdot N_{kl} \Big) \\
&= I_k \cdot \left((I_k)_- \cdot Y - \sum_{(k,l)\in\mathcal{T}} (I_k)_- Z_{kl} \cdot (\lambda_{kl} \cdot \tau) \right) + 0
\end{aligned}
$$

$(I_k)_- \cdot Y$ is predictable, and $\sum_{(k,l)\in\mathcal{T}} (I_k)_- Z_{kl} \cdot (\lambda_{kl} \cdot \tau)$ is an adapted continuous process, therefore $(I_k)_- I_k \cdot X$ is predictable. $\square$

Chapter 7

Thiele's differential equation

The following additions and alterations are made to the model in chapter 5:

(i) $(\mathcal{F}_t)$ is the filtration generated by a point process $((Z_n),(T_n))$. $\mathcal{S}$ represents the possible states of the policy (e.g. "active", "retired", "dead"), and $\mathcal{T}$ the possible transitions. $((Z_n),(T_n))$ is assumed to be such that for every $(k,l) \in \mathcal{T}$, there exists a (k,l)-intensity λ_{kl}.

(ii) For every $k \in \mathcal{S}$, a predictable integrator process a_k, and for every $(k,l) \in \mathcal{T}$, a predictable locally N_{kl}-integrable process A_{kl}, such that

$$P = \sum_{k \in \mathcal{S}} I_k \cdot a_k + \sum_{(k,l) \in \mathcal{T}} A_{kl} \cdot N_{kl}.$$

a_k represents the net payment stream while the policyholder is in state k (e.g. premium received while "active", pension paid while "retired") , and A_{kl} the net payment made upon transition from state k to state l (e.g. a benefit paid upon transition from "active" to "dead").

Let

$$\tilde{a}_k = v \cdot a_k, \quad \tilde{A}_{kl} = vA_{kl}, \quad \tilde{V} = vV.$$

$\tilde{a}_k$ and $\tilde{A}_{kl}$ are again a predictable integrator process and a predictable locally N_{kl}-integrable process respectively.

By the martingale representation theorem, for every $(k,l) \in \mathcal{T}$ there exists a predictable locally $M_{kl}, \lambda_{kl} \cdot \tau$-integrable process X_{kl} such that

$$\tilde{V} = v \cdot P - M = v \cdot P - \sum_{(k,l) \in \mathcal{T}} X_{kl} \cdot M_{kl} + M(0),$$

therefore $\tilde{V}$ is an integrator process.

Because $P^{\oplus}, P^{\ominus}$ are distribution processes, and v is a nonnegative $P^{\oplus}, P^{\ominus}$-integrable process, $v \cdot P^{\oplus}, v \cdot P^{\ominus}$ are nonnegative right continuous subclosed submartingales, and therefore also of class D (proposition 4.11(b)). M is a uniformly integrable martingale, and therefore also a submartingale of class D (proposition 4.11(c)). It follows that

$$\tilde{V} = v \cdot P - M = v \cdot P^{\oplus} - (v \cdot P^{\ominus} + M)$$

is the difference between two submartingales of class D.

Because the entire history of the policy can be constructed from Z, T, U, it is reasonable to assume that the exists $f \in \mathbb{R}^{\mathcal{S} \times \mathbb{R}_+ \times \mathcal{S}^{\mathbb{R}_+}}$ such that f is measurable and $\tilde{V} = f \circ (Z, T, U)$. For every $k \in \mathcal{S}$, let

$$\tilde{V}_k = f \circ (k, I_k T + (1 - I_k)\tau, U), \quad \tilde{V}'_k = f \circ (k, \tau, U).$$

By proposition 6.3(b), $\tilde{V}'_k$ is predictable. For every $(k,l) \in \mathcal{T}$, let

$$\tilde{R}_{kl} = \tilde{A}_{kl} + (\tilde{V}_l)_- - (\tilde{V}_k)_-.$$

$\tilde{R}_{kl}$ represents the *sum at risk of the transition* (k,l).

Proposition 7.1 (Thiele's differential equation). *If for every* $(k,l) \in \mathcal{T}$, $\tilde{R}_{kl}$ *is locally* $M_{kl}, \lambda_{kl} \cdot \tau$*-integrable, then*

$$\sum_{k \in \mathcal{S}} I_k \cdot (\tilde{a}_k + \tilde{V}_k) + \sum_{(k,l) \in \mathcal{T}} \tilde{R}_{kl} \cdot (\lambda_{kl} \cdot \tau) = a_0(0) + V_0(0).$$

Proof.

$$M = v \cdot P + \tilde{V} = \sum_{k \in \mathcal{S}} I_k \cdot \tilde{a}_k + \sum_{(k,l) \in \mathcal{T}} \tilde{A}_{kl} \cdot N_{kl} + \sum_{k \in \mathcal{S}} I_k \tilde{V}_k.$$

By integration by parts,

$$\begin{aligned}
\sum_{k \in \mathcal{S}} I_k \tilde{V}_k &= \sum_{k \in \mathcal{S}} (I_k \cdot \tilde{V}_k + (\tilde{V}_k)_- \cdot I_k) \\
&= \sum_{k \in \mathcal{S}} \left(I_k \cdot \tilde{V}_k + (\tilde{V}_k)_- \cdot \left(\delta_{0k} + \sum_{l \in \mathcal{S}, l \neq k} (N_{lk} - N_{kl}) \right) \right) \\
&= \sum_{k \in \mathcal{S}} I_k \cdot \tilde{V}_k + \sum_{(k,l) \in \mathcal{T}} (\tilde{V}_k)_- \cdot (N_{lk} - N_{kl}) \\
&= \sum_{k \in \mathcal{S}} I_k \cdot \tilde{V}_k + \sum_{(k,l) \in \mathcal{T}} ((\tilde{V}_l)_- - (\tilde{V}_k)_-) \cdot N_{kl},
\end{aligned}$$

therefore

$$M - \sum_{(k,l) \in \mathcal{T}} \tilde{R}_{kl} \cdot M_{kl} - M(0) = \sum_{k \in \mathcal{S}} I_k \cdot \tilde{a}_k + \sum_{k \in \mathcal{S}} I_k \cdot \tilde{V}_k + \sum_{(k,l) \in \mathcal{T}} \tilde{R}_{kl} \cdot (\lambda_{kl} \cdot \tau) - M(0).$$

The left side of the above equation is a zero at zero local martingale, while the right side is of finite variation. It is therefore sufficient (by proposition 4.12) to show that the right side is predictable. The first summation is predictable (proposition 6.6(b)), and the third summation is continuous and adapted, therefore also predictable. As for the second summation, for every $k \in \mathcal{S}$,

$$I_k \cdot \tilde{V}_k = \left(\sum_{l \in \mathcal{S}} (I_l)_- \right) I_k \cdot \tilde{V}_k = \sum_{l \in \mathcal{S}, l \neq k} (I_l)_- I_k \cdot \tilde{V}_k + (I_k)_- I_k \cdot \tilde{V}_k.$$

The last term here is predictable (proposition 6.6(c)), and by the definition of $\tilde{V}_k'$ and proposition 6.6(a),

$$\begin{aligned}
\sum_{l \in \mathcal{S}, l \neq k} (I_l)_- I_k \cdot \tilde{V}_k &= \sum_{l \in \mathcal{S}, l \neq k} \sum_{i=0}^{\infty} 1(Z_{i-1} = l) 1(Z_i = k) 1([T_i, \infty))(\Delta \tilde{V}_k)(T_i) \\
&= \sum_{l \in \mathcal{S}, l \neq k} \sum_{i=0}^{\infty} 1(Z_{i-1} = l) 1(Z_i = k) 1([T_i, \infty))(\Delta \tilde{V}_k')(T_i) \\
&= \sum_{l \in \mathcal{S}, l \neq k} (\Delta \tilde{V}_k') \cdot N_{lk} \\
&= 0.
\end{aligned}$$

□

Bibliography

[1] Bremaud, P. (1981). *Point processes and queues, martingale dynamics.* Springer-Verlag, New York, Heidelberg, Berlin.

[2] Cohn, D.L. (1980). *Measure theory.* Birkhäuser, Boston, Basel, Stuttgart.

[3] Dellacherie, C. & Meyer, P.A. (1978). *Probabilities and potential.* North-Holland, Amsterdam, New York, Oxford.

[4] Dellacherie, C. & Meyer, P.A. (1982). *Probabilities and potential B: theory of martingales.* North-Holland, Amsterdam, New York, Oxford.

[5] Haberman, S. & Sibbett, T.A. (Editors). (1995). *History of actuarial science: volume I.* William Pickering, London.

[6] Haberman, S. & Sibbett, T.A. (Editors). (1995). *History of actuarial science: volume IV.* William Pickering, London.

[7] Norberg, R. (1996). Addendum to Hattendorff's theorem and Thiele's differential equation generalized. *Scandinavian Actuarial Journal.* **1996**, 50-53.

[8] Norberg, R. (1992). Hattendorff's theorem and Thiele's differential equation generalized. *Scandinavian Actuarial Journal.* **1992**, 2-14.

[9] Norberg, R. (1990). Payment measures, interest, and discounting — an axiomatic approach with applications to insurance. *Scandinavian Actuarial Journal.* **1990**, 14-33.

[10] Protter, P. (1992). *Stochastic integration and differential equations: a new approach.* Springer-Verlag, New York, Berlin, Heidelberg.